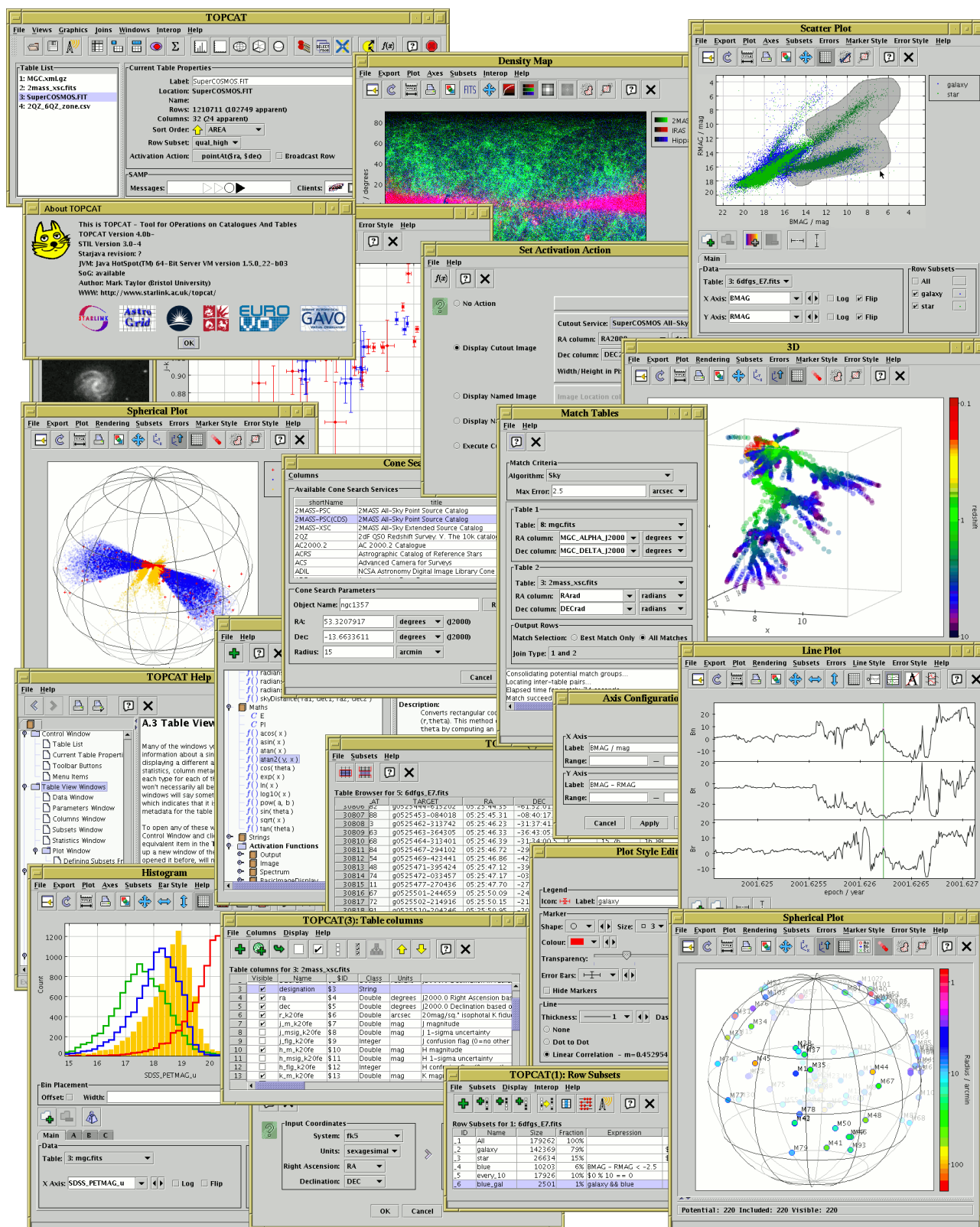


# TOPCAT - Tool for Operations on Catalogues And Tables

## Version 4.3-5



Starlink User Note253  
Mark Taylor  
23 September 2016

## Abstract

TOPCAT is an interactive graphical viewer and editor for tabular data. It has been designed for use

with astronomical tables such as object catalogues, but is not restricted to astronomical applications. It understands a number of different astronomically important formats, and more formats can be added. It is designed to cope well with large tables; a million rows by a hundred columns should not present a problem even with modest memory and CPU resources.

It offers a variety of ways to view and analyse the data, including a browser for the cell data themselves, viewers for information about table and column metadata, tools for joining tables using flexible matching algorithms, and extensive 2- and 3-d visualisation facilities. Using a powerful and extensible Java-based expression language new columns can be defined and row subsets selected for separate analysis. Selecting a row can be configured to trigger an action, for instance displaying an image of the catalogue object in an external viewer. Table data and metadata can be edited and the resulting modified table can be written out in a wide range of output formats.

A number of options are provided for loading data from external sources, including Virtual Observatory (VO) services, thus providing a gateway to many remote archives of astronomical data. It can also interoperate with other desktop tools using the SAMP protocol.

TOPCAT is written in pure Java and is available under the GNU General Public Licence. Its underlying table processing facilities are provided by STIL, the Starlink Tables Infrastructure Library.

## Contents

<b>Abstract.....</b>	<b>1</b>
<b>1 Introduction.....</b>	<b>9</b>
<b>2 Quick Start Guide.....</b>	<b>11</b>
<b>3 Apparent Table.....</b>	<b>13</b>
3.1 Row Subsets.....	13
3.1.1 Defining Subsets.....	14
3.2 Row Order.....	15
3.3 Column Set.....	15
<b>4 Table I/O.....</b>	<b>17</b>
4.1 Table Formats.....	17
4.1.1 Supported Input Formats.....	17
4.1.1.1 FITS.....	17
4.1.1.2 Column-oriented FITS.....	18
4.1.1.3 VOTable.....	18
4.1.1.4 CDF.....	19
4.1.1.5 ASCII.....	19
4.1.1.6 IPAC.....	20
4.1.1.7 Comma-Separated Values.....	21
4.1.1.8 GBIN.....	21
4.1.1.9 Tab-Separated Table.....	22
4.1.1.10 SQL Database Queries.....	23
4.1.1.11 World Data Center.....	23
4.1.2 Supported Output Formats.....	24
4.1.2.1 FITS.....	24
4.1.2.2 Column-oriented FITS.....	24
4.1.2.3 VOTable.....	25
4.1.2.4 ASCII.....	25
4.1.2.5 Text.....	26
4.1.2.6 Comma-Separated Values.....	26
4.1.2.7 IPAC.....	26
4.1.2.8 Tab-Separated Table.....	26
4.1.2.9 SQL Tables.....	26
4.1.2.10 HTML.....	27

4.1.2.11 LaTeX.....	27
4.1.2.12 Mirage Format.....	27
4.1.3 Custom I/O Formats.....	28
4.2 Table Locations.....	28
<b>5 Joins and Matches.....</b>	<b>30</b>
5.1 Concatenating Tables.....	30
5.2 Matching Rows Between Tables.....	30
5.3 Matching Rows Within a Table.....	32
5.4 Multi-Object Matches.....	32
5.5 Plotting Match Results.....	32
5.6 Notes on Matching.....	35
5.7 Matching against a Remote Table.....	35
<b>6 Virtual Observatory Access.....</b>	<b>36</b>
6.1 The Registry.....	36
6.2 Data Access Services.....	36
<b>7 Algebraic Expression Syntax.....</b>	<b>38</b>
7.1 Referencing Cell Values.....	38
7.2 Referencing Row Subset Flags.....	39
7.3 Referencing Table Parameters.....	40
7.4 Null Values.....	40
7.5 Operators.....	41
7.6 Functions.....	42
7.6.1 Technical Note.....	45
7.7 Instance Methods.....	45
7.8 Examples.....	45
7.9 Adding User-Defined Functions.....	47
<b>8 Activation Actions.....</b>	<b>49</b>
8.1 Activation Functions.....	49
<b>9 Tool Interoperability.....</b>	<b>51</b>
9.1 SAMP control.....	51
9.2 PLASTIC control.....	52
9.3 Messages Transmitted.....	53
9.4 Messages Received.....	55
<b>10 Invoking TOPCAT.....</b>	<b>57</b>
10.1 TOPCAT Command-line Arguments.....	57
10.2 Java Options.....	60
10.2.1 Class Path.....	60
10.2.2 Memory Size.....	61
10.2.3 System properties.....	61
10.3 JDBC Configuration.....	64
10.4 Tips for Large Tables.....	66
10.5 Examples.....	67
<b>Appendix A: TOPCAT Windows.....</b>	<b>69</b>
<b>A.1 Common Window Features.....</b>	<b>69</b>
A.1.1 Toolbar.....	69
A.1.2 Menus.....	69
A.1.3 JTables.....	71
A.1.4 Column Selector.....	72
<b>A.2 Control Window.....</b>	<b>72</b>
A.2.1 Table List.....	73
A.2.2 Current Table Properties panel.....	74
A.2.3 Memory Monitor.....	75
A.2.4 SAMP Panel.....	75
A.2.5 Toolbar Buttons.....	76
A.2.6 Menu Items.....	78

<b>A.3 Table View Windows.....</b>	<b>80</b>
A.3.1 Data Window.....	81
A.3.2 Parameters Window.....	83
A.3.3 Columns Window.....	84
A.3.4 Subsets Window.....	87
A.3.5 Statistics Window.....	90
<b>A.4 Plot Windows.....</b>	<b>93</b>
A.4.1 Differences From Old-Style Plot Windows.....	93
A.4.2 Window Overview.....	94
A.4.2.1 Navigation.....	100
A.4.2.2 Table Data Layer Controls.....	101
A.4.2.3 Defining Subsets by Region.....	103
A.4.2.4 Plot Export Window.....	106
A.4.3 Fixed Controls.....	108
A.4.3.1 Frame Control.....	108
A.4.3.2 Axes Control.....	109
A.4.3.3 Legend Control.....	109
A.4.3.4 Aux Axis Control.....	111
A.4.4 Layer Controls.....	113
A.4.4.1 Position Layer Control.....	113
A.4.4.2 Pair Position Layer Control.....	115
A.4.4.3 Healpix Layer Control.....	117
A.4.4.4 Histogram Layer Control.....	120
A.4.4.5 Spectrogram Layer Control.....	122
A.4.4.6 Function Layer Control.....	123
A.4.5 Plot Forms.....	126
A.4.5.1 Mark Form.....	126
A.4.5.2 Size Form.....	127
A.4.5.3 SizeXY Form.....	128
A.4.5.4 Vector Form.....	129
A.4.5.5 SkyVector Form.....	130
A.4.5.6 Error Bars Form.....	132
A.4.5.7 Ellipse Form.....	132
A.4.5.8 SkyEllipse Form.....	134
A.4.5.9 Line Form.....	135
A.4.5.10 Linear Fit Form.....	135
A.4.5.11 Text Label Form.....	137
A.4.5.12 Contour Form.....	138
A.4.5.13 Density Form.....	139
A.4.5.14 SkyDensity Form.....	140
A.4.5.15 Fill Form.....	141
A.4.5.16 Histogram Form.....	142
A.4.5.17 KDE Form.....	143
A.4.5.18 KNN Form.....	145
A.4.5.19 Densogram Form.....	148
A.4.5.20 Pair Marks Form.....	150
A.4.5.21 Pair Link Form.....	150
A.4.6 Shading Modes.....	151
A.4.6.1 Flat Mode.....	151
A.4.6.2 Translucent Mode.....	151
A.4.6.3 Transparent Mode.....	152
A.4.6.4 Auto Mode.....	152
A.4.6.5 Density Mode.....	153
A.4.6.6 Aux Mode.....	154
A.4.6.7 Weighted Mode.....	155

A.4.7 Histogram Plot Window.....	156
A.4.7.1 Histogram Navigation.....	157
A.4.7.2 Histogram Axes Control.....	158
A.4.7.3 Bins Control.....	158
A.4.8 Plane Plot Window.....	160
A.4.8.1 Plane Navigation.....	162
A.4.8.2 Plane Axes Control.....	163
A.4.9 Sky Plot Window.....	167
A.4.9.1 Sky Navigation.....	169
A.4.9.2 Sky Axes Control.....	170
A.4.10 Cube Plot Window.....	173
A.4.10.1 Cube Navigation.....	175
A.4.10.2 Cube Axes Control.....	176
A.4.11 Sphere Plot Window.....	182
A.4.11.1 Sphere Navigation.....	184
A.4.11.2 Sphere Axes Control.....	185
A.4.12 Time Plot Window.....	188
A.4.12.1 Plot Zones.....	189
A.4.12.2 Time Navigation.....	190
A.4.12.3 Time Axes Control.....	191
<b>A.5 Old-Style Plot Windows.....</b>	<b>194</b>
A.5.1 Common Features.....	195
A.5.1.1 Dataset Selectors.....	195
A.5.1.2 Axis Configuration and Zooming.....	197
A.5.1.3 Error Bars.....	199
A.5.1.4 Point Labels.....	201
A.5.1.5 Auxiliary Axes.....	203
A.5.1.6 Defining Subsets by Region.....	205
A.5.1.7 Exporting Graphics.....	207
A.5.2 Histogram (old-style).....	207
A.5.2.1 Histogram Style Editor.....	211
A.5.3 2D Plot (old-style).....	212
A.5.3.1 Plot Style Editor.....	215
A.5.4 Stacked Line Plot (old-style).....	218
A.5.4.1 Lines Style Editor.....	222
A.5.5 3D Plot (old-style).....	223
A.5.5.1 3D Plot Style Editor.....	227
A.5.6 Spherical Plot (old-style).....	229
A.5.7 Density Map (old-style).....	233
A.5.7.1 Density Style Editor.....	237
<b>A.6 Load Window.....</b>	<b>238</b>
A.6.1 Filestore Browser.....	240
A.6.2 System Browser.....	242
A.6.3 Hierarchy Browser.....	242
A.6.3.1 Navigation.....	244
A.6.3.2 Table Searches.....	245
A.6.4 SQL Query.....	245
A.6.5 Cone Search.....	246
A.6.6 SIA Query.....	246
A.6.7 SSA Query.....	246
A.6.8 TAP Query.....	246
A.6.9 Vizier Catalogue Service Query.....	246
A.6.10 GAVO Millennium Database Query.....	251
A.6.11 BaSTI Theory Database Query.....	253
A.6.12 Example Tables.....	255

<b>A.7 Save Window</b>	<b>255</b>
A.7.1 Content Panel	257
A.7.1.1 Current Table	257
A.7.1.2 Multiple Tables	257
A.7.1.3 Session	257
A.7.2 Destination Panel	258
A.7.2.1 Enter Location	258
A.7.2.2 Filestore Browser	258
A.7.2.3 System Browser	259
A.7.2.4 SQL Output Dialogue	260
<b>A.8 Match Windows</b>	<b>261</b>
A.8.1 Common Features	261
A.8.1.1 Match Criteria	261
A.8.1.2 Column Selection Boxes	264
A.8.1.3 Tuning	264
A.8.2 Pair Match Window	265
A.8.2.1 Output Rows Selector Box	267
A.8.3 Internal Match Window	270
A.8.3.1 Internal Match Action Box	272
A.8.4 Multiple Match Window	273
A.8.4.1 Multiple Match Action Box	274
<b>A.9 VO Data Access Windows</b>	<b>274</b>
A.9.1 Common Features	274
A.9.1.1 Registry Query Panel	274
A.9.1.2 Single Positional Search Panel	277
A.9.1.3 Multiple Positional Search Panel	278
A.9.2 Cone Search	280
A.9.3 Simple Image Access (SIA) Query	282
A.9.4 Simple Spectral Access (SSA) Query	283
A.9.5 Multiple Cone Search	284
A.9.6 Multiple SIA Query	286
A.9.7 Multiple SSA Query	288
A.9.8 Table Access Protocol (TAP) Query	290
A.9.8.1 Select Service tab	294
A.9.8.2 Use Service tab	296
A.9.8.3 Running Jobs tab	302
A.9.8.4 Resume Job tab	305
<b>A.10 Other Windows</b>	<b>306</b>
A.10.1 CDS Upload X-Match Window	306
A.10.2 Concatenation Window	309
A.10.3 Activation Window	310
A.10.3.1 Image Viewer Applications	313
A.10.3.2 Spectrum Viewers	315
A.10.3.3 Web Browsers	316
A.10.4 SAMP Window	317
A.10.5 Help Window	320
A.10.6 New Parameter Window	321
A.10.7 Synthetic Column Window	322
A.10.8 Sky Coordinates Window	324
A.10.9 Algebraic Subset Window	325
A.10.10 Column Classification Window	326
A.10.11 Available Functions Window	328
A.10.12 Log Window	329
<b>Appendix B: Algebraic Functions</b>	<b>331</b>
<b>B.1 General Functions</b>	<b>331</b>

B.1.1 Tilings.....	331
B.1.2 Arithmetic.....	332
B.1.3 Conversions.....	333
B.1.4 Distances.....	334
B.1.5 KCorrections.....	335
B.1.6 Times.....	338
B.1.7 TrigDegrees.....	339
B.1.8 Maths.....	340
B.1.9 Arrays.....	341
B.1.10 Fluxes.....	343
B.1.11 Strings.....	345
B.1.12 Formats.....	346
B.1.13 CoordsRadians.....	346
B.1.14 Coverage.....	348
B.1.15 Lists.....	349
B.1.16 CoordsDegrees.....	349
<b>B.2 Activation Functions.....</b>	<b>350</b>
B.2.1 Browsers.....	350
B.2.2 Mgc.....	351
B.2.3 System.....	351
B.2.4 Sdss.....	351
B.2.5 SuperCosmos.....	352
B.2.6 BasicImageDisplay.....	353
B.2.7 TwoQZ.....	353
B.2.8 Image.....	353
B.2.9 Output.....	354
B.2.10 Sog.....	354
<b>Appendix C: Release Notes.....</b>	<b>355</b>
<b>C.1 Acknowledgements.....</b>	<b>355</b>
<b>C.2 Version History.....</b>	<b>357</b>





## 1 Introduction


TOPCAT is an interactive graphical program which can examine, analyse, combine, edit and write out tables. A table is, roughly, something with columns and rows; each column contains objects of the same type (for instance floating point numbers) and each row has an entry for each of the columns (though some entries might be blank). A common astronomical example of a table is an object catalogue.

TOPCAT can read in tables in a number of formats from various sources, allow you to inspect and manipulate them in various ways, and if you have edited them optionally write them out in the modified state for later use, again in a variety of formats. Here is a summary of its main capabilities:

- View/edit table data in a scrollable browser
- View/edit table metadata (parameters)
- View/edit column metadata (column names, units, UCDs...)
- Re-order and hide/reveal columns
- Insert 'synthetic' columns defined by algebraic expression
- Sort rows on the values in a given column
- Define row subsets in various ways, including algebraically and graphically
- Plot columns against each other in 1, 2 and 3 dimensions, distinguishing different subsets
- Calculate statistics on each column for some or all rows
- Trigger a configurable action (e.g. object image display) when a column is selected
- Perform flexible matching of rows in the same or different tables
- Concatenate the rows of existing tables to create new ones
- Acquire tables from web services, external filestores or other customisable sources
- Perform single and multiple queries to Virtual Observatory services including TAP services
- Interoperate with other desktop tools using SAMP
- Write modified tables out in original or different format

Considerable effort has gone into making it work with large tables; a few million rows and hundreds of columns is usually quite manageable.

The general idea of the program is quite straightforward. At any time, it has a list of tables it knows about - these are displayed in the Control Window which is the first thing you see when you start up the program. You can add to the list by loading tables in, or by some actions which create new tables from the existing ones. When you select a table in the list by clicking on it, you can see general information about it in the control window, and you can also open more specialised view windows which allow you to inspect it in more detail or edit it. Some of the actions you can take, such as changing the current Sort Order, Row Subset or Column Set change the Apparent Table (Section 3), which is a view of the table used for things such as saving it and performing row matches. Changes that you make do not directly modify the tables on disk (or wherever they came from), but if you want to save the changes you have made, you can write the modified table(s) to a new location.

The main body of this document explains these ideas and capabilities in more detail, and Appendix A gives a full description of all the windows which form the application. While the program is running, this document is available via the online help system - clicking the **Help** () toolbar

button in any window will pop up a help browser open at the page which describes that window. This document is heavily hyperlinked, so you may find it easier to read in its HTML form than on paper.

Recent news about the program can be found on the TOPCAT web page (<http://www.starlink.ac.uk/topcat/>). It was initially developed within the now-terminated Starlink

and then AstroGrid projects, and has subsequently been supported by the UK's PPARC and STFC research councils, various Euro-VO and FP7 projects, GAVO and ESA. The underlying table handling facilities are supplied by the Starlink Tables Infrastructure Library STIL (<http://www.starlink.ac.uk/stil/>), which is documented more fully in SUN/252. The software is written in pure Java, and should run on any J2SE platform version 1.6 or later. This makes it highly portable, since it can run on any machine which has a suitable Java installation, which is available for MS Windows, Mac OS X and most flavours of Unix amongst others. Some of the external viewer applications it talks to rely on non-Java code however so one or two facilities, such as displaying spectra, may be absent in some cases.

The TOPCAT application is available under the terms of the GNU General Public License, since some of the libraries it uses are also GPL. However, all of the original code and many of the libraries it uses may alternatively be used under more permissive licenses such as the GNU Lesser General Public License, see documentation of the STILTS package for more details.

## 2 Quick Start Guide

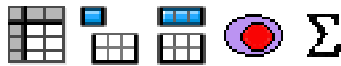
This manual aims to give detailed tutorial and reference documentation on most aspects of TOPCAT's capabilities, and reading it is an excellent way to learn about the program. However, it's quite a fat document, and if you feel you've got better things to do with your time than read it all, you should be able to do most things by playing around with the software and dipping into the manual (or equivalently the online help) when you can't see how to do something or the program isn't behaving as expected. This section provides a short introduction for the impatient, explaining how to get started.

To start the program, you will probably type `topcat` or something like `java -jar topcat-full.jar` (see Section 10 for more detail). To view a table that you have on disk, you can either give its name on the command line or load it using the **Load** button from the GUI. FITS, VOTable, CDF and GBIN files are recognised automatically; if your data is in another format such as ASCII (see Section 4.1.1) you need to tell the program (e.g. `-f ascii` on the command line). If you just want to try the program out, `topcat -demo` will start with a couple of small tables for demonstration purposes.

The first thing that you see is the Control Window (Appendix A.2). This has a list of the loaded table(s) on the left. If one of these is highlighted by clicking on it, information about it will be shown on the right; some of this (table name, sort order) you can change here. Along the top is a toolbar with a number of buttons, most of which open up new windows. These fall into a few groups:



Load/Save/Send Table(s).



Display various aspects of information about the table's data and metadata.




Open plotting/visualisation windows of various kinds.



Join tables in various ways including spatial crossmatching, and access remote databases.



Help and information

The **Help** (  ) button appears in most windows - if you click it a help browser will be displayed

showing an appropriate part of this manual. As well as the tool bar there are a number of menus along the top - some of the options just repeat those appearing on the toolbar, but several less common ones are available as well. The **Help** menu gives you a few more options along the same lines, including displaying the help information in your usual web browser rather than in TOPCAT's (somewhat scrappy) help viewer. All the windows follow roughly this pattern. For some of the toolbar buttons you can probably guess what they do from their icons, for others probably not - to find out you can hover with the mouse to see the tooltip, look in the menus, read the manual, or just push it and see.

Some of the windows allow you to make changes of various sorts to the tables, such as performing sorts, selecting rows, modifying data or metadata. None of these affect the table on disk (or

database, or wherever), but if you subsequently save the table the changes will be reflected in the table that you save.

A notable point to bear in mind concerns memory. TOPCAT is fairly efficient in use of memory, but in some cases when dealing with large tables you might see an `OutOfMemoryError`. It is usually possible to work round this by using the `-Xmx NNN M` flag on startup - see Section 10.2.2.

Finally, if you have queries, comments or requests about the software, and they don't appear to be addressed in the manual, consult the TOPCAT web page, use the `topcat-user` mailing list, or contact the author - user feedback is always welcome.

### 3 Apparent Table

The **Apparent Table** is a particular view of a table which can be influenced by some of the viewing controls.

When you load a table into TOPCAT it has a number of characteristics like the number of columns and rows it contains, the order of the rows that make up the data, the data and metadata themselves, and so on. While manipulating it you can modify the way that the table appears to the program, by changing or adding data or metadata, or changing the order or selection of columns or rows that are visible. For each table its "apparent table" is a table which corresponds to the current state of the table according to the changes that you have made.

In detail, the apparent table consists of the table as it was originally imported into the program plus any of the following changes that you have made:


- Selection of rows changed by changing the current Row Subset (Section 3.1)
- Changes to the current Row Order (Section 3.2) caused by doing a sort
- Changes to the current Column Set (Section 3.3) caused by adding, hiding or moving columns
- Changes to cell data by editing cells in the Data window
- Changes to table metadata by editing cells in the Parameter window
- Changes to column metadata by editing cells in the Columns window

The apparent table is used in the following contexts:

#### Data Window

The Data window always shows the rows and columns of the apparent table, so if you are in doubt about what form a table will get exported in, you can see what it looks like there.

#### Exports

When you save a table, or export it by dragging it off the Table List panel in the Control Window, or create a duplicate table, it is the apparent table which is copied. So for instance if you define a subset containing only the first ten rows of a table and then save it to a new table, or create a duplicate within TOPCAT using the **Duplicate Table** () toolbar button, the

resulting table will contain only those ten rows.

#### Joins

When you use the Match Window or Concatenation Window to construct a new table on the basis of one or more existing ones, the new table will be built on the basis of the apparent versions of the tables being operated on. The same applies to the join-like functionality provided by table uploads in the TAP window, CDS Upload X-Match window and the multiple positional search (Cone, SIA, SSA) windows.

Some of the other table view windows are affected too, for instance the Columns window displays its columns in the order that they appear in the Apparent Table.

### 3.1 Row Subsets

An important feature of TOPCAT is the ability to define and use **Row Subsets**. A Row Subset is a selection of the rows within a whole table being viewed within the application, or equivalently a new table composed from some subset of its rows. You can define these and use them in several different ways; the usefulness comes from defining them in one context and using them in another. The Subset Window displays the currently defined Row Subsets and permits some operations on them.

At any time each table has a **current** row subset, and this affects the Apparent Table. You can



always see what it is by looking at the "Row Subset" selector in the Control Window when that table is selected; by default it is one containing all the rows. You can change it by choosing from this selector or as a result of some other actions.

Other contexts in which subsets can be used are picking a selection of rows from which to calculate in the Statistics Window and marking groups of rows to plot using different markers in the various plotting (and old-style plotting) windows.

### 3.1.1 Defining Subsets

You can define a Row Subset in one of the following ways:

#### Selecting rows in the browser

You can select a single row in the Data Window by clicking on it, or select a group of adjacent rows by dragging the mouse over them. You can add more rows to the selection by keeping the <Control> button pressed while you do it. Once you have a set of rows selected you can use the **Subset From Selected Rows** () or **Subset From Unselected Rows** () buttons to

create a new subset based on the set of highlighted rows or their complement.



Combining this with sorting the rows in the table can be useful; if you do a Sort Up on a given column and then drag out the top few rows of the table you can easily create a subset consisting of the highest values of a given column.

#### Defining an algebraic expression

From the Subset Window using the **Add New Subset** () button will pop up the Algebraic


Subset Window which allows you to define a new subset using an algebraic expression based on the values of the cells in each row. The format of such expressions is described in Section 7.

#### Visible plotted points

In some of the plotting (and old-style plotting) windows you can plot columns against each other, and subsequently zoom in and out using the mouse. If you zoom to display only some of the plotted points and then use the **New Subset From Visible** () or () button then a

new subset will be created containing only rows represented by points in the field of view of the plot at the time.

#### Selected plotted points

For more control over which plotted points are to be included in a subset, you can use the **Draw Subset Region** () button in some of the plotting windows. This allows you to trace

out with the mouse a region or regions of any shape, creating a new subset containing only those rows represented by the points within those regions.

#### Boolean columns

Any column which has a boolean (true/false) type value can be used as a subset; rows in which it has a true value are in the subset and others are not. Any boolean column in a table is made available as a row subset with the same name when the table is imported.

In all these cases you will be asked to assign a name for the subset. As with column names, it is a good idea to follow a few rules for these names so that they can be used in algebraic expressions. They should be:

- Distinct from other subset and column names, even apart from upper/lower case distinctions
- In the form of a java identifier (starts with a letter, continues with alphanumerics including underscore, no spaces)


- Not too long

When you choose a name, you can either type one in, or select one from the drop-down list, which gives the names of all the existing subsets. This allows you to redefine existing subsets. Note if you do select or type in one of the existing names, any previous content of that subset will be lost.

In the first subset definition method above, the **current** subset will be set immediately to the newly created one. In other cases the new subset may be highlighted appropriately in other windows, for instance by being plotted in scatter plot windows.

### 3.2 Row Order

You can sort the rows of each table according to the values in a selected column. Normally you will want to sort on a numeric column, but other values may be sortable too, for instance a String column will sort alphabetically. Some kinds of columns (e.g. array ones) don't have any well-defined order, and it is not possible to select these for sorting on.

At any time, each table has a **current** row order, and this affects the Apparent Table. You can always see what it is by looking under the "Sort Order" item in the Control Window when that table is selected; by default it is "(none)", which means the rows have the same order as that of the table they were loaded in from. The little arrow () indicates whether the sense of the sort is up or down. You can change the sort order by selecting a column name from this control, and change the sense by clicking on the arrow. The sort order can also be changed by using menu items in the Columns Window or right-clicking popup menus in the Data Window.

Selecting a column to sort by calculates the new row order by performing a sort on the cell values there and then. If the table data change somehow (e.g. because you edit cells in the table) then it is possible for the sort order to become out of date.

The current row order affects the Apparent Table, and hence determines the order of rows in tables which are exported in any way (e.g. written out) from TOPCAT. You can always see the rows in their currently sorted order in the Data Window.

### 3.3 Column Set





When each table is imported it has a list of columns. Each column has header information which determines the kind of data which can fill the cells of that column as well as a name, and maybe some additional information like units and Unified Content Descriptor. All this information can be viewed, and in some cases modified, in the Columns Window.

During the lifetime of the table within TOPCAT, this list of columns can be changed by adding new columns, hiding (and perhaps subsequently revealing) existing columns, and changing their order. The current state of which columns are present and visible and what order they are in is collectively known as the **Column Set**, and affects the Apparent Table. The current Column Set is always reflected in the order in which columns are displayed in the Data Window and Statistics Window. The Columns Window shows all the known columns, including hidden ones, in Column Set order; whether they are currently visible is indicated by the (leftmost) "Visible" column.

You can affect the current Column Set in the following ways:

#### Hide/Reveal columns

In the Columns Window you can toggle columns between hidden and visible by clicking on their box in the **Visible** column. To make a group of columns hidden or visible at once, select the corresponding rows (drag the mouse over them to select a contiguous group; hold the Control button down to add more single rows or contiguous groups to the selection) and hit the

 ) or **Reveal Selected** (  ) button in the toolbar or menu. Note when selecting rows, don't drag the mouse over the Visible column, do it somewhere in the middle of the table. The **Hide All** (  ) and **Reveal All** (  ) buttons set all columns in the table invisible or visible

- a useful convenience if you've got a very wide table.

You can also hide a column by right-clicking on it in the Data Window, which brings up a popup menu - select the **Hide** option. To make it visible again you have to go to the Columns Window as above.

### Move Columns


In the Data Window you can move columns around by dragging the grey column header left or right to a new position (as usual in a JTable). Alternatively, you can drag the rows in the Columns Window by grabbing the grey row header (numbered cell at the left) and dragging it up or down. Either of these affects the Column Set, as you can see by looking at one window while moving columns in the other.

### Add Columns

You can use the **New Synthetic Column** (  ) or **New Sky Coordinate Columns** (  )


buttons in the Columns Window or the (right-click) popup menu in the Data Window to add new columns derived from existing ones.

### Replace a Column

If a column is selected in the Columns Window or from the Data Window popup menu you can use the **Replace Column with Synthetic** (  ) button. This is similar to the **Add a**

**Synthetic Column** described in the previous item, but it pops up a new column dialogue with similar characteristics (name, units etc) to those of the column that's being replaced, and when completed it slots the new column in to the table hiding the old one.

### Add a Subset Column

If you have defined a Row Subset somehow and you want it to appear explicitly in the table (for instance so that when you write the table out the selection is saved) you can select that subset in the Subsets Window and use the **To Column** (  ) button, which will add a new

boolean column to the table with the value **true** for rows part of that subset and **false** for the other rows.



## 4 Table I/O

### 4.1 Table Formats

TOPCAT supports a wide variety of tabular data formats. In most cases these are file formats for tables stored as single files on a disk or at the end of a URL, but there are other possibilities, for instance a table you have opened could be the result of an SQL query on a database.

Since you can load a table from one format and save it in a different one, TOPCAT can be used to convert a table from one format to another. If this is all you want to do however, you may find it more convenient to use the `tcopy` command line utility in the STILTS package.

The format handling is extensible, so new formats can be added fairly easily. All the table input/output is handled by STIL, the Starlink Tables Infrastructure Library; more detailed descriptions of the I/O capabilities can be found in its documentation (<http://www.starlink.ac.uk/stil/>).

The following subsections describe the available formats for reading and writing tables. The two operations are separate, so not all the supported input formats have matching output formats and vice versa.

#### 4.1.1 Supported Input Formats

Loading tables into TOPCAT is done either from the command line when you start the program up or using the Load Table dialogue. For FITS, VOTable, CDF and GBIN formats the file format can be detected automatically (note this is done by looking at the file content, it has nothing to do with filename extensions). For other formats though, for instance ASCII or Comma-Separated Values, you will have to specify the format that the file is in. In the Load Window, there is a selection box from which you can choose the format, and from the command line you use the `-f` flag - see Section 10 for details. You can always specify the format rather than using automatic detection if you prefer - this can be a good idea if a table appears to be failing to load in a surprising way, since it may give you a more detailed error message.

In either case, table locations may be given as filenames or as URLs, and any data compression (gzip, unix compress and bzip2) will be automatically detected and dealt with - see Section 4.2.

**Note:** in some earlier versions of TOPCAT, ASCII format tables could be detected automatically, so you could load them by typing something like `"topcat table.txt"`. In the current version, you have to signal that this is an ASCII table, for instance by typing `"topcat -f ascii table.txt"`.

The following sections describe the table formats which TOPCAT can read.

##### 4.1.1.1 FITS

FITS binary and ASCII table extensions can be read. Normally, TOPCAT will load all TABLE or BINTABLE extensions in a given file.

If only a single extension is required from a multi-extension FITS file, this is indicated by giving an identifier after a '#' at the end of the table location. The identifier can be in one of two forms:

- The numeric index of the HDU. The first extension (first HDU after the primary HDU) is numbered 1. Thus in a compressed FITS table named `"spec23.fits.gz"` with one primary HDU and two BINTABLE extensions, you would view the first one using the name `"spec23.fits.gz"` or `"spec23.fits.gz#1"` and the second one using the name `"spec23.fits.gz#2"`. The suffix `"#0"` is never used for a legal FITS file, since the primary HDU cannot contain a table.

- The name of the extension. This is the value of the EXTNAME header in the HDU, or alternatively the value of EXTNAME followed by "-" followed by the value of EXTVER. This follows the recommendation in the FITS standard that EXTNAME and EXTVER headers can be used to identify an HDU. So in a multi-extension FITS file "cat.fits" where a table extension has EXTNAME='UV\_DATA' and EXTVER=1, it could be referenced as "cat.fits#UV\_DATA" or "cat.fits#UV\_DATA-1". Matching of these names is case-insensitive.

You can select which extension to use interactively if you use the Hierarchy Browser to load the table.

If the table has been written using TOPCAT's "fits-plus" output format (see Section 4.1.2.1) then the metadata will be read in from the primary HDU as well.

For normal FITS files, header cards in the table's HDU header will be made available as table parameters (see Appendix A.3.2). Only header cards which are not used to specify the table format itself are visible as parameters (e.g. NAXIS, TTYPE\* etc cards are not). HISTORY and COMMENT cards are run together as one multi-line value.

If the table is stored in a FITS binary table extension in a file on local disk in uncompressed form, then the table is 'mapped' into memory - this generally means fast loading and low memory use.

#### 4.1.1.2 Column-oriented FITS

As well as normal binary and ASCII FITS tables, STIL supports FITS files which contain tabular data stored in column-oriented format. This means that the table is stored in a BINTABLE extension HDU, but that BINTABLE has a single row, with each cell of that row holding a whole column's worth of data. The final (slowest-varying) dimension of each of these cells (declared via the TDIMn header cards) is the same, namely, the number of rows in the table that is represented. The point of this is that all the cells of a given column are stored contiguously, which for very large, and especially very wide tables means that certain access patterns (basically, ones which access only a small proportion of the columns in a table) can be much more efficient since they require less I/O overhead in reading data blocks.

Such tables are perfectly legal FITS files, but most non-STIL software will probably not recognise them as tables in the usual way. This format is mostly intended for the case where you have a large table in some other format (possibly the result of an SQL query) and you wish to cache it in a way which can be read efficiently by a STIL-based application.

Like normal FITS, two variants are supported; with (colfits-plus) and without (colfits-basic) metadata stored as a VOTable byte array in the primary HDU.

For performance reasons, it is advisable to access colfits files uncompressed on disk. Reading them from a remote URL, or in gzipped form, may be rather slow (in earlier versions it was not supported at all).

#### 4.1.1.3 VOTable

VOTable is an XML-based format for tabular data endorsed by the International Virtual Observatory Alliance; while the tabular data which can be encoded is by design close to what FITS allows, it provides for much richer encoding of structure and metadata. TOPCAT is believed to read any table which conforms to the VOTable 1.0, 1.1, 1.2 or 1.3 specifications (<http://www.ivoa.net/Documents/latest/VOT.html>). This includes tables in which the cell data are included in-line as XML elements (VOTable/TABLEDATA format), or included/referenced as a FITS table (VOTable/FITS) or included/referenced as a raw binary stream (VOTable/BINARY or VOTable/BINARY2). TOPCAT does not attempt to be fussy about input VOTable documents, and

it will have a good go at reading VOTables which violate the standards in various ways.

VOTable documents can have a complicated hierarchical structure, and may contain more than one actual table (TABLE element). Normally, TOPCAT will load all tables it finds in the document. If you just want to load a single table, indicate the zero-based index of the TABLE element in a breadth-first search after a '#' character at the end of the table specification. Here is an example VOTable document:

```
<VOTABLE>
  <RESOURCE>
    <TABLE name="Star Catalogue"> ... </TABLE>
    <TABLE name="Galaxy Catalogue"> ... </TABLE>
  </RESOURCE>
</VOTABLE>
```

If this is available in a file named "cats.xml" then loading the document in the usual way from the filestore browser or command line will open two new tables in TOPCAT. If you just want the Star Catalogue or the Galaxy Catalogue you can use the name "cats.xml#0" or "cats.xml#1" respectively.

#### 4.1.1.4 CDF

NASA's Common Data Format, described at <http://cdf.gsfc.nasa.gov/>, is a binary format for storing self-describing data. It is typically used to store tabular data for subject areas like space and solar physics. Since v4.0-1 TOPCAT is capable of reading CDF files, though it can't write them. Currently, no attempt is made to represent values representing times in a human-friendly fashion. This may change in future releases.

#### 4.1.1.5 ASCII

In many cases tables are stored in some sort of unstructured plain text format, with cells separated by spaces or some other delimiters. There is a wide variety of such formats depending on what delimiters are used, how columns are identified, whether blank values are permitted and so on. It is impossible to cope with them all, but TOPCAT attempts to make a good guess about how to interpret a given ASCII file as a table, which in many cases is successful. In particular, if you just have columns of numbers separated by something that looks like spaces, you should be just fine.

Here are the detailed rules for how the ASCII-format tables are interpreted:

- Bytes in the file are interpreted as ASCII characters
- Each table row is represented by a single line of text
- Lines are terminated by one or more contiguous line termination characters: line feed (0x0A) or carriage return (0x0D)
- Within a line, fields are separated by one or more whitespace characters: space (" ") or tab (0x09)
- A field is either an unquoted sequence of non-whitespace characters, or a sequence of non-newline characters between matching single (') or double (") quote characters - spaces are therefore allowed in quoted fields
- Within a quoted field, whitespace characters are permitted and are treated literally
- Within a quoted field, any character preceded by a backslash character ("\") is treated literally. This allows quote characters to appear within a quoted string.
- An empty quoted string (two adjacent quotes) or the string "null" (unquoted) represents the null value
- All data lines must contain the same number of fields (this is the number of columns in the table)
- The data type of a column is guessed according to the fields that appear in the table. If all the fields in one column can be parsed as integers (or null values), then that column will turn into

an integer-type column. The types that are tried, in order of preference, are: Boolean, Short Integer, Long, Float, Double, String

- Some special values are permitted for floating point columns: NaN for not-a-number, which is treated the same as a null value for most purposes, and Infinity or inf for infinity (with or without a preceding +/- sign). These values are matched case-insensitively.
- Empty lines are ignored
- Anything after a hash character "#" (except one in a quoted string) on a line is ignored as far as table data goes; any line which starts with a "!" is also ignored. However, lines which start with a "#" or "!" at the start of the table (before any data lines) will be interpreted as metadata as follows:
  - The last "#"/"!"-starting line before the first data line may contain the column names. If it has the same number of fields as there are columns in the table, each field will be taken to be the title of the corresponding column. Otherwise, it will be taken as a normal comment line.
  - Any comment lines before the first data line not covered by the above will be concatenated to form the "description" parameter of the table.

If the list of rules above looks frightening, don't worry, in many cases it ought to make sense of a table without you having to read the small print. Here is an example of a suitable ASCII-format table:

```
#
# Here is a list of some animals.
#
# RECNO   SPECIES      NAME           LEGS   HEIGHT/m
# 1       pig          "Pigling Bland" 4    0.8
# 2       cow          Daisy          4      2
# 3       goldfish     Dobbin         " "    0.05
# 4       ant          " "            6      0.001
# 5       ant          " "            6      0.001
# 6       ant          ' '            6      0.001
# 7       "queen ant"  'Ma\'am'       6      2e-3
# 8       human        "Mark"         2      1.8
```

In this case it will identify the following columns:

Name	Type
----	----
RECNO	Short
SPECIES	String
NAME	String
LEGS	Short
HEIGHT/m	Float

It will also use the text "Here is a list of some animals" as the Description parameter of the table. Without any of the comment lines, it would still interpret the table, but the columns would be given the names col1..col5.

If you understand the format of your files but they don't exactly match the criteria above, the best thing is probably to write a simple free-standing program or script which will convert them into the format described here. You may find Perl or awk suitable languages for this sort of thing.

This format is not detected automatically - you must specify that you wish to load a table in `ascii` format.

#### 4.1.1.6 IPAC

CalTech's Infrared Processing and Analysis Center use a text-based format for storage of tabular data, defined at [http://irsa.ipac.caltech.edu/applications/DDGEN/Doc/ipac\\_tbl.html](http://irsa.ipac.caltech.edu/applications/DDGEN/Doc/ipac_tbl.html). Tables can store column name, type, units and null values, as well as table parameters. They typically have a

filename extension ".tbl" and are used for Spitzer data amongst other things. An example looks like this:

```
\title='Animals'
\ This is a table with some animals in it.
|  RECNO | SPECIES | NAME | LEGS | HEIGHT |
|  char  |  char   |  char |  int  | double |
|        |         |      |      | m       |
|        |         |      |      |         |
|    1   | pig     | Pigling Bland | 4 | 0.8 |
|    2   | cow     | Daisy | 4 | 2 |
|    3   | goldfish | Dobbin | 0 | 0.05 |
|    4   | ant     | null | 6 | 0.001 |
```

#### 4.1.1.7 Comma-Separated Values

Comma-separated value ("CSV") format is a common semi-standard text-based format in which fields are delimited by commas. Spreadsheets and databases are often able to export data in some variant of it. The intention is that TOPCAT can read tables in the version of the format spoken by MS Excel amongst other applications, though the documentation on which it was based was not obtained directly from Microsoft.

The rules for data which it understands are as follows:

- Each row must have the same number of comma-separated fields.
- Whitespace (space or tab) adjacent to a comma is ignored.
- Adjacent commas, or a comma at the start or end of a line (whitespace apart) indicates a null field.
- Lines are terminated by any sequence of carriage-return or newline characters ('\r' or '\n') (a corollary of this is that blank lines are ignored).
- Cells may be enclosed in double quotes; quoted values may contain linebreaks (or any other character); a double quote character within a quoted value is represented by two adjacent double quotes.
- The first line *may* be a header line containing column names rather than a row of data. Exactly the same syntactic rules are followed for such a row as for data rows.

Note that you can *not* use a "#" character (or anything else) to introduce "comment" lines.

This format is not detected automatically - you must specify that you wish to load a table in `csv` format.

#### 4.1.1.8 GBIN

GBIN format is a special-interest file format used within DPAC, the Data Processing and Analysis Consortium working on data from the Gaia astrometry satellite. It is based on java serialization, and in all of its various forms has the peculiarity that you only stand any chance of decoding it if you have the Gaia data model classes on your java classpath at runtime. Since the set of relevant classes is very large, and also depends on what version of the data model your GBIN file corresponds to, those classes will not be packaged as part of the standard TOPCAT distribution, so some additional setup is required to read GBIN files.

As well as the data model classes, you must provide on the runtime classpath the GaiaTools classes required for GBIN reading. The table input handler accesses these by reflection, partly because GaiaTools is targeted at java 8 while STIL is at time of writing java 6 compatible, and partly to avoid an additional large library dependency for a rather niche requirement. It is likely that since you have to supply the required data model classes you will also have the required GaiaTools classes to hand as well, so this shouldn't constitute much of an additional burden for usage.

In practice, if you have a jar file or files for pretty much any java library or application which is capable of reading a given GBIN file, just adding it or them to TOPCAT's classpath at run time ought to do the trick. Examples of such jar files are the `MDBExplorerStandalone.jar` file available from <http://gaia.esac.esa.int/mdbexp/>, or the `gbcats.jar` file you can build from the `CU9/software/gbcats/` directory in the DPAC subversion repository.

Suppose you have the `MDBExplorerStandalone.jar` file mentioned above, you can do it by starting TOPCAT like this:

```
topcat -classpath MDBExplorerStandalone.jar
```

or like this:

```
java -classpath topcat-full.jar:MDBExplorerStandalone.jar uk.ac.starlink.topcat.Driver
```

Note you will need to be using whatever version of java is required by GaiaTools, probably java 8.

The GBIN format is recognised automatically, so you do not need to set the file type explicitly when loading GBIN files.

The GBIN format doesn't really store tables, it stores arrays of java objects, so the input handler has to make some decisions about how to flatten these into table rows.

In its simplest form, the handler basically looks for public instance methods of the form `getXxx()` and uses the `xxx` as column names. If the corresponding values are themselves objects with suitable getter methods, those objects are added as new columns instead. This more or less follows the practice of the `gaia.cu1.tools.util.GbinInterrogator/gbcats` tool. Method names are sorted alphabetically. Arrays of complex objects are not handled well, and various other things may trip it up. See the source code (e.g. `uk.ac.starlink.gbin.GbinTableProfile`) for more details.

If the object types stored in the GBIN file are known to the special metadata-bearing class `gaia.cu9.tools.documentationexport.MetadataReader` and its dependencies, and if that class is on the runtime classpath, then the handler will be able to extract additional metadata as available, including standardised column names, table and column descriptions, and UCDs. An example of a jar file containing this metadata class alongside data model classes is `GaiaDataLibs-18.3.1-r515078.jar`. Note however at time of writing there are some deficiencies with this metadata extraction functionality related to unresolved issues in the upstream `gaia` class libraries and the relevant interface control document (GAIA-C9-SP-UB-XL-034-01, "External Data Centres ICD"). Currently columns appear in the output table in a more or less random order, units and Utypes are not extracted, and using the GBIN reader tends to cause a 700kbyte file "temp.xml" to be written in the current directory. If the upstream issues are fixed, this behaviour may improve.

Note that support for GBIN files is somewhat experimental. Please contact the author (who is not a GBIN expert) if it doesn't seem to be working properly or you think it should do things differently.

#### 4.1.1.9 Tab-Separated Table

Tab-Separated Table, or TST, is a text-based table format used by a number of astronomical tools including Starlink's GAIA and ESO's SkyCat on which it is based. A definition of the format can be found in Starlink Software Note 75. The implementation here ignores all comment lines: special comments such as the `"#column-units:"` are not processed.

An example looks like this:

```
Simple TST example; stellar photometry catalogue.
A.C. Davenhall (Edinburgh) 26/7/00.
```

Catalogue of U,B,V colours.  
 UBV photometry from Mount Pumpkin Observatory,  
 see Sage, Rosemary and Thyme (1988).

```
# Start of parameter definitions.
EQUINOX: J2000.0
EPOCH: J1996.35
```

```
id_col: -1
ra_col: 0
dec_col: 1
```

```
# End of parameter definitions.
ra<tab>dec<tab>V<tab>B_V<tab>U_B
--<tab>---<tab>---<tab>---<tab>---
5:09:08.7<tab> -8:45:15<tab> 4.27<tab> -0.19<tab> -0.90
5:07:50.9<tab> -5:05:11<tab> 2.79<tab> +0.13<tab> +0.10
5:01:26.3<tab> -7:10:26<tab> 4.81<tab> -0.19<tab> -0.74
5:17:36.3<tab> -6:50:40<tab> 3.60<tab> -0.11<tab> -0.47
[EOD]
```

#### 4.1.1.10 SQL Database Queries

With appropriate configuration, TOPCAT can be used to examine the results of queries on an SQL-compatible relational database.

Database queries can be specified as a string in the form:

```
jdbc:driver-specific-url#sql-query
```

The exact form is dependent on the driver. Here is an example for MySQL:

```
jdbc:mysql://localhost/astrol?user=mbt#SELECT ra, dec FROM swaa WHERE vmag<18
```

which would get a two-column table (the columns being "ra" and "dec"), constructed from certain rows from the table "swaa" in the database "astrol" on the local host, using the access privileges of user mbt.

Fortunately you don't have to construct this by hand, there is an SQL Query Dialogue (Appendix A.6.4) to assist in putting it together.

Note that TOPCAT does not view a table in the database directly, but the result of an SQL query on that table. If you want to view the whole table you can use the query

```
SELECT * FROM table-name
```

but be aware that such a query might be expensive on a large table.

Use of SQL queries requires some additional configuration of TOPCAT; see Section 10.3.

#### 4.1.1.11 World Data Center

Some support is provided for files produced by the World Data Centre for Solar Terrestrial Physics. The format itself apparently has no name, but files in this format look something like the following:

```
Column formats and units - (Fixed format columns which are single space seperated.)
-----
Datetime (YYYY mm dd HHMMSS)                %4d %2d %2d %6d      -
                                                %1s
aa index - 3-HOURLY (Provisional)              %3d                  nT

2000 01 01 000000 67
2000 01 01 030000 32
```

...

Support for WDC tables is experimental (*does the format still exist?*) - it may not be very robust.

This format is not detected automatically - you must specify that you wish to load a table in `wdc` format.

#### 4.1.2 Supported Output Formats

Writing out tables from TOPCAT is done using the Save Table Window. In general you have to specify the format in which you want the table to be output by selecting from the Save Window's **Table Output Format** selector; the following sections describe the possible choices. In some cases there are variants within each format - these are described as well.

The program has no "native" file format, but if you have no particular preference about which format to save tables to, FITS is a good choice. Uncompressed FITS tables do not in most cases have to be read all the way through (they are 'mapped' into memory), which makes them very fast to load up. The FITS format which is written by default (also known as "FITS-plus") also uses a trick to store extra metadata, such as table parameters and UCDs in a way TOPCAT can read in again later (see Section 4.1.2.1). These files are quite usable as normal FITS tables by other applications, but they will only be able to see the limited metadata stored in the FITS headers. For very large files, in some circumstances column-oriented FITS ("`colfits`") format can be more efficient for some applications, though this is unlikely to be understood except by STIL-based code (TOPCAT and STILTS). If you want to write to a format which retains all metadata in a portable format, then one of the Section 4.1.2.3 formats might be better.

##### 4.1.2.1 FITS

When saving in FITS format a new file is written consisting of two HDUs (Header+Data Units): a primary one (required by the FITS standard), and a single extension of type BINTABLE containing the table data.

There are two variants of this format:

###### **fits-basic**

The primary HDU contains only very minimal headers and no data.

###### **fits-plus**

The primary HDU contains an array of bytes which stores the full table metadata as the text of a VOTable document, along with headers that mark this has been done. Most FITS table readers will ignore this altogether and treat the file just as if it contained only the table. When TOPCAT (or other STIL-based applications) read it however, they read out the metadata and make it available for use. In this way you can store your data in the efficient and widely portable FITS format without losing the additional metadata such as table parameters, column UCDs, lengthy column descriptions etc that may be attached to the table. Other, more standard schemes exist for combining the benefits of FITS and VOTable, but suffer from some disadvantages: `votable-fits-inline` is hard to process efficiently (in particular the data cannot easily be mapped into memory) and `votable-fits-href` requires that you keep your data in two separate files, which can get separated from each other. If you want to ensure that the metadata are available to other VOTable-aware programs, you should use one of the normal VOTable formats (Section 4.1.2.3).

In general, you can just let TOPCAT detect the format automatically and not worry about which of these variants is being used - if `fits-plus` is being used you just get some hidden benefits.

##### 4.1.2.2 Column-oriented FITS



When saving in column-oriented FITS format a new file is written consisting of two HDUs (Header+Data Units); a primary one (required by the FITS standard) and a single extension of type BINTABLE containing the table data. Unlike normal FITS format however, this table consists of a single row in which each cell holds the data for an entire column. This can be a more efficient format to work with when dealing with very large, and especially very wide, tables. The benefits are greatest when the file size exceeds the amount of available physical memory and operations are required which scan through the table using only a few of the columns (much of TOPCAT's operations, for instance plotting two columns against each other, fit into this category). The overhead for reading and writing this format is somewhat higher than for normal FITS however, and other applications may not be able to work with it (though it is a legal FITS file), so in most cases normal FITS is a more suitable choice.

Like normal (row-oriented) FITS (see Section 4.1.2.1), there are two variants:

**colfits-plus**

The primary HDU contains an array of bytes which stores the table metadata in VOTable format.

**colfits-basic**

The primary HDU contains no data.

#### 4.1.2.3 VOTable

When a table is saved to VOTable format, a document conforming to the VOTable specification containing a single TABLE element within a single RESOURCE element is written. The version of the VOTable specification according to which the output is written can be controlled using the `votable.version` system property.

There are a number of variants which determine the form in which the table data (DATA element) is written:

**votable-tabledata**

TABLEDATA element (pure XML)

**votable-binary-inline**

BINARY element containing base64-encoded data within the document

**votable-fits-href**

FITS element containing a reference to an external newly-written FITS file (with a name derived from that of the VOTable document)

**votable-binary-href**

BINARY element containing a reference to an external newly-written binary file (with a name derived from that of the VOTable document)

**votable-fits-inline**

FITS element containing base64-encoded data within the document

See the VOTable specification (<http://www.ivoa.net/Documents/latest/VOT.html>) for more explanation of what these variants mean. They can all be read by the VOTable input handler.

If VOTable 1.3 output is in force, the `*-binary-*` options will be replaced by `*-binary2-*`.

#### 4.1.2.4 ASCII

Tables can be written using a format which is compatible with the ASCII input format. It writes as plainly as possible, so should stand a good chance of being comprehensible to other programs

which require some sort of plain text rendition of a table.

The first line is a comment (starting with a "#" character) which names the columns, and an attempt is made to line up data in columns using spaces. Here is an example of a short table written in this format:

```
# index Species Name Legs Height Mammal
1      pig      Bland  4    0.8    true
2      cow      Daisy  4    2.0    true
3     goldfish Dobbin  0    0.05   false
4       ant      " "    6    0.0010 false
5       ant      " "    6    0.0010 false
6     human     Mark  2    1.9    true
```

#### 4.1.2.5 Text

Tables can be written to a simple text-based format which is designed to be read by humans. No reader exists for this format.

Here is an example of a short table written in this format:

index	Species	Name	Legs	Height	Mammal
1	pig	Bland	4	0.8	true
2	cow	Daisy	4	2.0	true
3	goldfish	Dobbin	0	0.05	false
4	ant	" "	6	0.0010	false
5	ant	" "	6	0.0010	false
6	human	Mark	2	1.9	true

#### 4.1.2.6 Comma-Separated Values

Tables can be written to the semi-standard comma-separated value (CSV) format, described in more detail in Section 4.1.1.7. This can be useful for importing into certain external applications, such as some spreadsheets or databases.

There are two variants:

##### CSV

The first line is a header which contains the column names.

##### CSV-noheader

No header line is emitted, all lines represent data rows.

#### 4.1.2.7 IPAC

Tables can be written to the IPAC data format, described in more detail in Section 4.1.1.6.

#### 4.1.2.8 Tab-Separated Table

Tables can be written to TST format, which is described in more detail in Section 4.1.1.9. This can be useful for communicating with some other astronomical tools such as GAIA.

#### 4.1.2.9 SQL Tables

With appropriate configuration, TOPCAT can write out tables as new tables in an SQL-compatible relational database.

For writing, the location is specified as the following URL:

```
jdbc:driver-specific-url#new-table-name
```

The exact form is dependent on the driver. Here is an example for MySQL:

```
jdbc:mysql://localhost/astrol?user=mbt#newtab
```

which would write the current contents of the browser into a new table named "newtab" in the database "astrol" on the local host with the access privileges of user mbt.

Fortunately you do not have to construct this URL by hand, there is an SQL dialogue box to assist in putting it together.

Use of SQL queries requires some additional configuration of TOPCAT; see Section 10.3.

#### 4.1.2.10 HTML

A table can be written out as an HTML 4.01 TABLE element, suitable for use as a web page or insertion into one.

There are two variants:

##### **HTML**

A freestanding HTML document, complete with HTML, HEAD and BODY tags is output.

##### **HTML-element**

Only the TABLE element representing the table is output; this should normally be embedded in a larger HTML document before use.

#### 4.1.2.11 LaTeX

A table can be written out as a LaTeX `tabular` environment, suitable for insertion into a document intended for publication.

There are two variants:

##### **LaTeX**

The `tabular` element alone is output; this will have to be embedded in a larger LaTeX document before use.

##### **LaTeX-document**

A freestanding LaTeX document, consisting of the `tabular` within a `table` within a document is output.

Obviously, this isn't so suitable for very large tables.

#### 4.1.2.12 Mirage Format

Mirage (<http://www.bell-labs.com/project/mirage/index.html>) is a powerful standalone java tool developed at Bell Labs for analysis of multidimensional data. It uses its own file format for input. TOPCAT can write tables in the input format which Mirage uses, so that you can prepare tables in TOPCAT and write them out for subsequent use by Mirage.

It is also possible in principle to launch Mirage directly from within TOPCAT, using the **Export To Mirage** item on the Control Window's **File** menu; this will cause Mirage to start up viewing the currently selected Apparent Table. In order for this to work the Mirage classes must be on your classpath (see Section 10.2.1) when TOPCAT is run.

There appears to be a bug in Mirage which means this does not always work - sometimes Mirage starts up with no data loaded into it. In this case you will have to save the data to disk in Mirage format, start up Mirage separately, and load the data in using the **New Dataset** item in Mirage's **Console** menu.

Note that when Mirage has been launched from TOPCAT, exiting Mirage or closing its window will exit TOPCAT as well.

### 4.1.3 Custom I/O Formats

It is in principle possible to configure TOPCAT to work with table file formats other than the ones listed in this section. It does not require any upgrade of TOPCAT itself, but you have to write or otherwise acquire an input and/or output handler for the table format in question.

The steps that you need to take are:

1. Write java classes which constitute your input and/or output handler
2. Ensure that these classes are available on your classpath while TOPCAT is running (see Section 10.2.1)
3. Set the `startable.readers` and/or `startable.writers` system property to the name of the handler classes (see Section 10.2.3)

Explaining how to write such handlers is beyond the scope of this document - see the user document and javadocs for STIL (<http://www.starlink.ac.uk/stil/>).

## 4.2 Table Locations

In many cases loading and saving tables will be done using GUI dialogues such as the filestore load and save windows, where you just need to click on a filename or directory to indicate the load/save location. However in some cases, for instance specifying tables on the command line (Section 10.1) or typing pathnames directly into the load/save dialogue windows, you may want give the location of a table for input or output using only a single string.

Most of the time you will just want to type in a filename; either an absolute or relative pathname can be used. However, TOPCAT also supports direct use of URLs, including ones using some specialised protocols. Here is the list of URL types allowed:

**http:**

Read from HTTP resources.

**ftp:**

Read from anonymous FTP resources.

**file:**

Read from local files; not particularly useful since you can do much the same using just the filename.

**jar:**

Specialised protocol for looking inside Java Archive files - see `JarURLConnection` documentation.

**myspace:**

Accesses files in the AstroGrid "MySpace" virtual file store. These URLs look something like "myspace:/survey/iras\_psc.xml", and can access files in the myspace are that the user is currently logged into. These URLs can be used for both input and output of tables. To use them you must have an AstroGrid account and the AstroGrid WorkBench or similar must be running; if you're not currently logged in a dialogue will pop up to ask you for name and password.

**ivo:**

Understands ivo-type URLs which signify files in the AstroGrid "MySpace" virtual file store. These URLs look something like "ivo://uk.ac.le.star/filemanager#node-2583". These URLs can be used for both input and output of tables. To use them you must have an AstroGrid account and the AstroGrid WorkBench or similar must be running; if you're not currently logged in a dialogue will pop up to ask you for name and password.

**jdbc:**

Used for communicating with SQL-compliant relational databases. These are a bit different to normal URLs - see Section 4.1.1.10 and Section 4.1.2.9.

It is also possible to use the special value "-" to mean standard input; in this case the file format must be specified explicitly using the `-f` flag, and not all formats can be streamed from stdin. Finally, the form "<syscmd" or equivalently "syscmd|" may be used to read from the standard output of a shell pipeline (probably only works on Unix-like systems).

As with the GUI-based load dialogues, data compression in any of the supported formats (gzip, bzip2, Unix compress) is detected and dealt with automatically for input locations.

## 5 Joins and Matches

TOPCAT allows you to join two or more tables together to produce a new one in a variety of ways, and also to identify "similar" rows within a single table according to their cell contents. This section describes the facilities for performing these related operations.

There are two basic ways to join tables together: top-to-bottom and side-by-side. A top-to-bottom join (which here I call **concatenation**) is fairly straightforward in that it just requires you to decide which columns in one table correspond to which columns in the other. A side-by-side join is more complicated - it is rarely the case that row  $i$  in the first table should correspond to row  $i$  in the second one, so it is necessary to provide some criteria for deciding which (if any) row in the second table corresponds to a given row in the first. In other words, some sort of **matching** between rows in different tables needs to take place. This corresponds to what is called a *join* in database technology. Matching rows within a single table is a useful operation which involves many of the same issues, so that is described here too.

### 5.1 Concatenating Tables

Two tables can be concatenated using the Concatenation Window, which just requires you to specify the two tables to be joined, and for each column in the first ("Base") table, which column in the second ("Appended") table (if any) corresponds to it. The Apparent Table (Section 3) is used in each case. The resulting table, which is added to the list of known tables in the Control Window, has the same columns as the Base table, and a number of rows equal to the sum of the number of rows in the Base and Appended tables.

As a very simple example, concatenating these two tables:

Messier	RA	Dec	Name
-----	--	---	----
97	168.63	55.03	Owl Nebula
101	210.75	54.375	Pinwheel Galaxy
64	194.13	21.700	Black Eye Galaxy


and

RA2000	DEC2000	ID
-----	-----	--
185.6	58.08	M40
186.3	18.20	M85

with the assignments RA->RA2000, Dec->DEC2000 and Messier->ID would give:

Messier	RA	Dec	Name
-----	--	---	----
97	168.63	55.03	Owl Nebula
101	210.75	54.375	Pinwheel Galaxy
64	194.13	21.700	Black Eye Galaxy
M40	185.6	58.08	
M85	183.6	18.20	

Of course it is the user's responsibility to ensure that the correspondance of columns is sensible (that the two corresponding columns mean the same thing).

You can perform a concatenation using the Concatenation Window; obtain this using the **Joins|Concatenate Tables** () menu option in the Control Window.

### 5.2 Matching Rows Between Tables

When joining two tables side-by-side you need to identify which row(s) in one correspond to which

row(s) in the other. Conceptually, this is done by looking at each row in the first table, somehow identifying in the second table which row "refers to the same thing", and putting a new row in the joined table which consists of all the fields of the row in the first table, followed by all the fields of its matched row in the second table. The resulting table then has a number of columns equal to the sum of the number of columns in both input tables.

In practice, there are a number of complications. For one thing, each row in one table may be matched by zero, one or many rows in the other. For another, defining what is meant by "referring to the same thing" may not be straightforward. There is also the problem of actually identifying these matches in a relatively efficient way (without explicitly comparing each row in one table with each row in the other, which would be far too slow for large tables).

A common example is the case of matching two object catalogues - suppose we have the following catalogues:

Xpos	Ypos	Vmag
----	----	----
1134.822	599.247	13.8
659.68	1046.874	17.2
909.613	543.293	9.3

and

x	y	Bmag
-	-	----
909.523	543.800	10.1
1832.114	409.567	12.3
1135.201	600.100	14.6
702.622	1004.972	19.0

and we wish to combine them to create one new catalogue with a row for each object which appears in both tables. To do this, you have to specify what counts as a match - in this case let's say that a row in one table matches (refers to the same object as) a row in the other if the distance between the positions indicated by their X and Y coordinates matches to within one unit ( $\sqrt{(X_{\text{pos}}-x)^2 + (Y_{\text{pos}}-y)^2} \leq 1$ ). Then the catalogue we will end up with is:


Xpos	Ypos	Vmag	x	y	Bmag
----	----	----	-	-	----
1134.822	599.247	13.8	1135.201	600.100	14.6
909.613	543.293	9.3	909.523	543.800	10.1

There are a number of variations on this however - your match criteria might involve sky coordinates instead of Cartesian ones (or not be physical coordinates at all), you might want to match more than two tables, you might want to identify groups of matching objects in a single table, you might want the output to include rows which don't match as well...


The Match Window allows you to specify

- Which tables are to be matched
- What the criteria are for matching rows
- What to do for rows that don't have exactly one match
- What rows to include in the output table

and to start the matching operation. Depending on the type of match chosen, some additional columns may be appended to the resulting table giving additional details on how the match went. Usually, the 'match score' is one of these; The exact value and meaning of this column depends on the match, but it typically gives the distance between the matched points in some sensible units; the smaller the value, the better the match. You can find out exactly what this score means by examining the column's description in the Columns Window. Columns in the resulting table retain their original names unless that would lead to ambiguity, in which case a disambiguating suffix "\_1" or "\_2" is added to the column name.

To match two tables, use the **Pair Match** () button in the Control Window; to match more tables than two at once, use the other options on the Control Window's **Join** menu.

### 5.3 Matching Rows Within a Table

Although the effect is rather different, searching through a single table for rows which match each other (refer to the same object, as explained above) is a similar process and requires much of the same information to be specified, mainly, what counts as a match. You can do this using the Internal Match Window, obtained by using the **Internal Match** () button in the Control

Window's **Joins** menu.

### 5.4 Multi-Object Matches

The matching in TOPCAT is determined by specified match criteria, as described in Appendix A.8.1.1. These criteria give conditions for whether two items (table rows) count as matched with each other. In the case of a pair match, it is clear how this is to be interpreted.

However, some of the matching operations such as the Internal Match search for match groups which may have more than two members. This section explains how TOPCAT applies the pair-wise matching criteria it is given to identifying multi-object groups.

In a multi-object match context, the matcher identifies a matched group as the largest possible group of objects in which each is linked by a pair match to *any* other object in the group - it is a group of "friends of friends". Formally, the set of matched groups is a set of disjoint graphs whose nodes are input table rows and whose edges are successful pair matches, where no successful pair match exists between nodes in different elements of that set. Thus the set has a minimal number of elements, and each of its elements is a matched group of maximal size. The important point to note is that for any particular pair in a matched group, there is no guarantee that the two objects match each other, only that you can hop from one to the other via pairs which do match.

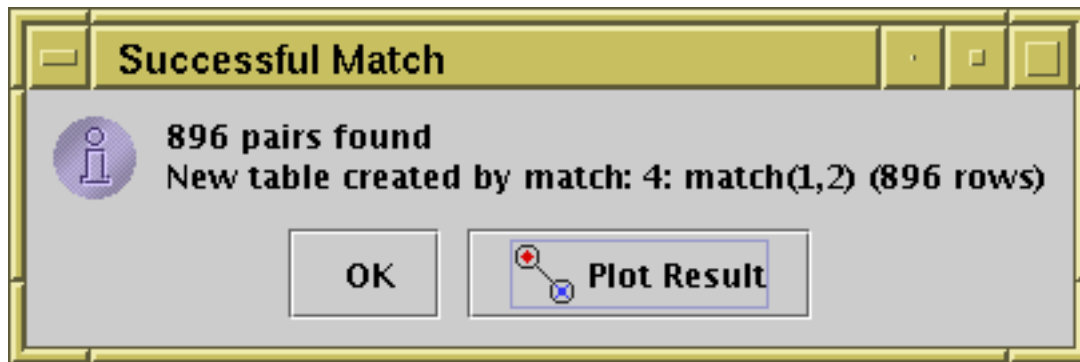
So in the case of a multi-object sky match on a field which is very crowded compared to the specified error radius, it is quite possible for *all* the objects in the input table(s) to end up as part of the same large matching group. Results at or near this percolation threshold are (a) probably not useful and (b) likely to take a long time to run. Some care should therefore be exercised when specifying match criteria in multi-object match contexts.

### 5.5 Plotting Match Results

Having performed a crossmatch, it is very often a good idea to look at the results graphically. If you can plot the input catalogues, and the output catalogue, with an indication of which rows in the inputs have been joined together, you can rapidly get a good idea of the result of the match, and whether it did what you expected. This is especially true in the presence of crowded fields, tentative match criteria, or other circumstances under which the match might not go as expected.

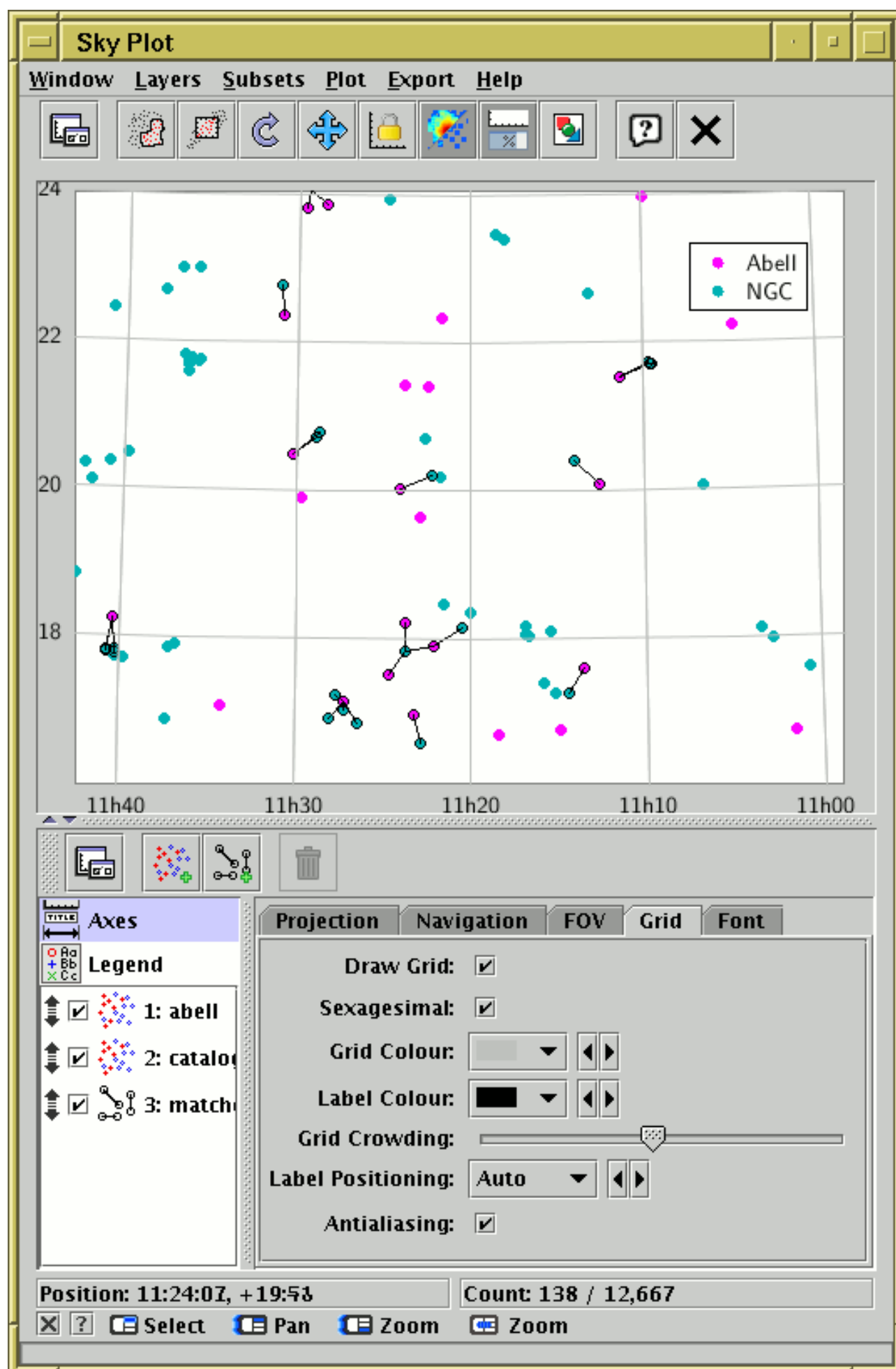
Since TOPCAT version 4.1, for most pair matches, you can do this automatically. When a suitable match completes, you may see a confirmation dialogue like this:





### Pair plot completion confirmation dialogue

If you click the **OK** button, it will simply dismiss the dialogue. However, if you click the **Plot Result** button, a new plot window will appear with all the points from the input catalogues and pair links (lines between the joined positions) for the joined rows. The result is something like this:



## Output from crossmatch Plot Result

It shows clearly which points have been joined. You can fiddle with the plot controls in the usual way to adjust the output (see Appendix A.4). It is also possible to set up such plots by hand.

## 5.6 Notes on Matching

This section provides a bit more detail on the how the row matching of local tables (sections Section 5.2 and Section 5.3) is done. It is designed to give a rough idea to interested parties; it is *not* a tutorial description from first principles of how it all works.

The basic algorithm for matching is based on dividing up the space of possibly-matching rows into an (indeterminate) number of bins. These bins will typically correspond to disjoint cells of a physical or notional coordinate space, but need not do so. In the first step, each row of each table is assessed to determine which bins might contain matches to it - this will generally be the bin that it falls into and any "adjacent" bins within a distance corresponding to the matching tolerance. A reference to the row is associated with each such bin. In the second step, each bin is examined, and if two or more rows are associated with it every possible pair of rows in the associated set is assessed to see whether it does in fact constitute a matched pair. This will identify all and only those row pairs which are related according to the selected match criteria. During this process a number of optimisations may be applied depending on the details of the data and the requested match.

The matching algorithm described above is roughly an  $O(N \log(N))$  process, where  $N$  is the total number of rows in all the tables participating in a match. This is good news, since the naive interpretation would be  $O(N^2)$ . This can break down however if the matching tolerance is such that the number of rows associated with some or most bins gets large, in which case an  $O(M^2)$  component can come to dominate, where  $M$  is the number of rows per bin. The average number of rows per bin is reported in the logging while a match is proceeding, so you can keep an eye on this.

For more detail on the matching algorithms, see the javadocs for the `uk.ac.starlink.table.join` package, or contact the author.

## 5.7 Matching against a Remote Table

TOPCAT provides a number of facilities for positional crossmatches against tables which are exposed via Virtual Observatory web services (Cone Search, Simple Image Access and Simple Spectral Access) and general SQL-like matching (TAP). These work rather differently from the other functions described in this section, which operate on local tables, though conceptually the result is similar. See Section 6 for more details.

It also provides access to the X-Match service provided by the CDS, as described in Appendix A.10.1, which allows fast matching against any VizieR table or SIMBAD.

In many cases the CDS Upload X-Match window is the easiest, fastest and best way to match against remote tables. TAP is more flexible, but has a somewhat steeper learning curve, and may not be as fast or scalable for large local tables. The multi-SIA and multi-SSA windows may be useful for performing per-row searches of modest size against remote image or spectral archives. The multi-cone window is (since version 4.2) present mostly for historical reasons, and is not usually a good choice, since it is much less efficient than TAP or CDS X-Match.

## 6 Virtual Observatory Access

Several of the windows in TOPCAT allow you to interface with so-called *Data Access Layer* services provided within the Virtual Observatory (VO). The buzzwords are not important, but the basic idea is that this allows you to locate a service providing data which may be of interest to you, and then query that service to obtain the data. The VO is not a single monolithic entity, but a set of protocols which allow a general purpose tool such as TOPCAT to talk to services made available by many different participating data providers in a uniform way, without having to have prior knowledge of what services are out there or the details of how their data is arranged.

The basic operation is similar for all of TOPCAT's access to these services:

1. Select the *registry* to query for services (or use the default one)
2. Query that registry for services of interest
3. From the returned list of services, select one that you wish to query for data
4. Specify the query to send to the data service
5. Start the query and wait for the result

These ideas are explained in more detail in the following subsections. The windows from which this is done are documented in Appendix A.9.

**Note:** For information on SAMP or PLASTIC, which are protocols developed within the Virtual Observatory context, but are not necessarily concerned with remote data access, see Section 9.

### 6.1 The Registry

The **Registry** is fundamental to the way that the VO works. A registry is a list of all the services made available by different data providers. Each entry records some information about the type of service, who provided it, what kind of data it contains, and so on (registries may contain other types of entry as well, but we will not discuss these further here). Any data provider can add an entry to the registry to advertise that it has certain datasets available for access.

Several registries exist; they tend to be maintained by different regional VO organisations. At the time of writing, there are registries maintained for public access by GAVO, ESA, STScI and (remnants of) the UK's AstroGrid, amongst others. Particular projects may also maintain their own registries with limited holdings for internal use. The main public access registries talk to each other to synchronize their contents, so to a first approximation, they contain similar lists of entries to each other, and it shouldn't matter too much which one you use. In practice, there are some differences of format and content between them, so one may work better than another for you or may contain a record that you need. In most cases though, using the default registry (currently the GAVO one) will probably do what you want it to.

### 6.2 Data Access Services

A number of different service types are defined and listed in the registry; the ones that TOPCAT currently knows how to access are the following:

**Cone Search:**

retrieve entries in a certain region of the sky from a remote catalogue

**Simple Image Access (SIA):**

find image data (often in FITS format) in a certain region of the sky from a remote image archive

**Simple Spectral Access (SSA):**

find spectral data, usually in a certain region of the sky, from a remote archive of spectral

observations or theoretical models

**Table Access Protocol (TAP):**

make free-form queries to a remote database using an SQL-like query language

Detailed technical information about these protocols can be found at the IVOA web site (<http://www.ivoa.net/>) in the Cone Search, SIA, SSA and TAP documents, but these are by no means required reading for users of the services. These protocols (apart from Cone Search) are quite complex and have many specialised and optional features. The options offered for Cone Search and TAP are reasonably complete, but for SIA and SSA TOPCAT only provides a fairly basic level of interaction, and many features (for instance SSA queries by wavelength, or non-positional queries for theoretical data) are not accessible from it.

Cone Search, SIA and SSA are *positional* protocols meaning that they query a single region of the sky. TOPCAT provides access to these service types in two main ways:

**Single positional query**

If you enter by hand a sky position (RA, Dec) and radius, you can download a table containing the results of a search for a single (usually, small) region on the sky. See the sections on Cone, SIA, SSA load dialogues in Appendix A.9.

**Multiple positional query**

You can define how each row of an input table selects a region on the sky. This will usually correspond to selecting a column for RA and a column for Dec, and either a fixed radius or a column giving the radius. Then a positional query can be made for each row of the input table. The result is effectively a join between the input Apparent Table and the remote table of object data, images or spectra. See the sections on Cone, SIA, SSA multiple query windows in Appendix A.9.

TAP is a much more powerful protocol not restricted to positional queries, and has its own interface. See the TAP load dialogue section in Appendix A.9.

## 7 Algebraic Expression Syntax

TOPCAT allows you to enter algebraic expressions in a number of contexts:

1. To define a new column in terms of existing columns in the Synthetic Column dialogue
2. To define a new Row Subset (Section 3.1) on the basis of table data in the Algebraic Subset dialogue
3. To define a custom Activation Action (Section 8) in the Activation dialogue.
4. When faced with a column selector for plotting or other purposes, in some cases you can type in an expression rather than selecting or typing a simple column name.

This is a powerful feature which permits you to manipulate and select table data in very flexible ways - you can think of it like a sort of column-oriented spreadsheet. The syntax for entering these expressions is explained in this section.

What you write are actually expressions in the Java language, which are compiled into Java bytecode before evaluation. However, this does not mean that you need to be a Java programmer to write them. The syntax is pretty similar to C, but even if you've never programmed in C most simple things, and some complicated ones, are quite intuitive.

The following explanation gives some guidance and examples for writing these expressions. Unfortunately a complete tutorial on writing Java is beyond the scope of this document, but it should provide enough information for even a novice to write useful expressions.

The expressions that you can write are basically any function of all the column values and subset inclusion flags which apply to a given row; the function result can then define the per-row value of a new column, or the inclusion flag for a new subset, or the action to be performed when a row is activated by clicking on it. If the built-in operators and functions are not sufficient, or it's unwieldy to express your function in one line of code, you can add new functions by writing your own classes - see Section 7.9.

**Note:** if Java is running in an environment with certain security restrictions (a security manager which does not permit creation of custom class loaders) then algebraic expressions won't work at all, and the buttons which allow you to enter them will be disabled.

### 7.1 Referencing Cell Values

To create a useful expression for a cell in a column, you will have to refer to other cells in different columns of the same table row. You can do this in three ways:

#### By Name

The Name of the column may be used if it is unique (no other column in the table has the same name) and if it has a suitable form. This means that it must have the form of a Java variable - basically starting with a letter and continuing with letters or numbers. In particular it cannot have any spaces in it. The underscore and currency symbols count as letters for this purpose. Column names are treated case-insensitively.

#### By \$ID

The "\$ID" identifier of the column may always be used to refer to it; this is a useful fallback if the column name isn't suitable for some reason (for instance it contains spaces or is not unique). This is just a "\$" sign followed by a unique integer assigned by the program to each column when it is first encountered. You can find out the \$ID identifier by looking in the Columns Window.

#### By ucd\$ specifier

If the column has a Unified Content Descriptor (this will usually only be the case for VOTable or possibly FITS format tables) you can refer to it using an identifier of the form

`ucd$<ucd-spec>`". Depending on the version of UCD scheme used, UCDs can contain various punctuation marks such as underscores, semicolons and dots; for the purpose of this syntax these should all be represented as underscores ("\_"). So to identify a column which has the UCD "phot.mag;em.opt.R", you should use the identifier "ucd\$phot\_mag\_em\_opt\_r". Matching is not case-sensitive. Furthermore, a trailing underscore acts as a wildcard, so that the above column could also be referenced using the identifier "ucd\$phot\_mag\_". If multiple columns have UCDs which match the given identifier, the first one will be used.

Note that the same syntax can be used for referencing table parameters (see Section 7.3); columns take preference so if a column and a parameter both match the requested UCD, the column value will be used.

### By `utype$` specifier

If the column has a **Utype** (this will usually only be the case for VOTable or possibly FITS format tables) you can refer to it using an identifier of the form "utype\$<utype-spec>". Utypes can contain various punctuation marks such as colons and dots; for the purpose of this syntax these should all be represented as underscores ("\_"). So to identify a column which has the Utype "ssa:Access.Format", you should use the identifier "utype\$ssa\_Access\_Format". Matching is not case-sensitive. If multiple columns have Utypes which match the given identifier, the first one will be used.

Note that the same syntax can be used for referencing table parameters (see Section 7.3); columns take preference so if a column and a parameter both match the requested Utype, the column value will be used.

### With the `Object$` prefix

If a column is referenced with the prefix "Object\$" before its identifier (e.g. "Object\$BMAG" for a column named BMAG) the result will be the column value as a java Object. Without that prefix, numeric columns are evaluated as java primitives. In most cases, you *don't want to do this*, since it means that you can't use the value in arithmetic expressions. However, if you need the value to be passed to a (possibly user-defined activation) method, and you need that method to be invoked even when the value is null, you have to do it like this. Null-valued primitives otherwise cause expression evaluation to abort.

There is also a special column:

#### `$index`

The value of this is the current row number in the unsorted table (the first row is 1). This is the value seen in the grey numbers at the left of the grid in the Data Window). You can alternatively use the form `$0` . (The form `index` is also permitted, but deprecated). Note that this value is a `long` (8-byte integer); when using it in certain expressions you may find it necessary to convert it to an `int` (4-byte integer) using the `toInteger()` function.

The value of the variables so referenced will be a primitive (boolean, byte, short, char, int, long, float, double) if the column contains one of the corresponding types. Otherwise it will be an Object of the type held by the column, for instance a String. In practice this means: you can write the name of a column, and it will evaluate to the numeric (or string) value that that column contains in each row. You can then use this in normal algebraic expressions such as "B\_MAG-U\_MAG" as you'd expect.

## 7.2 Referencing Row Subset Flags

If you have any Row Subsets defined you can also access the value of the boolean (true/false) flag indicating whether the current row is in each subset. Again there are two ways of doing this:

### By Name

The name assigned to the subset when it was created can be used if it is unique and if it has a

suitable form. The same comments apply as to column names above.

### By `_ID`

The "`_ID`" identifier of the subset may always be used to refer to it. Like the "`$ID`" identifier for columns above, this is a unique integer preceded by a special symbol, this time the underscore, "`_`".

**Note:** in early versions of TOPCAT the hash sign ("`#`") was used instead of the underscore for this purpose; the hash sign no longer has this meaning.

In either case, the value will be a boolean value; these can be useful in conjunction with the conditional "`? :`" operator or when combining existing subsets using logical operators to create a new subset.

## 7.3 Referencing Table Parameters

Some tables have constant values associated with them; these may represent such things as the epoch at which observations were taken, the name of the catalogue, an angular resolution associated with all observations, or any number of other things. Such constants are known as *table parameters* and can be viewed and modified in the Parameter Window. The values of such parameters can be referenced in algebraic expressions as follows:

### **`param$name`**

If the parameter name has a suitable form (starting with a letter and continuing with letters or numbers) it can be referenced by prefixing that name with the string `param$`.

### **`ucd$ucd-spec`**

If the parameter has a Unified Content Descriptor it can be referenced by prefixing the UCD specifier with the string `ucd$`. Any punctuation marks in the UCD should be replaced by underscores, and a trailing underscore is interpreted as a wildcard. See Section 7.1 for more discussion.

### **`utype$utype-spec`**

If the parameter has a Utype, it can be referenced by prefixing the Utype specifier with the string `utype$`. Any punctuation marks in the Utype should be replaced by underscores. See Section 7.1 for more discussion.

Note that if a parameter has a name in an unsuitable form (e.g. containing spaces) and has no UCD then it cannot be referenced in an expression.

There are also a couple of special values:

### **`$ncol`**

The number of columns in the table. This figure refers to the underlying table, not the apparent table, so it is not affected by hiding columns.

### **`$nrow`**

The number of rows in the table. This figure refers to the underlying table, not the apparent table, so it is not affected by the value of the current subset. Note that this value is a `long` (8-byte integer); when using it in certain expressions you may find it necessary to convert it to an `int` (4-byte integer) using the `toInteger()` function.

## 7.4 Null Values

When no special steps are taken, if a null value (blank cell) is encountered in evaluating an expression (usually because one of the columns it relies on has a null value in the row in question) then the result of the expression is also null.



It is possible to exercise more control than this, but it requires a little bit of care, because the expressions work in terms of primitive values (numeric or boolean ones) which don't in general have a defined null value. The name "null" in expressions gives you the java `null` reference, but this cannot be matched against a primitive value or used as the return value of a primitive expression.

For most purposes, the following two tips should enable you to work with null values:

### Testing for null

To test whether a column contains a null value, prepend the string "NULL\_" (use upper case) to the column name or \$ID. This will yield a boolean value which is true if the column contains a blank or a floating point NaN (not-a-number) value, and false otherwise.

### Returning null

To return a null value from a numeric expression, use the name "NULL" (upper case). To return a null value from a non-numeric expression (e.g. a String column) use the name "null" (lower case).

Null values are often used in conjunction with the conditional operator, "? :"; the expression

```
test ? tval : fval
```

returns the value `tval` if the boolean expression `test` evaluates true, or `fval` if `test` evaluates false. So for instance the following expression:

```
Vmag == -99 ? NULL : Vmag
```

can be used to define a new column which has the same value as the `Vmag` column for most values, but if `Vmag` has the "magic" value -99 the new column will contain a blank. The opposite trick (substituting a blank value with a magic one) can be done like this:

```
NULL_Vmag ? -99 : Vmag
```

Some more examples are given in Section 7.8.

## 7.5 Operators

The operators are pretty much the same as in the C language. The common ones are:

### Arithmetic

- +** (add)
- (subtract)
- \*** (multiply)
- /** (divide)
- %** (modulus)

### Logical

- !** (not)
- &&** (and)
- ||** (or)
- ^** (exclusive-or)
- ==** (numeric identity)
- !=** (numeric non-identity)
- <** (less than)
- >** (greater than)
- <=** (less than or equal)
- >=** (greater than or equal)

**Bitwise**

**& (and)**  
**| (or)**  
**^ (exclusive-or)**  
**<< (left shift)**  
**>> (right shift)**  
**>>> (logical right shift)**

**Numeric Typecasts**

**(byte) (numeric -> signed byte)**  
**(short) (numeric -> 2-byte integer)**  
**(int) (numeric -> 4-byte integer)**  
**(long) (numeric -> 8-byte integer)**  
**(float) (numeric -> 4-byte floating point)**  
**(double) (numeric -> 8-byte floating point)**

Note you may find the numeric conversion functions in the **Maths** class described in Appendix B.1 below more convenient for numeric conversions than these.

**Other**

**+ (string concatenation)**  
**[] (array dereferencing - first element is zero)**  
**?: (conditional switch)**  
**instanceof (class membership)**

**7.6 Functions**

Many functions are available for use within your expressions, covering standard mathematical and trigonometric functions, arithmetic utility functions, type conversions, and some more specialised astronomical ones. You can use them in just the way you'd expect, by using the function name (unlike column names, this is case-sensitive) followed by comma-separated arguments in brackets, so

```
max ( IMAG , JMAG )
```

will give you the larger of the values in the columns IMAG and JMAG, and so on.

The functions are grouped into the following classes:

**Tilings**

Pixel tiling functions for the celestial sphere.

The *k* parameter for the HEALPix functions is the HEALPix order, which can be in the range  $0 \leq k \leq 29$ . This is the logarithm to base 2 of the HEALPix NSIDE parameter. At order *k*, there are  $12 \cdot 4^k$  pixels on the sphere.

**Arithmetic**

Standard arithmetic functions including things like rounding, sign manipulation, and maximum/minimum functions.

**Conversions**

Functions for converting between strings and numeric values.

**Distances**

Functions for converting between different measures of cosmological distance.

The following parameters are used:

- **z**: redshift
- **H0**: Hubble constant in km/sec/Mpc (example value ~70)
- **omegaM**: Density ratio of the universe (example value 0.3)
- **omegaLambda**: Normalised cosmological constant (example value 0.7)

For a flat universe,  $\text{omegaM} + \text{omegaLambda} = 1$

The terms and formulae used here are taken from the paper by D.W.Hogg, *Distance measures in cosmology*, astro-ph/9905116 (<http://arxiv.org/abs/astro-ph/9905116>) v4 (2000).

## KCorrections

Functions for calculating K-corrections.

## Times

Functions for conversion of time values between various forms. The forms used are

### Modified Julian Date (MJD)

A continuous measure in days since midnight at the start of 17 November 1858. Based on UTC.

### ISO 8601

A string representation of the form `yyyy-mm-ddThh:mm:ss.s`, where the `T` is a literal character (a space character may be used instead). Based on UTC.

### Julian Epoch

A continuous measure based on a Julian year of exactly 365.25 days. For approximate purposes this resembles the fractional number of years AD represented by the date. Sometimes (but not here) represented by prefixing a 'J'; J2000.0 is defined as 2000 January 1.5 in the TT timescale.

### Besselian Epoch

A continuous measure based on a tropical year of about 365.2422 days. For approximate purposes this resembles the fractional number of years AD represented by the date. Sometimes (but not here) represented by prefixing a 'B'.

### Decimal Year

Fractional number of years AD represented by the date. 2000.0, or equivalently 1999.99recurring, is midnight at the start of the first of January 2000. Because of leap years, the size of a unit depends on what year it is in.

Therefore midday on the 25th of October 2004 is `2004-10-25T12:00:00` in ISO 8601 format, 53303.5 as an MJD value, 2004.81588 as a Julian Epoch and 2004.81726 as a Besselian Epoch.

Currently this implementation cannot be relied upon to better than a millisecond.

## TrigDegrees

Standard trigonometric functions with angles in degrees.

## Maths

Standard mathematical and trigonometric functions. Trigonometric functions work with angles in radians.

## Arrays

Functions which operate on array-valued cells. The array parameters of these functions can only be used on values which are already arrays (usually, numeric arrays). In most cases that means on values in table columns which are declared as array-valued. FITS and VOTable tables can have columns which contain array values, but other formats such as CSV cannot.

If you want to calculate aggregating functions like sum, min, max etc on multiple values which are not part of an array, it's easier to use the functions from the `Lists` class.

Note that none of these functions will calculate statistical functions over a whole column of a

table.

The functions fall into a number of categories:

- Aggregating operations, which map an array value to a scalar, including `size`, `count`, `countTrue`, `maximum`, `minimum`, `sum`, `mean`, `median`, `quantile`, `stdev`, `variance`, `join`.
- Operations on one or more arrays which produce array results, including `add`, `subtract`, `multiply`, `divide`, `reciprocal`, `condition`.
- The function `array`, which lets you assemble an array value from a list of scalar numbers. This can be used with the aggregating functions here, but it's generally easier to use the corresponding functions from the `Lists` class.

## Fluxes

Functions for conversion between flux and magnitude values. Functions are provided for conversion between flux in Janskys and AB magnitudes.

Some constants for approximate conversions between different magnitude scales are also provided:

- Constants `JOHNSON_AB_*`, for Johnson  $\leftrightarrow$  AB magnitude conversions, from Frei and Gunn, *Astronomical Journal* 108, 1476 (1994), Table 2 (1994AJ....108.1476F) (<http://adsabs.harvard.edu/abs/1994AJ....108.1476F>).
- Constants `VEGA_AB_*`, for Vega  $\leftrightarrow$  AB magnitude conversions, from Blanton et al., *Astronomical Journal* 129, 2562 (2005), Eqs. (5) (2005AJ....129.2562B) (<http://adsabs.harvard.edu/abs/2005AJ....129.2562B>).

## Strings

String manipulation and query functions.

## Formats

Functions for formatting numeric values.

## CoordsRadians

Functions for angle transformations and manipulations, based on radians rather than degrees. In particular, methods for translating between radians and HH:MM:SS.S or DDD:MM:SS.S type sexagesimal representations are provided.

## Coverage

Functions related to coverage and footprints.

One coverage standard is Multi-Order Coverage maps, described at <http://www.ivoa.net/Documents/MOC/> (<http://www.ivoa.net/Documents/MOC/>). MOC positions are always defined in ICRS equatorial coordinates.

MOC locations may be given as either the filename or the URL of a MOC FITS file. Alternatively, they may be the identifier of a VizieR table, for instance "v/139/sdss9" (SDSS DR9). A list of all the MOCs available from VizieR can currently be found at <http://alasky.u-strasbg.fr/footprints/tables/vizier/> (<http://alasky.u-strasbg.fr/footprints/tables/vizier/>). You can search for VizieR table identifiers from the VizieR web page (<http://vizier.u-strasbg.fr/> (<http://vizier.u-strasbg.fr/>)); note you must use the *table* identifier (like "v/139/sdss9") and not the *catalogue* identifier (like "v/139").

## Lists

Functions which operate on lists of values.

Some of these resemble similar functions in the `Arrays` class, and in some cases are interchangeable, but these are easier to use on non-array values because you don't have to explicitly wrap up lists of arguments as an array. However, for implementation reasons, most of the functions defined here can be used on values which are already `double[]` arrays (for instance array-valued columns) rather than as comma-separated lists of floating point values.

## CoordsDegrees

Functions for angle transformations and manipulations, with angles generally in degrees. In particular, methods for translating between degrees and HH:MM:SS.S or DDD:MM:SS.S type sexagesimal representations are provided.

A listing of the functions in these classes is given in Appendix B.1, and complete documentation on them is available within TOPCAT from the Available Functions Window.

### 7.6.1 Technical Note

This note provides a bit more detail for Java programmers on what is going on here; only read on if you want to understand how the use of functions in TOPCAT algebraic expressions relates to normal Java code.

The expressions which you write are compiled to Java bytecode when you enter them (if there is a 'compilation error' it will be reported straight away). The functions listed in the previous subsections are all the `public static` methods of the classes which are made available by default. The classes listed are all in the packages `uk.ac.starlink.ttools.func` and `uk.ac.starlink.topcat.func` (`uk.ac.starlink.topcat.func.Strings` etc). However, the public static methods are all imported into an anonymous namespace for bytecode compilation, so that you write `sqrt(x)` and not `Maths.sqrt(x)`. The same happens to other classes that are imported (which can be in any package or none) - their public static methods all go into the anonymous namespace. Thus, method name clashes are a possibility.

This cleverness is all made possible by the rather wonderful JEL (<http://www.gnu.org/software/jel/>).

### 7.7 Instance Methods

There is another category of functions which can be used apart from those listed in the previous section. These are called, in Java/object-oriented parlance, "instance methods" and represent functions that can be executed on an object.

It is possible to invoke any of its public instance methods on any object (though not on primitive values - numeric and boolean ones). The syntax is that you place a "." followed by the method invocation after the object you want to invoke the method on, hence `NAME.substring(3)` instead of `substring(NAME,3)`. If you know what you're doing, feel free to go ahead and do this. However, most of the instance methods you're likely to want to use have equivalents in the normal functions listed in the previous section, so unless you're a Java programmer or feeling adventurous, you are probably best off ignoring this feature.

### 7.8 Examples

Here are some general examples. They could be used to define synthetic columns or (where numeric) to define values for one of the axes in a plot.

#### Average

```
(first + second) * 0.5
```

#### Square root

```
sqrt(variance)
```

#### Angle conversion

```
radiansToDegrees(DEC_radians)
degreesToRadians(RA_degrees)
```

### Conversion from string to number

```
parseInt($12)
parseDouble(ident)
```

### Conversion from number to string

```
toString(index)
```

### Conversion between numeric types

```
toShort(obs_type)
toDouble(range)
```

*or*

```
(short) obs_type
(double) range
```

### Conversion from sexagesimal to degrees

```
hmsToDegrees(RA1950)
dmsToDegrees(decDeg,decMin,decSec)
```

### Conversion from degrees to sexagesimal

```
degreesToDms($3)
degreesToHms(RA,2)
```

### Outlier clipping

```
min(1000, max(value, 0))
```

### Converting a magic value to null

```
jmag == 9999 ? NULL : jmag
```

### Converting a null value to a magic one

```
NULL_jmag ? 9999 : jmag
```

### Taking the third scalar element from an array-valued column

```
psfCounts[2]
```

and here are some examples of boolean expressions that could be used to define row subsets (or to create boolean synthetic columns):

#### Within a numeric range

```
RA > 100 && RA < 120 && Dec > 75 && Dec < 85
```

#### Within a circle

```
$2*$2 + $3*$3 < 1
skyDistanceDegrees(ra0,dec0,hmsToDegrees(RA),dmsToDegrees(DEC))<15./3600.
```

#### First 100 rows

```
index <= 100
```

### Every tenth row

```
index % 10 == 0
```

### String equality/matching

```
equals(SECTOR, "ZZ9 Plural Z Alpha")
equalsIgnoreCase(SECTOR, "zz9 plural z alpha")
startsWith(SECTOR, "ZZ")
contains(ph_qual, "U")
```

### String regular expression matching

```
matches(SECTOR, "[XYZ] Alpha")
```

### Combining subsets

```
(_1 && _2) && ! _3
```

### Test for non-blank value


```
! NULL_ellipticity
```

## 7.9 Adding User-Defined Functions

The functions provided by default for use with algebraic expressions, while powerful, may not provide all the operations you need. For this reason, it is possible to write your own extensions to the expression language. In this way you can specify arbitrarily complicated functions. Note however that this will only allow you to define new columns or subsets where each cell is a function only of the other cells in the same row - there is no way to define a value in one row as a function of values in other rows.

In order to do this, you have to write and compile a (probably short) program in the Java language. A full discussion of how to go about this is beyond the scope of this document, so if you are new to Java and/or programming you may need to find a friendly local programmer to assist (or mail the author). The following explanation is aimed at Java programmers, but may not be incomprehensible to non-specialists.

The steps you need to follow are:

1. Write and compile a class containing one or more static public methods representing the function(s) required
2. Make this class available on the application's classpath at runtime as described in Section 10.2.1
3. Specify the class's name to the application, *either* as the value of the `jel.classes` or `jel.classes.activation` system properties (colon-separated if there are several) as described in Section 10.2.3 *or* during a run using the Available Function Window's **Add Class** ()

button

Any public static methods defined in the classes thus specified will be available for use in the Synthetic Column, Algebraic Subset or (in the case of activation functions only) Activation Window windows. They should be defined to take and return the relevant primitive or Object types

for the function required (in the case of activation functions the return value should normally be a short log string).

For example, a class written as follows would define a three-value average:

```
public class AuxFuncs {  
    public static double average3(double x, double y, double z) {  
        return (x + y + z) / 3.0;  
    }  
}
```

and the expression "average3(\$1,\$2,\$3)" could then be used to define a new synthetic column, giving the average of the first three existing columns. Exactly how you would build this is dependent on your system, but it might involve doing something like the following:

1. Writing a file named "AuxFuncs.java" containing the above code
2. Compiling it using a command like "javac AuxFuncs.java"
3. Starting up TOPCAT with the flags: "topcat -Djel.classes=AuxFuncs -classpath ."

Note that (in versions later than TOPCAT 4.3-2) variable-length argument lists are supported where the final declared argument is an array, so for instance a method declared like:

```
public static double sum(double[] values) {  
    double total = 0;  
    for (int i = 0; i < values.length; i++) {  
        total += values[i];  
    }  
    return total;  
}
```

can be invoked like "sum(UMAG,GMAG,RMAG,IMAG,ZMAG)". The alternative form "double... values" can be used in the declaration with identical effect.



## 8 Activation Actions

As well as seeing the overview of table data provided by a plot or statistics summary, it is often necessary to focus on a particular row of the table, which according to the nature of the table may represent an astronomical object, an event or some other entity. In the Data Window a table row is simply a row of the displayed JTable, and in a plot it corresponds to one plotted point.

If you click on plotted point in one of the graphics windows, or on a row in the Data Window, the corresponding table row will be *activated*. When a row is activated, three things happen:

1. If that row is represented by a point in any open 2- or 3-dimensional scatter plot windows, a visible cursor marker will be drawn centred on that point.
2. If the Data Window is visible, the table will be scrolled to show the row and it will be highlighted
3. If an *activation action* has been defined, it will be invoked

The first two of these mean that you can easily see which point in a plot corresponds to which row in the table and vice versa - just click on one and the other will be highlighted.

The third one can be more complicated. By default, no activation action is set, so nothing else happens, and this may very well be what you want. However, by clicking on the **Activation Action** selector in the Control Window you can bring up the Activation Window which enables you to choose an additional action to take place. There are various options here and various ways to achieve them (see Appendix A.10.3 for more details) but the kinds of actions which are envisaged are to display one or more images or spectra relating to the row you have identified. One of the options available for instance retrieves a postage-stamp image of a few arcminutes around the sky position defined by the row from a SuperCOSMOS all-sky image survey and pops it up in a viewer window. So for instance having spotted an interesting point in a plot of a galaxy catalogue you can click on it, and immediately see a picture to identify its morphological type.

The exact actions you want to perform may be closely tailored to the data you have, for instance you may have a set of spectra on disk named by object ID. It's impossible to cater for such possibilities with a set of pre-packaged options, so you are able to define your own custom actions here. This is done by writing a expression using the syntax described in Section 7. A number of special functions (described in the following subsection) are provided to do things like display an image or a spectrum in a browser (given its filename or URL), or access data from certain data servers on the web, but there is nothing to stop the adventurous plugging in their own external programs so in principle you can configure pretty much anything to happen on the basis of the values in the row that you have activated.

### 8.1 Activation Functions

When defining custom activation actions in the Activation Window, you enter an expression to be invoked on a row when it is activated. This expression uses the syntax defined in Section 7 and can make use of the functions listed in Appendix B.1. However in this case there is an additional list of functions you can use which cause something to happen (for instance displaying an image) rather than just returning a value. The following classes of functions are available:

#### Browsers

Displays URLs in web browsers.

#### Mgc

Specialist functions for use with data from the the Millennium Galaxy Survey.

#### System

Executes commands on the local operating system. These are executed as if typed in from the shell, or command line.

**Sdss**

Specialist display functions for use with the Sloan Digital Sky Server.

**SuperCosmos**

Specialist display functions for use with the SuperCOSMOS survey. These functions display cutout images from the various archives hosted at the SuperCOSMOS Sky Surveys (<http://www-wfau.roe.ac.uk/sss/> (<http://www-wfau.roe.ac.uk/sss/>)). In most cases these cover the whole of the southern sky.

**BasicImageDisplay**

Functions for display of graphics-format images in a no-frills viewing window (an `ImageWindow`). Supported image formats include GIF, JPEG, PNG and FITS, which may be compressed.

**TwoQZ**

Specialist functions for use with data from the the 2QZ survey. Spectral data are taken directly from the 2QZ web site at <http://www.2dfquasar.org/> (<http://www.2dfquasar.org/>).

**Image**

Functions for display of images in a window. Supported image formats include GIF, JPEG, PNG and FITS, which may be compressed. The SoG program (<http://www.starlink.ac.uk/sog/> (<http://www.starlink.ac.uk/sog/>)) will be used if it is available, otherwise a no-frills image viewer will be used instead.

**Output**

Functions for simple logging output.

**Sog**

Functions for display of images in external viewer SOG (<http://www.starlink.ac.uk/sog/> (<http://www.starlink.ac.uk/sog/>)).

A listing of the functions in these classes is given in Appendix B.2, and complete documentation on them is available within TOPCAT from the Available Functions Window.

## 9 Tool Interoperability

TOPCAT is able to communicate with other tools using one or other of two messaging protocols:

- SAMP (Simple Applications Messaging Protocol)
- PLASTIC (PLatform for AStronomical InterConnection)

An example of the kind of thing which can be done might be:

1. TOPCAT sends a catalogue to an image display tool
2. The image display tool shows the catalogue entries as markers placed appropriately on a displayed image
3. User actions which highlight one of the points in one tool can then automatically highlight the same point in the other

Examples of the kind of tool which TOPCAT can interoperate with in this way are image analysis tools (Aladin, GAIA, ds9) table analysis tools (VisIVO, STILTS, other instances of TOPCAT itself), spectrum analysis tools (SPLAT, VOSpec), sky visualisation tools (World Wide Telescope, VirGO), scripting languages (SAMPy), and others.

SAMP and PLASTIC do much the same job as each other, and work in much the same way. SAMP is an evolution of PLASTIC with a number of technical improvements, and PLASTIC has been deprecated in favour of SAMP since around 2009. PLASTIC is therefore of mainly historical interest, though TOPCAT can still run in PLASTIC mode on request, if you need it to communicate with older tools that can only speak PLASTIC.


The communication architecture of the two protocols is basically the same: all tools communicate with a central "Hub" process, so a hub must be running in order for the messaging to operate. If a hub is running when TOPCAT starts, or if one starts up while TOPCAT is in operation, it will connect to it automatically. If no SAMP hub is running, TOPCAT will set one up during application startup.

TOPCAT can work in either SAMP or PLASTIC mode, but not both at once. It determines which mode to use at startup: if the `-samp` or `-plastic` flag is supplied on the command line the corresponding mode will be used; otherwise it will try to use SAMP. It is easy to tell which mode is being used by looking at the Control Window; in SAMP mode the SAMP panel displaying connection and message status is visible at the bottom of the right hand panel (there are a few other changes to the GUI as well).

This communication has two aspects to it: on the one hand TOPCAT can send messages to other applications which causes them to do things, and on the other hand TOPCAT can receive and act on such messages sent by other applications. The sent and received messages are described separately in the subsections below. There are also sections on the, somewhat different, ways to control and monitor messaging operation for the cases of SAMP and PLASTIC.

### 9.1 SAMP control

When running in SAMP mode, the Control Window features several ways to view and control SAMP status.

**SAMP Status button** (  )

Pops up the SAMP Window which shows a detailed view of the status of SAMP communications, and allows you to register or unregister with the hub, and to start a hub if required.

#### **SAMP Status panel**

The bottom part of the right hand panel has an area labelled **SAMP**. This gives a summary




view of registration status, other connected applications, and messages recently sent and received. It is described in more detail in Appendix A.2.4.

### Table Broadcast () and Table Send () menu options

The **Interop** menu provides these options which allow you to send the currently selected table to one or all connected clients. Note they will only be enabled if the hub is running, and there is at least one connected client which knows how to receive a table. The **Broadcast** option is also available on the toolbar.

### Stop Internal Hub () menu option

By default, when TOPCAT starts up, it will look for a SAMP hub, and if none appears to be running it will start one internally, which will normally run until TOPCAT exits. This is usually not problematic, but if you would prefer to run a hub external to TOPCAT, then you may need to shut down TOPCAT's before starting a new one. Using this option shuts down TOPCAT's internal hub.

A number of other windows feature an **Interop** menu with **Broadcast** () and **Send** () operations for data types appropriate to that window. Sometimes Broadcast appears in the toolbar as well. In some cases there are also **Accept** () toggle options in the Interop menu. These control whether TOPCAT will respond to appropriate messages sent by other SAMP applications.

## 9.2 PLASTIC control

*Note: the PLASTIC protocol has been deprecated in favour of SAMP since about 2009, and this functionality, though still present, is of mostly historical interest.*

You can view and control operations relating to the PLASTIC hub from the **Interop** menu in the Control Window. It contains the following options:

### Register with PLASTIC

Attempts to locate a running PLASTIC hub and register with it.

### Unregister with PLASTIC

If currently registered with a PLASTIC hub, unregisters with it.

### Show registered applications

Pops up a small window which displays what applications are currently registered with any PLASTIC hub with which TOPCAT is registered. This will be kept up to date as applications register and unregister.

### Start internal PLASTIC hub

Starts a PLASTIC hub as part of the process within which TOPCAT is running. This will only work if a hub is not already running. The hub will terminate when TOPCAT exits.

### Start external PLASTIC hub

Starts a PLASTIC hub as a separate process. This will only work if a hub is not already running. The hub will continue running until it is explicitly terminated (e.g. using a `kill` command). Because this has some system-dependent features, it's not guaranteed to work, especially in non-Unix environments.

### Help on interoperability

Opens an appropriate help section in the help browser.



### Broadcast Table

Broadcasts the currently selected table to all listening applications using PLASTIC.



### Send Table to ...

Sends the currently selected table to a selected listening application using PLASTIC. Select the desired recipient application from the submenu.

## 9.3 Messages Transmitted

This section describes the messages which TOPCAT can transmit to other tools which understand the SAMP or PLASTIC protocol, and how to cause them to be sent. Approximately the same behaviour results for either SAMP or PLASTIC, except as noted.

In most cases you can choose two ways to transmit a message from TOPCAT:



### Broadcast

Broadcasts the message to all other applications currently registered with the hub which understand that message.



### Send

Sends the message to a single application which you select. The suitable applications (ones which are registered with the hub and claim to understand that message) are listed and you can choose one.

Examining the list of applications in the **Send** menu gives you an indication of which ones a **Broadcast** would broadcast to. Note however that just because an application appears in this list doesn't necessarily mean it will do something substantial with the message, for instance some applications register with the hub just to monitor traffic. In general the **Broadcast** and **Send** actions will be disabled (greyed-out) if TOPCAT is not registered with a hub, or if there are no applications listening which claim to support the relevant message.

Below is a list of places you can cause TOPCAT to transmit messages. The SAMP *MTypes* and PLASTIC *message IDs* are listed along with the descriptions; unless you are a tool developer you can probably ignore these.

### Transmit Table

The Control Window's **Interop** menu provides **Broadcast Table** and **Send Table** options which cause the currently selected table to be transmitted to other listening applications. They are invited to load the table in its current ("apparent") form. The **Broadcast** action is also available in the toolbar.

SAMP MTypes: `table.load.votable` or `table.load.fits`

PLASTIC            Message            IDs:            `ivo://votech.org/votable/load`            or  
`ivo://votech.org/votable/loadFromURL`

### Transmit Subset

The Subset Window (Appendix A.3.4)'s **Interop** menu contains **Broadcast Subset** and **Send Subset** options. These cause the selected subset to be sent to other listening applications (these actions are only available when one of the subsets is currently selected). Applications are invited to highlight the rows corresponding to that subset. Note this will only have an effect if the other application(s) are displaying the table that this subset relates to. This will be the case in one of two situations: (1) the table has been loaded from the same URL/filename by the other tool(s) or (2) the other tool(s) have acquired the table because it has already been

broadcast using SAMP/PLASTIC.

Also, whenever a new subset is created, for instance by entering an algebraic expression or tracing out a region on a plot (see Section 3.1.1), you have the option of transmitting the subset directly to one or all listening applications as an alternative to adding the new subset to the table's subset list.

SAMP MType: `table.select.rowList`

PLASTIC Message ID: `ivo://votech.org/votable/showObjects`

### Transmit Row

The ControlWindow has a **Broadcast Row** checkbox by the **Activation Action** selector. If this is selected, then any time you activate a row (e.g. by clicking on the corresponding point in a plot) other applications who are looking at the same table are invited to highlight that row in their copy of the table. As for **Transmit Subset** above, this will only have an effect if the other application(s) are displaying the same table.

A similar facility is available as the **Transmit Row** option in the Activation Actions (Section 8) window itself - the checkbox is just a convenience for a commonly-used option.

SAMP MType: `table.highlight.row`

PLASTIC Message ID: `ivo://votech.org/votable/highlightObject`

### Transmit Coordinates

Another Activation Action (Section 8) is **Transmit Coordinates**. In this case when you activate a row (e.g. by clicking on the corresponding point in a plot) other applications are invited to point out the sky position corresponding to the row which is activated, for instance by placing a cursor over it. The table columns which correspond to Right Ascension and Declination must be selected for this to work.

SAMP MType: `coord.pointAt.sky`

PLASTIC Message ID: `ivo://votech.org/sky/pointAtCoords`

### Transmit Image

The Density Plot (Appendix A.5.7) produces a 2-d histogram which is actually an image. Its **Interop** menu provides **Broadcast Image** and **Send Image** options which will send this (as a FITS image) to other applications which can display it in some way. This is a useful supplement to the facilities of the Density Plot window, since it doesn't have very sophisticated image display features (variable colour maps, contour plots etc).

Also, the **View URL as Image** Activation Action option allows you to identify a column which contains an image URL and to send it to a suitable image viewer or viewers.

SAMP MType: `image.load.fits`

PLASTIC Message ID: `ivo://votech.org/fits/image/loadFromURL`

### Transmit Spectrum

The **View URL as Spectrum** Activation Action option allows you to identify a column which contains a spectrum URL and to send it to a suitable spectrum viewer or viewers.

SAMP MType: `spectrum.load.ssa-generic`

PLASTIC Message ID: `ivo://votech.org/spectrum/loadFromURL`

### Transmit Resource List

The Registry Query Panel present in most of the Virtual Observatory windows allows you to send lists of VO registry resource identifiers to other applications which can make use of them. Note this only works in SAMP mode, not for PLASTIC.

SAMP MTypes: `voresource.loadlist.cone,` `voresource.loadlist.siap,`  
`voresource.loadlist.ssap`

## 9.4 Messages Received

This section describes the messages which TOPCAT will respond to when it receives them from other applications via the SAMP/PLASTIC hub. The SAMP MTypes and PLASTIC message IDs are listed along with the descriptions; unless you are a tool developer you can probably ignore these.

### Load Table

Loads a table sent by another application. It is added to the list of tables in the Control Window.

SAMP MType: `table.load.votable`, `table.load.fits`, `table.load.cdf` or `table.load.stil`

PLASTIC Message ID: `ivo://votech.org/votable/load` or `ivo://votech.org/votable/loadFromURL`

Note the non-standard MType `table.load.stil` is supported for loading tables with SAMP. This is like the other `table.load.*` MTypes, but any table format supported by the application is permitted. This MType has an additional parameter "format", which must contain the table format name (not case sensitive).

### New Subset

Loads a row selection sent from another application. If an external application is looking at the same table as one that TOPCAT has loaded, it can send a row selection. In this case, TOPCAT will add that selection as a new Row Subset for the table. The name of the received subset will be that of the sending application, for instance "Aladin". If a subset of that name already exists (which is probably because the same application has sent a row selection previously) then its content will be overwritten by the new selection. In either case, receiving the message causes plots displaying data from the table in question to show the points from the new subset.

Note: this behaviour differs from the behaviour in TOPCAT versions prior to v3.0. Different options for handling exactly how a received row selection is treated may be made available in future versions.

SAMP MType: `table.select.rowList`

PLASTIC Message ID: `ivo://votech.org/votable/showObjects`

### Highlight Row

If TOPCAT already has loaded the table identified by the request, it will activate (Section 8) the requested row. This will result in the row being marked in any currently visible plots and in the Data Window if it is visible, as well as performing any additional actions you have configured in the Activation Window.

SAMP MType: `table.highlight.row`

PLASTIC Message ID: `ivo://votech.org/votable/highlightObject`

### Load Resource Identifiers

Receives a list of VO registry resource identifiers from another application. Several of TOPCAT's Virtual Observatory access windows display a list of resources in their Registry Query Panel. Normally, the displayed resources are a default set or are those you have selected by entering keywords. However, another application can send a list of resources, and if they are of an appropriate type for the window in question, and if the window is currently visible, its currently displayed list will be replaced with the ones which have been sent.

SAMP MTypes: `voresource.loadlist`, `voresource.loadlist.cone`, `voresource.loadlist.siap`, `voresource.loadlist.ssap`

### Get Table

Allows clients to retrieve tables currently loaded into TOPCAT. This MType has a mandatory parameter `format` giving the (case-insensitive) table format name required for the output table. It also takes an optional integer-like parameter `id`, giving the ID number (as shown in the Control Window table list) of the table required. If `id` is missing or non-positive, the current table is used. The returned response has an output parameter `url` giving the URL of the apparent table in question. This MType is experimental and may be modified or withdrawn in the future.

SAMP MType: `table.get.stil`.

System-level messages are also responded to. For SAMP these are:

- `samp.hub.disconnect`
- `samp.hub.event.shutdown`
- `samp.hub.event.register`
- `samp.hub.event.unregister`
- `samp.hub.event.metadata`
- `samp.hub.event.subscriptions`
- `samp.app.ping`

and for PLASTIC:

- `ivo://votech.org/test/echo`
- `ivo://votech.org/info/getName`
- `ivo://votech.org/info/getIconURL`
- `ivo://votech.org/hub/event/ApplicationRegistered`
- `ivo://votech.org/hub/event/ApplicationUnregistered`
- `ivo://votech.org/hub/event/HubStopping`



## 10 Invoking TOPCAT

Starting up TOPCAT may just be a case of typing "topcat" or clicking on an appropriate icon and watching the Control Window pop up. If that is the case, and it's running happily for you, you can probably ignore this section. What follows is a description of how to start the program up, and various command line arguments and configuration options which can't be changed from within the program. Some examples are given in Section 10.5. Actually obtaining the program is not covered here; please see the TOPCAT web page <http://www.starlink.ac.uk/topcat/>.

There are various ways of starting up TOPCAT depending on how (and whether) it has been installed on your system; some of these are described below.

There may be some sort of short-cut icon on your desktop which starts up the program - in this case just clicking on it will probably work. Failing that you may be able to locate the jar file (probably named `topcat.jar`, `topcat-full.jar` or `topcat-lite.jar`) and click on that. These files would be located in the `java/lib/topcat/` directory in a standard Starjava installation. Note that when you start by clicking on something you may not have the option of entering any of the command line options described below - to use these options, which you may need to do for serious use of the program, you will have to run the program from the command line.

Alternatively you will have to invoke the program from the command line. On Unix-like operating systems, you can use the `topcat` script. If you have the full starjava installation, this is probably in the `starjava/java/bin` directory. If you have one of the standalone jar files (`topcat-full.jar` or `topcat-lite.jar`), it should be in the same directory as it/them. If it's not there, you can unpack it from the jar file itself, using a command like `unzip topcat-lite.jar topcat`. If that directory (and java) is on your path then you can write:

```
topcat [java-args] [topcat-args]
```

In this case any arguments which start `-D` or `-X` are assumed to be arguments to the java command, any arguments which start `-J` are stripped of the `-J` and then passed as arguments to the java command, a `-classpath path` defines a class path to be used in addition to the TOPCAT classes, and any remaining arguments are used by TOPCAT.

If you're not running Unix then to start from the command line you will have to use the `java` command itself. The most straightforward way of doing this will look like:

```
java [java-args] -jar path/to/topcat.jar [topcat-args]
```

(or the same for `topcat-full.jar` etc). However **NOTE**: using java's `-jar` flag ignores any other class path information, such as the `CLASSPATH` environment variable or java's `-classpath` flag - see Section 10.2.1.

Note that Java Web Start can also be used to invoke the program without requiring any prior download/installation - sorry, this isn't documented properly here yet.

The meaning of the optional `[topcat-args]` and `[java-args]` sequences are described in Section 10.1 and Section 10.2 below respectively.

### 10.1 TOPCAT Command-line Arguments

You can start TOPCAT from the command line with no arguments - in this case it will just pop up the command window from which you can load in tables. However you may specify flags and/or table locations and formats.

If you invoke the program with the `"-help"` flag you will see the following usage message:

Usage: topcat <flags> [[-f <format>] <table> ...]

General flags:

-help	print this message and exit
-version	print component versions etc and exit
-verbose	increase verbosity of reports to console
-debug	add debugging information to console log messages
-demo	start with demo data
-running	use existing TOPCAT instance if one is running
-memory	use memory storage for tables
-disk	use disk backing store for large tables
-samp	use SAMP for tool interoperability
-plastic	use PLASTIC for tool interoperability
-[no]hub	[don't] run internal SAMP/PLASTIC hub
-exthub	run external SAMP/PLASTIC hub
-noserv	don't run any services (PLASTIC or SAMP)
-nocheckvers	don't check latest version
-stilts <args>	run STILTS not TOPCAT
-jsamp <args>	run JSAMP not TOPCAT

Useful Java flags:

-classpath jar1:jar2..	specify additional classes
-XmxnnnM	use nnn megabytes of memory
-Dname=value	set system property

Auto-detected formats:

fits-plus, colfits-plus, colfits-basic, fits, votable, cdf, gbin

All known formats:

fits-plus, colfits-plus, colfits-basic, fits, votable, cdf, gbin, ascii, csv, tst, ipa

Useful system properties (-Dname=value - lists are colon-separated):

java.io.tmpdir	temporary filespace directory
jdbc.drivers	JDBC driver classes
jel.classes	custom algebraic function classes
jel.classes.activation	custom action function classes
star.connectors	custom remote filestore classes
startable.load.dialogs	custom load dialogue classes
startable.readers	custom table input handlers
startable.writers	custom table output handlers
startable.storage	storage policy
mark.workaround	work around mark/reset bug
(see topcat -jsamp -help for more)	

The meaning of the flags is as follows:

**-f <format>**

Signifies that the file(s) named after it on the command line are in a particular file format. Some file formats (VOTable, FITS) can be detected automatically by TOPCAT, but others (including Comma-Separated Values) cannot - see Section 4.1.1. In this case you have to specify with the -f flag what format the named files are in. Any table file on the command line following a -f <format> sequence must be in the named format until the next -f flag. The names of both the auto-detected formats (ones which don't need a -f) and the non-auto-detected formats (ones which do) are given in the usage message you can see by giving the -help flag (this message is shown above). You may also use the classname of a class on the classpath which implements the `TableBuilder` interface - see SUN/252.

**-help**

If the -help (or -h) flag is given, TOPCAT will write a usage message and exit straight away.

**-version**

If the -version flag is given, TOPCAT will print a summary of its version and the versions and availability of some its components, and exit straight away.

**-verbose**

The -verbose (or -v) flag increases the level of verbosity of messages which TOPCAT writes to standard output (the console). It may be repeated to increase the verbosity further. The messages it controls are currently those written through java's standard logging system - see

the description of the Log Window (Appendix A.10.12) for more information about this.

#### **-debug**

The `-debug` flag affects how logging messages appear (whether they appear is affected by the `-verbose` flag). If present, these messages will provide more detail about where each message was generated from.

#### **-demo**

The `-demo` flag causes the program to start up with a few demonstration tables loaded in. You can use these to play around with its facilities. Note these demo tables are quite small to avoid taking up a lot of space in the installation, and don't contain particularly sensible data, they are just to give an idea.

#### **-running**

The `-running` flag can be used when specifying tables on the command line. If an existing instance of TOPCAT is already running, the tables are loaded into it. If no instance of TOPCAT is running (or at least one cannot be discovered), then the `-running` flag has no effect and a new instance is started up as usual.

#### **-memory**

If the `-memory` flag is given then the program will store table data in memory, rather than the default which is to store small tables in memory, and larger ones in temporary disk files.

#### **-disk**

If the `-disk` flag is given then the program will store table data in temporary disk files, rather than the default which is to store small tables in memory, and larger ones on disk. If you get out of memory messages, running with the `-disk` flag might help, though the default policy should work fairly well. The temporary data files are written in the default temporary directory (defined by the `java.io.tmpdir` system property - often `/tmp` - and deleted when the program exits, unless it exits in an unusual way).


#### **-samp**

Forces TOPCAT to use SAMP for tool interoperability (see Section 9). SAMP rather than PLASTIC is the default, so this flag has no effect.

#### **-plastic**

Forces TOPCAT to use PLASTIC instead of SAMP for tool interoperability (see Section 9).

#### **-[no]hub**

The `-hub` flag causes TOPCAT to run an internal SAMP or PLASTIC hub (SAMP by default), if no hub is already running, and the `-nohub` flag prevents this from happening. The default behaviour, when neither of these flags is given, is to start a hub automatically for SAMP but not for PLASTIC. The hub will terminate when TOPCAT does, or can be shut down manually with the **Interop|Stop Internal Hub** ( 

#### **-exthub**

Causes TOPCAT to attempt to run an external SAMP or PLASTIC hub, if no hub is already running. The hub will show up in its own window, and can be stopped by closing the window. The hub will continue to run after TOPCAT terminates. Invoking external processes from Java is inherently unreliable, so this is done on a best-efforts basis. See Section 9.

#### **-noserv**

Prevents TOPCAT from running any services. Currently the services which it may run are SAMP or PLASTIC (see above).

#### **-checkvers**

Prevents TOPCAT from checking an external URL so it can determine whether the latest version is being run.

#### **-stilts <stilts-args>**

This flag is a convenience which allows you to run the STILTS command-line tool instead of

TOPCAT. This is possible because TOPCAT is built on top of STILTS and contains a superset of its code. See the STILTS manual (or `topcat -stilts -help`) for the form of the `<stilts-args>`.

### **-jsamp <jsamp-args>**

This flag is a convenience which allows you to run the jsamp command-line tool, from the JSAMP package, instead of TOPCAT. This is possible because TOPCAT contains the JSAMP library. JSAMP provides various SAMP-related utilities, such as a freestanding hub and diagnostic tools. See the JSAMP documentation, or do `topcat -jsamp -help` for more information.

Other arguments on the command line are taken to be the locations of tables. Any tables so specified will be loaded into TOPCAT at startup. These locations are typically filenames, but could also be URLs or SQL queries, or perhaps something else. In addition they may contain "fragment identifiers" (with a "#") to locate a table within a given resource, so that for instance the location

```
/my/data/cat1.fits#2
```

means the second extension in the multi-extension FITS file `/my/data/cat1.fits`. Section 4.2 describes in more detail the kinds of URLs which can be used here.

Note that options to Java itself may also be specified on the command-line, as described in the next section.

## **10.2 Java Options**

As described above, depending on how you invoke TOPCAT you may be able to specify arguments to Java itself (the "Java Virtual Machine") which affect how it runs. These may be defined on the command line or in some other way. The following subsections describe how to control Java in ways which may be relevant to TOPCAT; they are however somewhat dependent on the environment you are running in, so you may experience OS-dependent variations.

### **10.2.1 Class Path**

The classpath is the list of places that Java looks to find the bits of compiled code that it uses to run an application. When running TOPCAT this always has to include the TOPCAT classes themselves - this is part of the usual invocation and is described in Section 10. However, for certain things Java might need to find some other classes, in particular for:

- JDBC drivers (see Section 10.3)
- Providing extended algebraic functions (see Section 7.9)
- Launching Mirage (see Section 4.1.2.12)
- Installing I/O handlers for new table formats (see Section 4.1.3)

If you are going to use these facilities you will need to start the program with additional class path elements that point to the location of the classes required. How you do this depends on how you are invoking TOPCAT. If you are using the `topcat` startup script, you can write:

```
topcat -classpath other-paths ...
```

(this adds the given paths to the standard ones required for TOPCAT itself). If you are invoking java directly, then you can either write on the command line:

```
java -classpath path/to/topcat.jar:other-paths
    uk.ac.starlink.topcat.Driver ...
```

or set the CLASSPATH environment variable something like this:

```
setenv CLASSPATH path/to/topcat.jar:other-paths
```

In any case, multiple (extra) paths should be separated by colons in the *other-paths* string.

**Note** that if you are running TOPCAT using java's `-jar` flag, any attempt you make to specify the classpath will be ignored! This is to do with Java's security model. If you need to specify a classpath which includes more than the TOPCAT classes themselves, you can't use `java -jar` (use `java -classpath topcat-lite.jar:... uk.ac.starlink.topcat.Driver` instead).

### 10.2.2 Memory Size

If TOPCAT fails during operation with a message that says something about a `java.lang.OutOfMemoryError`, then your heap size is too small for what you are trying to do. You will have to run java with a bigger heap size using the `-Xmx` flag. Invoking TOPCAT from the `topcat` script you would write something like:

```
topcat -Xmx256M ...
```

or using java directly:

```
java -Xmx256M ...
```

which means use up to 256 megabytes of memory (don't forget the "M" for megabyte). JVMs often default to a heap size of 64M. You probably don't want to specify a heap size larger than the physical memory of the machine that you are running on.

There are other types of memory and tuning options controlled using options of the form `-X<something-or-other>`; if you're feeling adventurous you may be able to find out about these by typing `"java -X"`.

**Note** however: using the `-disk` flag described in Section 10.1 may be a better solution; this makes the program store data from large tables on disk rather than in memory.

### 10.2.3 System properties

System properties are a way of getting information into Java (they are the Java equivalent of environment variables). The following ones have special significance within TOPCAT:

#### **apple.laf.useScreenMenuBar**

On the Apple Macintosh platform only, this property controls whether menus appear at the top of the screen as usual for Mac, or at the top of individual windows as usual for Java. By default it is set to `true` for TOPCAT, so menus mostly appear at the top of the screen (though it's not true to say that TOPCAT obeys the Mac look and feel completely); if you prefer the more Java-like look and feel, set it to `false`.

#### **http.proxyHost**

If you are operating inside a firewall which prohibits direct HTTP connections, you can set this to the name of an HTTP proxy server in order to access remote servers as required for VO functionality etc. There are a number of related system properties which you may or may not need to use in this case: `http.proxyPort`, `https.proxyHost` etc. See the appropriate java documentation (e.g. by googling for "http.proxyHost") for details.

#### **java.io.tmpdir**

The directory in which TOPCAT will write any temporary files it needs. This is usually only done if the `-disk` flag has been specified (see Section 10.1).

#### **jdbc.drivers**

Can be set to a (colon-separated) list of JDBC driver classes using which SQL databases can

be accessed (see Section 10.3).

#### **jel.classes**

Can be set to a (colon-separated) list of classes containing static methods which define user-provided functions for synthetic columns or subsets. (see Section 7.9).

#### **jel.classes.activation**

Can be set to a (colon-separated) list of classes containing static methods which define user-provided functions for use in custom activation expressions. (see Section 7.9).

#### **jsamp.hub.profiles**

This property determines what profiles a SAMP hub will use if you run an internal or external hub from TOPCAT (either with the `-hub/-exthub` flag or from the GUI). The value is a comma-separated list of profile specifiers; options are "std" for Standard Profile, "web" for Web Profile or the name of a class implementing the `org.astrogrid.samp.hub.HubProfile` interface. The default setting runs just a Standard Profile hub, but, for instance, setting it to "std,web" would run a Web Profile as well. Note you should include `std` in the list, otherwise TOPCAT will not be able to talk to the hub. See the JSAMP documentation for more detail.

#### **jsamp.localhost**

Sets the hostname by which the local host is to be identified in URLs, for instance server endpoints. If unset, the default is currently the loopback address "127.0.0.1". However, if this property is set (presumably to the local host's fully- or partly-qualified domain name) its value will be used instead. Two special values may also be set: "[hostname]" for the host's fully qualified domain name, and "[hostnumber]" for the IP number.

#### **jsamp.server.port**

Gives a preferred port number on which to open TOPCAT's internal HTTP server, used for SAMP communications amongst other things. If this port is unavailable, some other port will be used instead.

#### **jsamp.xmlrpc.impl**

Indicates which pluggable XML-RPC implementation should be used for SAMP communications. By default an internal implementation is used. If it is set to "xml-log" or "rpc-log" then all XML-RPC communications will be logged in very or fairly verbose terms respectively to standard output. The classname of an `org.astrogrid.samp.xmlrpc.XmlRpcKit` implementation may be given instead to use a custom implementation.

#### **lut.files**

Can be set to a (colon-separated on \*nix, semicolon-separated on Windows) list of files giving custom lookup tables for auxiliary axis and density map colour maps. Each file should be a text file containing a number of space-separated lines, each containing red, green, blue samples in the range 0-1. For instance the file

```
1.0  1.0  0.0
1.0  0.0  1.0
```

would give a colour map that fades from yellow to magenta. Any number of samples may be given; the scale is interpolated.

#### **mark.workaround**

If set to "true", this will work around a bug in the `mark()/reset()` methods of some java `InputStream` classes. These are rather common, including in Sun's J2SE system libraries. Use this if you are seeing errors that say something like "Resetting to invalid mark". Currently defaults to "false".

#### **myspace.cache**

If set to "true", this will prevent directories in the MySpace file browser from being read more than once. This is a workaround for MySpace performance problems at time of writing (April 2006); setting it true may lead to out of date file listings in MySpace, but it may be much

faster. This behaviour may be withdrawn in future versions if MySpace performance improves. Currently defaults to "false".

**service.maxparallel**

Raises the maximum number of concurrent queries that may be made during a multi-cone operation. You should only increase this value **with great care** since you risk overloading servers and becoming unpopular with data centres. As a rule, you should only increase this value if you have obtained permission from the data centres whose services on which you will be using the increased parallelism.

**star.basicauth.user****star.basicauth.password**

If set, these will provide username and password for HTTP Basic Authentication. Any time the application attempts to access an HTTP URL and is met by a 401 Unauthorized response, it will try again supplying these user credentials. This is a rather blunt instrument, since the same identity is supplied regardless of which URL is being accessed, but it may be of some use in accessing basic-authentication protected services.

If these properties are not set, then on encountering a 401 in the application will pop up a dialogue window asking for username and password.

**star.connectors**

Can be set to a (colon-separated) list of classes which provide access to remote filesystem implementations. Thus-named classes should implement the `uk.ac.starlink.connect.Connector` interface which specifies how you can log on to such a service and provides a hierarchical view of the filesystem it contains.

**startable.load.dialogs**

Can be set to a (colon-separated) list of custom table load dialogue classes. Briefly, you can install your own table import dialogues at runtime by providing classes which implement the `uk.ac.starlink.table.gui.TableLoadDialog` interface and naming them in this property. See STIL documentation for more detail.

**startable.readers**

Can be set to a (colon-separated) list of custom table format input handler classes (see Section 4.1.3).

**startable.storage**

Can be set to determine the default storage policy. Setting it to "disk" has basically the same effect as supplying the "-disk" argument on the TOPCAT command line (see Section 10.1). Other possible values are "adaptive", "memory", "sideways" and "discard"; see SUN/252. The default is "adaptive", which means storing smaller tables in memory, and larger ones on disk.

**startable.unmap**

Determines whether and how unmapping of memory mapped buffers is done. Possible values are "sun" (the default) or "none". In most cases you are advised to leave this alone, but in the event of unmapping-related JVM crashes (not expected!), setting it to `none` may help.

**startable.writers**

Can be set to a (colon-separated) list of custom table format output handler classes (see Section 4.1.3).

**topcat.extttools**

Defines extension tools to be used by TOPCAT. The value is a colon-separated list of class names, each of which must implement the `uk.ac.starlink.topcat.TopcatToolAction` interface and have a no-arg constructor. The actions corresponding to any such classes will be added to toolbar. This is an experimental extensibility feature, which may be modified or withdrawn in a future release.

**user.dir**

Sets the current working directory. This determines the default from which the file browsers will start.

**votable.namespacing**

Determines how namespacing is handled in input VOTable documents. Known values are "none" (no namespacing, xmlns declarations in VOTable document will probably confuse parser), "lax" (anything that looks like it is probably a VOTable element will be treated as a VOTable element) and "strict" (VOTable elements must be properly declared in one of the correct VOTable namespaces). May also be set to the classname of a `uk.ac.starlink.votable.Namespacing` implementation. The default is "lax".

**votable.strict**

Set true for strict enforcement of the VOTable standard when parsing VOTables. This prevents the parser from working round certain common errors, such as missing `arraysize` attributes on FIELD/PARAM elements with `datatype="char"`. False by default.

**votable.version**

Selects the version of the VOTable standard which output VOTables will conform to by default. May take the values "1.0", "1.1", "1.2" or "1.3". By default, version 1.2 VOTables are written.

To define these properties on the command line you use the `-D` flag, which has the form

```
-D<property-name>=<value>
```

If you're using the TOPCAT startup script, you can write something like:

```
topcat -Djdbc.drivers=org.postgresql.Driver ...
```

or if you're using the java command directly:

```
java -Djdbc.drivers=org.postgresql.Driver ...
```

Alternatively you may find it more convenient to write these definitions in a file named `.starjava.properties` in your home directory; the above command-line flag would be equivalent to inserting the line:

```
jdbc.drivers=org.postgresql.Driver
```

in your `.starjava.properties` file.

### 10.3 JDBC Configuration

This section describes additional configuration which must be done to allow TOPCAT to access SQL-compatible relational databases for reading (see Section 4.1.1.10) or writing (see Section 4.1.2.9) tables. If you don't need to talk to SQL-type databases, you can ignore the rest of this section. The steps described here are the standard ones for configuring JDBC (which sort-of stands for Java Database Connectivity), described in more detail on Sun's JDBC web page.

To use TOPCAT with SQL-compatible databases you must:

- Have access to an SQL-compatible database locally or over the network
- Have a JDBC driver appropriate for that database
- Install this driver for use with TOPCAT
- Know the format the driver uses for URLs to access database tables
- Have appropriate privileges on the database to perform the desired operations

Installing the driver consists of two steps:



1. Set the `jdbc.drivers` system property to the name of the driver class as described in Section 10.2.3
2. Ensure that the classpath you are using includes this driver class as described in Section 10.2.1

These steps are all standard for use of the JDBC system.

To the author's knowledge, TOPCAT has so far successfully been used with the following RDBMSs and corresponding JDBC drivers:

### MySQL

MySQL has been tested on Linux with the Connector/J driver and seems to work; tested versions are server 3.23.55 with driver 3.0.8 and server 4.1.20 with driver 5.0.4. Sometimes tables with very many (hundreds of) columns cannot be written owing to SQL statement length restrictions. Note there is known to be a column metadata bug in version 3.0.6 of the driver which can cause a `ClassCastException` error when tables are written. Check the driver's documentation for additional parameters, for instance `"useUnicode=true&characterEncoding=UTF8"` may be required to handle some non-ASCII characters.

### PostgreSQL

PostgreSQL 7.4.1 apparently works with its own driver. Note the performance of this driver appears to be rather poor, at least for writing tables.

### Oracle

You can use Oracle with the JDBC driver that comes as part of its Basic Instant Client Package. However, you *can't* currently use the SQL load/SQL save dialogue boxes to do it. You have to specify a JDBC URL specifying the query to read/table to write as a string in the **Location** field of the normal table load/save dialogue boxes. The URL will look something like

```
jdbc:oracle:thin:@//hostname:1521/database#SELECT ...
```

for querying an existing database (loading) and

```
jdbc:oracle:thin:@//hostname:1521/database#new-table-name
```

for writing a new table (saving).

### SQL Server

There is more than one JDBC driver known to work with SQL Server, including jTDS and the Microsoft JDBC driver. Some evidence suggests that jTDS may be the better choice, but your mileage may vary.

### Sybase ASE

There has been a successful use of Sybase 12.5.2 and jConnect (jconn3.jar) using a JDBC URL like `"jdbc:sybase:Tds:hostname:port/dbname?user=XXX&password=XXX#SELECT..."`. An earlier attempt using Sybase ASE 11.9.2 failed.

Other RDBMSs and drivers may or may not work - please let us know the results of any experiments you carry out. Sun maintain a list of JDBC drivers for various databases; it can be found at <http://servlet.java.sun.com/products/jdbc/drivers>.

Here are example command lines to start up TOPCAT using databases known to work.

### PostgreSQL

```
java -classpath topcat-full.jar:pg73jdbc3.jar \
-Djdbc.drivers=org.postgresql.Driver \
uk.ac.starlink.topcat.Driver
```

## MySQL

```
java -classpath topcat-full.jar:mysql-connector-java-3.0.8-bin.jar \
-Djdbc.drivers=com.mysql.jdbc.Driver \
uk.ac.starlink.topcat.Driver
```

## Oracle

```
java -classpath topcat-full.jar:ojdbc14.jar \
-Djdbc.drivers=oracle.jdbc.driver.OracleDriver \
uk.ac.starlink.topcat.Driver
```

## SQL Server with jTDS

```
java -classpath topcat-full.jar:jtds-1.1.jar \
-Djdbc.drivers=net.sourceforge.jtds.jdbc.Driver \
uk.ac.starlink.topcat.Driver
```

## 10.4 Tips for Large Tables

Considerable effort has gone into making TOPCAT capable of dealing with large datasets. In particular it does not in general have to read entire files into memory in order to do its work, so it's not restricted to using files which fit into the java virtual machine's 'heap memory' or into the physical memory of the machine. As a rule of thumb, the program will work at reasonable speed with tables up to about 1-10 million rows, depending on the machine it's running on. The number of columns is less of an issue, though see below concerning performance.

However, the way you invoke the program affects how well it can cope with large tables; you may in some circumstances get a message that TOPCAT has run out of memory (either a popup or a terse "OutOfMemoryError" report on the console), and there are some things you can do about this:

### Increase Java's heap memory

When a Java program runs, it has a fixed maximum amount of memory that it will use; Java does not really make use of virtual memory. The default maximum is typically 64Mb, though that depends on your java setup. You can increase this by using the `-Xmx` flag, followed by the maximum heap memory, for instance `"topcat -Xmx256M"` or `"java -Xmx256M -jar topcat-full.jar"`. Don't forget the "M" to indicate megabytes. It's generally reasonable to increase this value up to nearly the amount of free physical memory in your machine if you need to (taking account of the needs of other processes running at the same time) but attempting any more will usually result in abysmal performance. See Section 10.2.2.

### Use FITS files

Because of the way that FITS files are organised, TOPCAT is able to load tables which are stored as uncompressed FITS binary tables on disk almost instantly regardless of their size (in this case loading just reads the metadata, and any data elements are only read if and when they are required). So if you have a large file stored in VOTable or ASCII-based form which you use often and takes a long time to load, it's a good idea to convert it to FITS format once for subsequent use. You can do this either by loading it into TOPCAT once and saving it as FITS, or using the separate command-line package STILTS. Note that the "fits-plus" variant which TOPCAT uses by default retains all the metadata that would be stored in a corresponding VOTable, so you won't normally lose information by doing this (see Section 4.1.2.1). As well as speeding things up, using FITS files will also reduce the need to use `-Xmx` flags as above.

### Run in 64-bit mode

If you are working with a file or files whose total size approaches or exceeds about 2 Gbyte,

you should use a 64-bit version of java. This means that you will need a 64-bit operating system, and also a 64-bit version of the Java Virtual Machine. Executing "java -version" (or "topcat -version") will probably say something about 64-bit-ness if it is 64-bit.

### Use column-oriented storage

For really large tables storing them in the colfits output format can significantly improve performance. This stores all the elements of a single column contiguously on disk, which means that scanning through all the values in one or a few columns can proceed with much less unnecessary I/O than in normal (row-oriented) FITS format. It will make most difference when the table is larger than the amount of physical memory available, and the table has many columns. Be aware however that operations which require all the cells in all the rows (for instance, calculating row statistics) may be somewhat slower using this format.

It is also possible to use column-oriented storage for non-FITS files by specifying the flag `-Dstartable.storage=sideways`. This is like using the `-disk` flag but uses column-oriented rather than row-oriented temporary files. However, using it for such large files means that the conversion is likely to be rather slow, so you may be better off converting the original file to colfits format in a separate step and using that.

### Use the -disk flag

The way TOPCAT stores table data is configurable, and the details can be controlled by setting its *Storage Policy*. The default storage policy (since version 3.5), "adaptive", means that the data for relatively small tables are stored in memory, and for larger ones in temporary disk files. This usually works fairly well and you're not likely to need to change it. However, you can experiment if you like, and a small amount of memory may be saved if you encourage it to store all table data on disk, by specifying the `-disk` flag on the command line. You can achieve the same effect by adding the line `startable.storage=disk` in the `.starjava.properties` in your home directory. See Section 10.1, Section 10.2.3.

As far as performance goes, the memory size of the machine you're using does make a difference. If the size of the dataset you're dealing with (this is the size of the FITS HDU if it's in FITS format but may be greater or less than the file size for other formats) will fit into unused physical memory then generally everything will run very quickly because the operating system can cache the data in memory; if it's larger than physical memory then the data has to keep being re-read from disk and most operations will be much slower, though use of column-oriented storage can help a lot in that case.

## 10.5 Examples

Here are some examples of invoking TOPCAT from the command line. In each case two forms are shown: one using the `topcat` script, and one using the jar file directly. In the latter case, the `java` command is assumed to be on the your path, and the jar file itself, assumed in directory `my/tcdir`, might be named `topcat.jar`, `topcat-full.jar`, or something else, but the form of the command is the same.

### No arguments

```
topcat
java -jar topcat.jar
```

### Output usage message

```
topcat -h
java -jar topcat.jar -h
```

### Load a FITS file

```
topcat testcat.fits
java -jar my/tcdir/topcat.jar testcat.fits
```

### Loading files of various formats

```
topcat t1.fits -f ascii t2.txt t3.txt -f votable t4.xml
java -jar my/tcdir/topcat.jar t1.fits -f ascii t2.txt t3.txt -f votable t4.xml
```

### Use disk storage format and boosted heap memory

```
topcat -Xmx256M -disk
java -Xmx256M -jar my/tcdir/topcat.jar -disk
```

### Make custom functions available

```
topcat -classpath my/funcdir/funcs.jar -Djel.classes=my.ExtraFuncs t1.fits
java -classpath my/tcdir/topcat.jar:my/funcdir/funcs.jar \
  -Djel.classes=func.ExtraFuncs \
  uk.ac.starlink.topcat.Driver t1.fits
```

### Make PostgreSQL database connectivity available

```
topcat -classpath my/jdbcdir/pg73jdbc3.jar -Djdbc.drivers=org.postgresql.Driver
java -classpath my/tcdir/topcat.jar:my/jdbcdir/pg73jdbc3.jar \
  -Djdbc.drivers=org.postgresql.Driver uk.ac.starlink.topcat.Driver
```


### Use custom I/O handlers

```
topcat -classpath my/driverdir/drivers.jar \
  -Dstartable.readers=my.MyTableBuilder \
  -Dstartable.writers=my.MyTableWriter \
java -classpath my/tcdir/topcat.jar:my/driverdir/drivers.jar \
  -Dstartable.readers=my.MyTableBuilder \
  -Dstartable.writers=my.MyTableWriter \
  uk.ac.starlink.topcat.Driver
```

The `-Dx=y` definitions can be avoided by putting equivalent `x=y` lines into the `.starjava.properties` in your home directory.

## A TOPCAT Windows

This appendix gives a tour of all the windows that form the TOPCAT application, explaining the anatomy of the windows and the various tools, menus and other controls. Attributes common to many or all windows are described in Appendix A.1, and the subsequent sections describe each of the windows individually.

When the application is running, the **Help For Window** () toolbar button will display the appropriate description for the window on which it is invoked.

### A.1 Common Window Features

This section describes some features which are common to many or all of the windows used in the TOPCAT program.

#### A.1.1 Toolbar

Each window has a toolbar at the top containing various buttons representing actions that can be invoked from the window. Most of them contain the following buttons:



**Close**

Closes the window. This convenience button has the same effect as closing the window in whatever way your graphics platform normally allows. In most cases, closing the window does not lose state associated with it (such as fields filled in); if you recover the window later it will look the same as when you closed it.



**Help**

Pops up a Help browser window, or makes sure it is visible if it has already been opened. The window will display help information relevant to the window in which you pushed this button.

Buttons in the toolbar often appear in menus of the same window as well; you can identify them because they have the same icon. This is a convenience; invoking the action from the toolbar or from the menu will have the same effect.

Often an action will only be possible in certain circumstances, for instance if some rows in the associated JTable have been selected. If the action is not possible (i.e. it would make no sense to invoke it) then the button in the toolbar and the menu option will be greyed out, indicating that it cannot be invoked in the current state.

#### A.1.2 Menus

Most windows have a menu bar at the top containing one or more menus. These menus will usually provide the actions available from the toolbar (identifiable because they have the same icons), and may provide some other less-commonly-required actions too.

Here are some of the menus common to several windows:

##### Window menu

Nearly all windows have this menu. At least the following options are available:



**Control Window**

Ensures that the Control Window is visible on the screen, deiconifying and raising it if necessary. This can be useful if you 'lose' the window behind a proliferation of other ones.



### Scrollable

If selected, this causes the entire content of the window to be wrapped in scrollbars. It is not generally recommended to use this option, since in general the windows are arranged so that resizing them will resize sensible parts of them, but it may be useful if using some of the larger windows on an unusually small screen.



### Close

Closes the window. This convenience button has the same effect as closing the window in whatever way your graphics platform normally allows. In most cases, closing the window does not lose state associated with it (such as fields filled in); if you recover the window later it will look the same as when you closed it.



### Exit

Exits TOPCAT. You will be prompted to confirm this action if tables are loaded, since it might result in loss of data.

## Help menu

Nearly all windows have this menu. The following options are available:



### Help

Pops up the Help Window.



### Help for Window

Pops up the Help Window; the window will display help information relevant to the window in which the menu appears.



### Help in Browser

Attempts to show the application help in a web browser. This is somewhat system dependent and is not guaranteed to work.



### Help in Browser (single page)

Attempts to show the application help in a web browser as a single (long page). This can be useful if you want to search for a given word or string in the text. This is somewhat system dependent and is not guaranteed to work.



### Help for Window in Browser

Attempts to show the help page relevant to the window in which the menu appears in a web browser. This is somewhat system dependent and is not guaranteed to work.



### About TOPCAT

Pops up a little window giving information on the version and authorship of the program. It also reports on availability of some optional components.

## Display menu

This menu is available for most windows which display their data using a JTable component. If present, it contains a list of the columns in the JTable with tickboxes next to them - clicking on a column name in this menu toggles whether the column is visible in the window.

### A.1.3 JTables

MU_ACOSD	MU_D	SIGMU_A	SIGMU_D
-40.194374	0.488800	10.700078	10.003148
-0.470101	7.433875	11.458762	9.757084
-31.978386	3.366049	86.749847	78.256889
5.828093	-16.916819	10.965803	11.553731
-43.659512	-33.585552	40.006008	37.997807
-6.994719	86.585754	11.505673	9.632096
-8.585894	5.851262	11.452976	10.542965
-1.857549	-7.694178	13.431982	12.487694
14.381025	-40.090954	15.317303	11.817266
3.195966	-4.708207	18.448475	18.928114

#### An example JTable

Many of the windows, including the Data Window, display their data in a visual component called a **JTable**. This displays a grid of values, with headings for each column, in a window which you can scroll around. Although JTables are used for a number of different things (for instance, showing the table data themselves in the Data Window and showing the column metadata in the Columns Window), the fact that the same widget is used provides a common look and feel.

Here are some of the things you can do with a JTable:

#### Scroll around

Using the scrollbars which may appear to the right and below the table you can scroll around it to see parts which are not initially visible. You can grab the sliders and drag them around by holding the mouse button down while you move it, or click in the slider "trough" one side or other of the current slider position to move a screenful. Under some circumstances the cursor arrow keys and PageUp/PageDown keys may move the table too. If the JTable is small enough to fit within the window the scrollbars may not appear.

#### Move columns

By clicking on the header (grey title bit at the top) of a column and dragging it left or right, you can change the order of columns as displayed. In some cases (the Data Window) this actually has the effect of changing the order of the columns in the table; in other cases it is just cosmetic.

#### Resize columns

By dragging on the line between row headers you can change the width of the columns in the table.

#### Edit cells

In some cases, cells are editable. If they are, then double-clicking in the cell will begin an edit session for that cell, and pressing Return will confirm that the edit has been made.

#### Select rows

Sometimes rows can be highlighted; you can select one row by clicking on it or a number of contiguous rows by clicking and dragging from the first to the last. To add further rows to a set already selected without deselecting the first lot, hold the "Control" key down while you do it.

#### Sort rows

In some, but not all cases, you may see a little up or down arrow in one of the column headers. This indicates that the rows of the table have been sorted according to the values in that column. If you click on another column header, the rows will be re-sorted according to the

values in that column. Clicking again on the same column reverses the sort order. This feature is only available for certain displayed JTables.

In some cases where a JTable is displayed, there will be a menu on the menu bar named **Display**. This permits you to select which columns are visible and which are hidden. Clicking on the menu will give you a list of all the available columns in the table, each with a checkbox by it; click the box to toggle whether the column is displayed or not.

#### A.1.4 Column Selector



##### Editable Column Selector

Several windows in TOPCAT invite you to select a column value from one of the loaded tables by presenting a selection box like the one in the figure. Examples are the Plot and Match windows.

In the simplest case you can simply select the value from the list of columns by clicking on the down-arrow at the right and then selecting one from the drop-down list of columns which is offered. Note that only appropriate columns will be offered - for instance if a numeric value is required, string-valued columns will not be included.

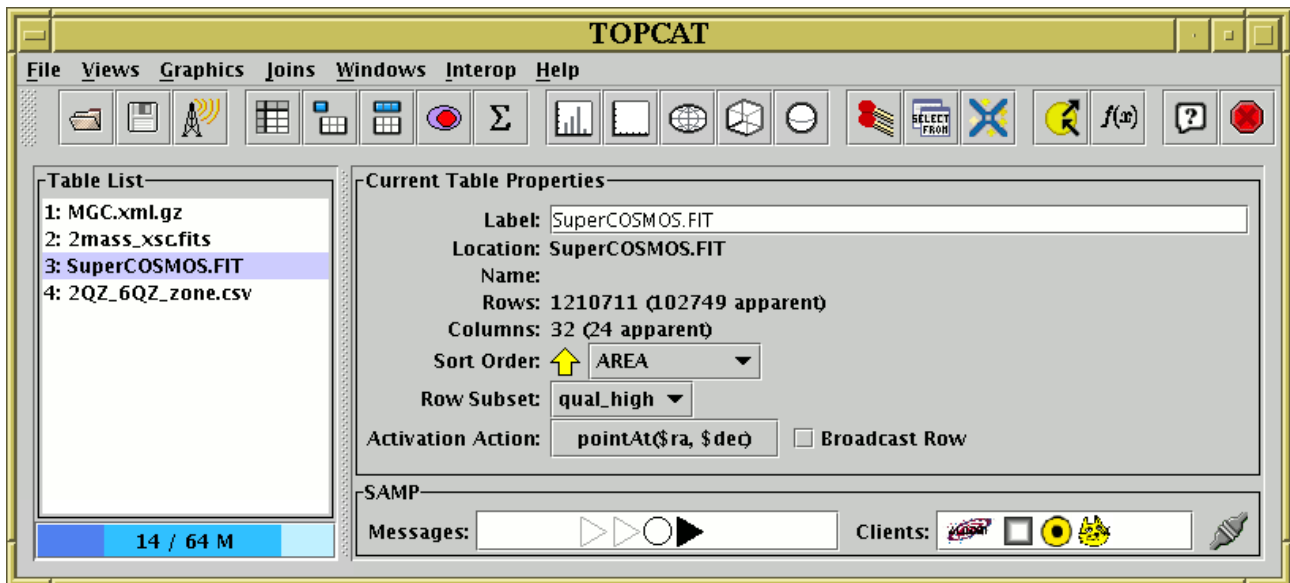
Another possibility is to use the little left/right arrows on the far right to cycle through all the known columns. This can be useful in plots, for instance to see a sequence of all the available plots against a given abscissa using only one click at a time.

Finally, you can type in an expression giving the required value. This can either be the name of a column just as if you had selected it from the drop-down list, or an expression based on column names, or even a constant value. The syntax of the expression language is explained in Section 7. Typing the column name rather than selecting it may be more convenient if the selection list is very long; typing an expression obviously gives you a lot more possibilities.

Note that depending on your platform the selection box may not look exactly like the figure above. However, if you can type into it (probably by clicking on it) then you should be able to use expressions as described above. Some selectors however only take column names; if you can't type a value into it, you have to choose one of the options given.

#### A.2 Control Window





## The Control Window

The Control Window is the main window from which all of TOPCAT's activities are controlled. It lists the known tables, summarises their characteristics, and allows you to open other windows for more specialised tasks. When TOPCAT starts up you will see this window - it may or may not have some tables loaded into it according to how you invoked the program.

The window consists of two main parts: the **Table List** panel on the left, and the **Current Table Properties** panel on the right. Tables loaded into TOPCAT are shown in the Table List, each identified by an index number which never changes for a given table, and a label which is initially set from its location, but can be changed for convenience.

One of the tables in the list is highlighted, which means it is the currently selected table; you can change the selection by clicking on an item in the list. Information about the selected table is shown in the properties panel on the right. This shows such things as the number of rows and columns, current sort order, current row subset selection and so on. It contains some controls which allow you to change these properties. Additionally, many of the buttons in the toolbar relate to the currently selected table.

Additionally there is a toolbar and some menus at the top which display windows and perform other actions, there is a memory monitor at the bottom left, and there may, depending on how TOPCAT was invoked, be a panel labelled **SAMP** at the bottom of the right hand panel.

The Table List, Current Table Properties panel, memory monitor, SAMP panel, and actions available from the Control Window's toolbar and menus are described in the following subsections.

### A.2.1 Table List

The Table List panel on the left of the Control Window is pretty straightforward - it lists all the tables currently known to TOPCAT. If a new table is introduced by loading it from the Load Window or as a result of some action such as table joining then its name and number will appear in this list. The currently selected table is highlighted - clicking on a different table name (or using the arrow keys if the list has keyboard focus) will change the selection. The properties of the selected table are displayed in the Current Table Properties panel to its right, and a number of the toolbar buttons and menu items refer to it.

If you double-click on a table in the list, or press Return while it is selected, that table's Data Window will appear.

Certain other applications (Treeview or even another instance of TOPCAT) can interoperate with TOPCAT using drag-and-drop, and for these the table list is a good place to drag/drop tables. For instance you can drag a table node off of the main panel of Treeview and drop it onto the table list of TOPCAT, and you will be able to use that table as if it had been loaded from disk. You can also paste the filename or URL of a table onto the table list, and it will be loaded.

Sometimes while a table is being loaded a greyed-out entry will appear in this list with text like "*Loading SAMP table*" or similar. Such entries cannot be selected, but they let you know that something is happening.

### A.2.2 Current Table Properties panel

The **Current Table Properties** panel on the right hand side of the Control Window contains a number of controls which relate to the currently selected table and its Apparent properties (Section 3); they will be blank if no table is selected. Here is what the individual items mean:

#### Label

The short name associated with the selected table. It is used in the Table List panel and in labelling view windows so you can see which table they refer to. It usually set initially according to where the table came from, but you can change it by typing into the text field.

#### Location

The original source of the selected table. This is typically a filename or URL (perhaps abbreviated), but may be something else depending on where the table came from.

#### Name

A name associated with the selected table. This may be derived from the table's filename if it had one or from some naming string stored within the table metadata.

#### Rows

The number of rows in the selected table. If the current Row Subset does not include all the rows, then an indication of how many are visible within that subset will be given too.

#### Columns

The number of columns in the selected table. If some are currently hidden (not included in the current Column Set), an indication of how many are visible will be given too.

#### Sort Order

The column (if any) which determines the current Row Order. A selector shows the column (if any) on which the rows of the Apparent Table are sorted and allows you to choose a different one. The little arrow beside it indicates whether the sort is ascending or descending.

#### Row Subset

The name of the current Row Subset. A selector shows the name of the subset which determines which rows are part of the Apparent Table and allows you to choose another one. "All" indicates that all rows are included. If you select a new value using this selector, then other windows which display subset-sensitive information for the current table will change their displayed subset to the newly-selected one. However the reverse does not happen - you can change the visible subset in the statistics window for instance or one of the graphics windows and it will not affect the value displayed here.

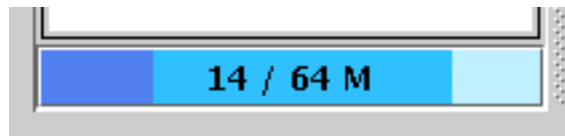
#### Activation Action

The currently selected Activation Action (Section 8). The action can be changed by clicking on this button to display the Activation Window.

The **Broadcast Row** checkbox to the right is a shortcut for one useful activation action; it corresponds to the Activation Window's option **Transmit Row** to All Listeners - that is it means that any other tools talking to TOPCAT using SAMP or PLASTIC will be informed about a row activation any time it happens. This checkbox will be greyed out if no suitable

messaging protocol is in use (see Section 9).

### A.2.3 Memory Monitor

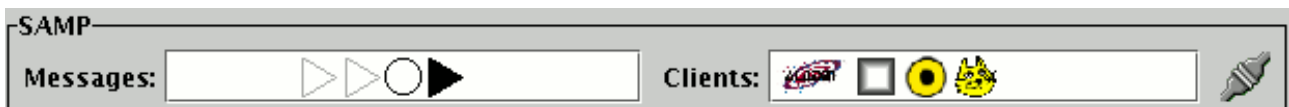


#### Control Window Memory Monitor

The memory monitor is a small widget at the bottom left of the Control Window which gives an indication of TOPCAT's memory usage. The numbers indicate the currently used and maximum heap size that Java will use (e.g. "64 M" for 64 megabytes), and the two darker colours filled in from the left indicate the current total and used proportions of this. It's not necessary to understand in detail what these mean, but if the darkest (used) colour comes to fill up the whole area, the application will slow down and then signal an Out Of Memory Error. See Section 10.4 for some tips on what to do if this happens.

If you click on the memory monitor, you will request that the Java Virtual Machine performs a garbage collection round, which may have the effect of reclaiming some memory and modifying the current usage. This is not really useful (Java will garbage collect at sensible times without prompting), but you can do it if you're bored.

### A.2.4 SAMP Panel



#### Control Window SAMP Panel

If TOPCAT is running in SAMP mode, the SAMP panel at the bottom of the Control Window gives a quick view of the current status of SAMP communications. For a discussion of the whats and whys of SAMP, see Section 9. **Note** that if not running in SAMP mode (e.g. if in PLASTIC or no-server mode) this panel will not appear. SAMP mode is the default under normal circumstances.

The panel is made up of the following main parts:

#### Message View

This shows a graphical representation of any messages which have been recently sent to or received from other applications by TOPCAT. Triangles to the left of the central circle represent incoming messages and triangles to the right represent outgoing ones. A filled triangle represents a message which is still waiting for an answer, and an open one represents a message which either is not expecting an answer or has already received one. Colour coding is used to indicate success or failure. The triangles disappear from the display shortly after they are no longer waiting for a reply.

#### Client View

An icon is shown in this panel for each application currently registered with the SAMP hub, including TOPCAT itself. If no icon is available for a registered application, a generic grey square is used.

#### Connection Indicator

An icon at the right may be visible to indicate whether or not TOPCAT is currently registered with the hub.

When TOPCAT is not connected to a SAMP hub (most likely because none is running) these panels

will be greyed out.

More detail and control for the information presented in this panel is available in the SAMP Window.

### A.2.5 Toolbar Buttons

There are a number of buttons on the Control Window's toolbar; these are the most convenient way of invoking most of TOPCAT's functions. They are grouped as follows:

#### Import and export



##### Load Table

Pops up the Load Table dialogue which allows you to load a table into TOPCAT. If a table is loaded it becomes the new current table.



##### Save Table(s)/Session

Pops up the Save Table dialogue which allows you to write out tables in various ways. You can either write out the current Apparent Table, or multiple tables, or save (much of) the current session state.



##### Broadcast Table

Broadcasts the current **Apparent Table** to any applications registered using the SAMP or PLASTIC protocol. See Section 9.

#### Table views (see Appendix A.3)



##### Data Window

Displays the table rows and columns in a scrollable viewer so you can see the cell contents themselves.



##### Parameters Window

Displays table "parameters", that is metadata which applies to the whole table.



##### Columns Window

Displays metadata about each column such as data type, units, description, UCDs etc.



##### Subsets Window

Displays the currently defined row subsets (Section 3.1) and enables new ones to be defined.



##### Statistics Window

Displays a window for calculating statistical quantities for the values in each column of the table.

#### Graphics windows



##### Histogram Plot

Displays a window for plotting 1D histograms and kernel density estimates.

**Plane Plot**

Displays a window for making various types of plot on 2D Cartesian axes.

**Sky Plot**

Displays a window for making various types of plot on the celestial sphere.

**Cube Plot**

Displays a window for making various types of plot on 3D Cartesian axes.

**Sphere Plot**

Displays a window for making various types of plot on 3D axes using spherical polar coordinates.

**Matching etc (see Section 5)****Pair Match Window**

Displays a dialog for joining tables side-by-side by locating rows which match between them.

**TAP Query**

Displays a dialogue window for querying remote databases using the Table Access Protocol (TAP). This is a very powerful way to access remote data by writing SQL-like queries, and can be used to do joins with remote tables as well as simply downloading data.

**CDS Upload X-Match**

Displays a dialog window which uses the X-Match provided by the Centre de Données astronomiques de Strasbourg to allow fast matching local tables with any tables in the remote VizieR or SIMBAD databases.

Note that other options for joining tables, including matches involving a single table or more than two tables, and joining tables top-to-bottom, are available from the **Joins** menu.

**Miscellaneous****SAMP Window**

Displays a window which displays detail about the current status, and allows configuration, of SAMP messaging (see Section 9). Note this button will not be visible if TOPCAT is running in PLASTIC mode.

**Available Functions**

Displays a window containing all the functions which can be used for writing algebraic expressions (see Section 7).

**Help**

Displays the help browser, open at the entry on the Control Window.

**Exit**

Queries for confirmation, then exits the application.

## A.2.6 Menu Items

This section describes actions available from the Control Window menus additional to those also available from the toolbar (described in the previous section) and those common to other windows (described in Appendix A.1.2).

The **File** menu contains the following additional actions:



### Duplicate Table

Adds a new copy of the current Apparent Table to the list of known tables. This is like loading in the current table again, except that its apparent characteristics become the basic characteristics of the copied one, so for instance whatever is the current row order becomes the natural order of the new one.



### Discard Table(s)

Removes the current table from the list and closes and forgets any view windows associated with it. A discarded table cannot be reinstated. You will be prompted to confirm this action. Discarding a table in this way *may* free up memory, for other operations, but often will not; whether it does or not depends on the details of where the table comes from. This action can also be triggered by hitting the **Delete** key on the keyboard when the table list has keyboard focus. If multiple tables are selected at the same time, you will be prompted to remove them all.



### Move Table Up

Moves the currently selected table up one slot in the tables list. This can be convenient for viewing, and it also influences the order in which tables are saved when doing a Multiple Table or Session save. This action can also be triggered by hitting **ALT-up-arrow** key on the keyboard when the table list has keyboard focus.



### Move Table Down

Moves the currently selected table down one slot in the tables list. This can be convenient for viewing, and it also influences the order in which tables are saved when doing a Multiple Table or Session save. This action can also be triggered by hitting **ALT-down-arrow** key on the keyboard when the table list has keyboard focus.



### Send Table to ...

Sends the currently selected table to a selected listening application using SAMP or PLASTIC. Select the desired recipient application from the submenu.



### View Log

Pops up the Log Window to display logging messages generated by the application. Intended mainly for debugging.

The **Views** menu contains actions for launching the windows which give certain views of the table metadata. These are all provided as toolbar buttons as well.

The **Graphics** menu contains actions for launching the windows which give various plotting and visualisation options. The most important ones are provided as toolbar buttons as well, but this menu contains some actions not available elsewhere. The ones marked "(old)" are the plotting

windows that were available before TOPCAT version 4; they have a different user interface, are generally less powerful, and are somewhat deprecated, but still work. The other non-toolbar options are somewhat experimental. Items available from this menu and not the toolbar are:



#### **Time Plot**

Displays the Time Plot window for plotting time series data. This is currently experimental and lacks some useful features.



#### **Histogram Plot (old)**

Displays the old-style Histogram window for plotting 1-d histograms.



#### **2D Plot (old)**

Displays the old-style 2D plot window for plotting 2-d scatter plots.



#### **3D Plot (old)**

Displays the old-style 3D plot window for plotting 3-d scatter plots on Cartesian axes.



#### **Sphere Plot (old)**

Displays the old-style sphere window for plotting 3-d scatter plots on spherical polar axes, with or without a radial coordinate.



#### **Stacked Line Plot (old)**

Displays the old-style Stacked Line Plot window for plotting multiple vertically stacked line plots against a single X coordinate.



#### **Density Map (old)**

Displays the old-style Density Map window for plotting 2-d density maps (image-like histograms on a 2-d grid of bins).

The **Joins** menu contains actions for various types of table join. The items provided additional to those on the toolbar are:



#### **Concatenation Window**

Displays the Contatenate Tables window, which allows you to join two tables top-to-bottom.



#### **Multiple Cone Search**

Displays the Multiple Cone Search Window which allows you to crossmatch a local table with a remote catalogue exposed via the Cone Search protocol. Note this is somewhat deprecated in favour of TAP or CDS X-Match Upload.



#### **Multiple SIA**

Displays the Multiple SIA Window which allows you to crossmatch with a remote image service.



#### **Multiple SSA**

Displays the Multiple SSA Window which allows you to crossmatch with a remote spectrum service.



#### **Internal Match**

Displays the Internal Match Window for finding internal matches between rows in the same table.



### N-Table Match

Displays a dialog for matching more than two tables together.

The **Windows** menu contains actions for controlling which table view windows (Appendix A.3) are currently visible on the screen. If you have lots of tables and are using various different views of several of them, the number of windows on the screen can get out of hand and it's easy to lose track of what window is where. The actions on this menu do some combination of either hiding or revealing all the various view windows associated with either the selected table or all the other ones. Windows hidden are removed from the screen but if reactivated (e.g. by using the appropriate toolbar button) will come back in the same place and the same state. Revealing all the windows associated with a given table means showing all the view windows which have been opened before (it won't display windows which have never explicitly been opened).

#### Show Selected Views Only

Reveal all view windows associated with the currently selected table and hide all others.

#### Show Selected Views

Reveal all view windows which are associated with the currently selected table.

#### Show All Views

Reveal all view windows associated with all tables.

#### Hide Unselected Views


Hide all view windows associated with tables other than the currently selected one.

#### Hide Selected Views

Hide all view windows associated with the currently selected table.

#### Hide All Views

Hide all the view windows. If you get really confused, this is a good one to select to clear up your screen prior to reinstating the ones that you actually want to look at.

Note that the **Control Window** item () on menus on all other windows is also useful for window management - it brings back the control window if it's been hidden.

The **VO** menu groups together those actions which access remote data sources. All of these options can also be found in either the Load Window toolbar or the **Joins** menu.

The **Interop** menu contains options relevant to tool interoperability (SAMP or PLASTIC). These items are available elsewhere on the toolbar or **File** menu.

## A.3 Table View Windows

Many of the windows you will see within TOPCAT display information about a single table. There are several of these, each displaying a different aspect of the table data - cell contents, statistics, column metadata etc. There is one of each type for each of the tables currently loaded, though they won't necessarily all be displayed at once. The title bar of these windows will say something like **TOPCAT(3): Table Columns**, which indicates that it is displaying information about the column metadata for the table labelled "3:" in the Control Window.

To open any of these windows, select the table of interest in the Control Window and click the appropriate toolbar button (or the equivalent item in the **Table Views** menu). This will either open up a new window of the sort you have requested, or if you have opened it before, will make sure it's

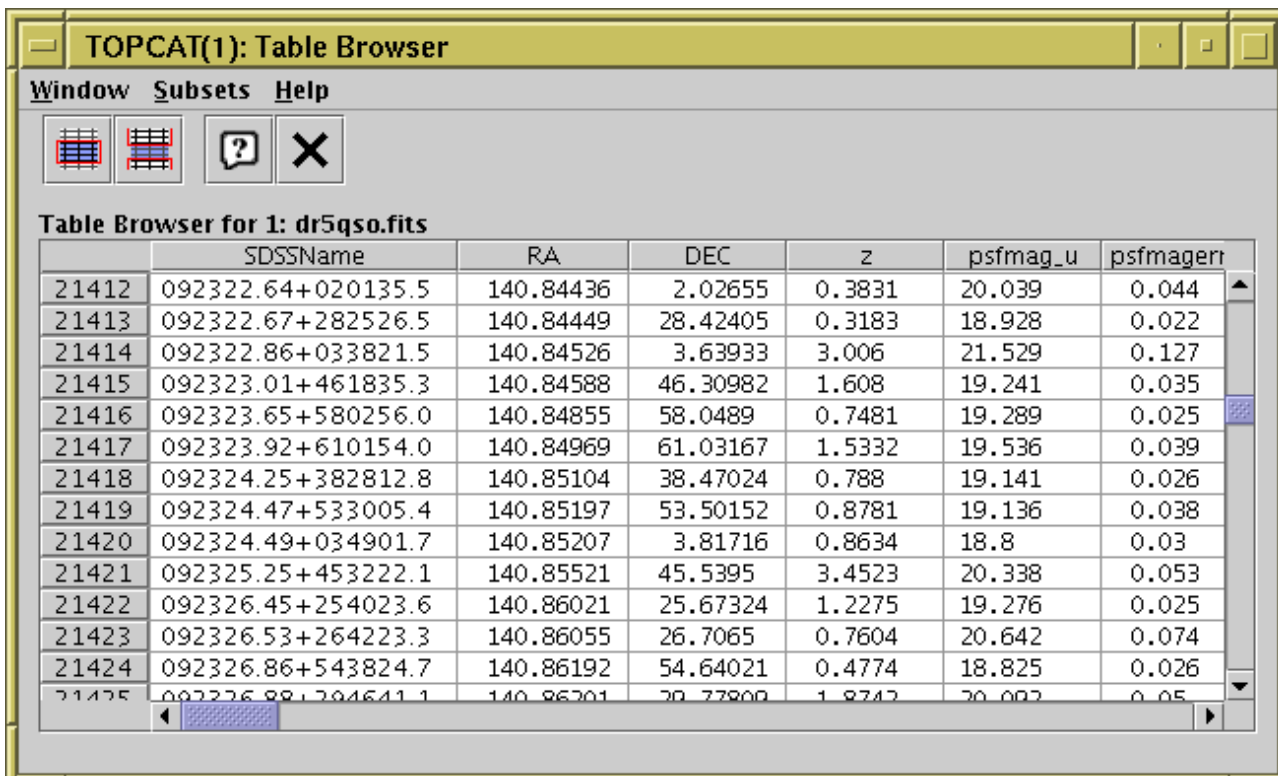


visible.

If you have lots of tables and are using various different views of several of them, the number of windows on the screen can get out of hand and it's easy to lose track of what window is where. In this case the Control Window's **Windows** menu (described in Appendix A.2.6), or the **Window|Control Window** menu item in any of the view windows can be handy to keep them under control.

The following sections describe each of these table view windows in turn.


### A.3.1 Data Window



The screenshot shows the TOPCAT(1): Table Browser window. It has a menu bar with 'Window', 'Subsets', and 'Help'. Below the menu bar are four icons: a grid, a grid with a subset, a question mark, and a close button. The main area is titled 'Table Browser for 1: dr5qso.fits'. It displays a table with 7 columns: an index column, 'SDSSName', 'RA', 'DEC', 'z', 'psfmag\_u', and 'psfmag\_r'. The table contains 17 rows of data. The index column values range from 21412 to 21425. The SDSSName column contains values like '092322.64+020135.5'. The RA column contains values like '140.84436'. The DEC column contains values like '2.02655'. The z column contains values like '0.3831'. The psfmag\_u column contains values like '20.039'. The psfmag\_r column contains values like '0.044'.

	SDSSName	RA	DEC	z	psfmag_u	psfmag_r
21412	092322.64+020135.5	140.84436	2.02655	0.3831	20.039	0.044
21413	092322.67+282526.5	140.84449	28.42405	0.3183	18.928	0.022
21414	092322.86+033821.5	140.84526	3.63933	3.006	21.529	0.127
21415	092323.01+461835.3	140.84588	46.30982	1.608	19.241	0.035
21416	092323.65+580256.0	140.84855	58.0489	0.7481	19.289	0.025
21417	092323.92+610154.0	140.84969	61.03167	1.5332	19.536	0.039
21418	092324.25+382812.8	140.85104	38.47024	0.788	19.141	0.026
21419	092324.47+533005.4	140.85197	53.50152	0.8781	19.136	0.038
21420	092324.49+034901.7	140.85207	3.81716	0.8634	18.8	0.03
21421	092325.25+453222.1	140.85521	45.5395	3.4523	20.338	0.053
21422	092326.45+254023.6	140.86021	25.67324	1.2275	19.276	0.025
21423	092326.53+264223.3	140.86055	26.7065	0.7604	20.642	0.074
21424	092326.86+543824.7	140.86192	54.64021	0.4774	18.825	0.026
21425	092326.88+294641.1	140.86201	29.77800	1.8742	20.002	0.05

#### Data Window

The Data Window presents a JTable containing the actual cells of the Apparent Table (Section 3). You can display it using the **Table Data** () button when the chosen table is selected in the

Control Window's Table List.

You can scroll around the table in the usual way. In most cases you can edit cells by double-clicking in them, though some cells (e.g. ones containing arrays rather than scalars) cannot currently be edited. If it looks like an edit has taken place, it has.

There is a grey column of numbers on the left of the JTable which gives the row index of each row. This is the value of the special Index column, which numbers each row of the original (not apparent) table starting at 1. If the table has been sorted these numbers may not be in order.

Note that reordering the columns by dragging their headings around will change the order of columns in the table's Column Set and hence the Apparent Table.

If you have table with very many columns it can be difficult to scroll the display sideways so that a column you are interested in is in view. In this case, you can go to the Columns Window and click on the description of the column you are after in the display there. This will have the effect of

scrolling the Data Window sideways so that your selected column is visible in the centre of the display here.

The following buttons are available in the toolbar:



#### **Subset From Selected Rows**

Defines a new Row Subset consisting of those rows which are currently highlighted. You can highlight a contiguous group of rows by dragging the mouse over them; further contiguous groups can be added by holding the Control key down while dragging. This action is only available when some rows have been selected.



#### **Subset From Unselected Rows**

Defines a new Row Subset consisting of those rows which are visible but currently not highlighted. You can highlight a contiguous group of rows by dragging the mouse over them; further contiguous groups can be added by holding the Control key down while dragging. This action is only available when some rows have been selected.

The **Subsets** menu additionally contains the following item:



#### **Highlight Subset**

This is a submenu in which all the currently defined row subsets are listed. Selecting one of them marks the rows corresponding to that subset as if they have been selected.

As well as the normal menu, right-clicking over one of the columns in the displayed table will present a **Column Popup Menu**, which provides a convenient way to do some things with the column in question:



#### **Replace Column**

Pops up a Synthetic Column dialogue to replace this column with a new synthetic one. The dialogue is initialised with the same name, units etc as the selected column, and with an expression that evaluates to its value. You can alter any of these, and the new column will replace the old one, which will be hidden and renamed by appending a suffix like "\_old" to its name.



#### **New Synthetic Column**

Pops up a Synthetic Column dialogue to insert a new synthetic column just after this one.



#### **Sort up**

Sorts the table rows according to ascending value of the contents of the column. Only available if some kind of order (e.g. numeric or alphabetic) can sensibly be applied to the column.



#### **Sort down**

Sorts the table rows according to descending value of the contents of the column. Only available if some kind of order (e.g. numeric or alphabetic) can sensibly be applied to the column.



#### **Hide**

Hides the column. It can be reinstated from the Columns window.



### Search Column

For string-valued columns, this option allows you to search for values in a column. If you select it you will be asked to enter a regular expression, and then any row which matches that expression in this column will be selected (highlighted). If there's just one matching column it will be activated as well. The expression obeys normal regular expression syntax, so for instance you'd enter `".*XYZ.*"` to find all rows which contain the string "XYZ".



### Explode Array Column

For columns which have an array value with a fixed number of elements, selecting this option will hide the original column and replace it by a set of scalar columns, one for each of its elements. For instance if a column PMAG contains a 5-element vector of type `float[]` representing magnitudes in 5 different bands, selecting this option will hide PMAG and insert 5 new `Float`-type columns PMAG\_1...PMAG\_5 in its place, each containing one of the magnitudes.

## A.3.2 Parameters Window

**TOPCAT(1): Table Parameters**

Window Parameters Display Help

+ - ? X

Table Parameters for 1: 6dfgs\_mini.xml.bz2

Name	Value	Units	UCD	Description
Name	6dfgs_E7_subset			Table name
URL	jar:file:/data/andromeda1/starjava/java/li...			URL of original table
Column Count	17			Number of columns
Row Count	875			Number of rows
Description	6dFGS master config file (version E7 March...			
Original Source	http://www-wfau.roe.ac.uk/6dFGS/6dfgs_...			URL of data file used to
Credits	Column explanations provided by Mike Re...			
Conversion	Converted from 6dfgs_E7.fld.gz by Mark T...			
RESOLUTION	15	arcsec	stat.error;pos.eq.ra	Nominal positional error

Name: Description

Class: String

Shape:


Units:

Description:

UCD:

Value: 6dFGS master config file (version E7 March 2004) - DEMO SUBSET. These data are taken from the 6dF Galaxy Redshift Survey Database, see astro-ph/0505068. Kindly provided by Mike Read, ROE. These data are for EXAMPLE PURPOSES ONLY, intended for demonstrations of some of TOPCAT's properties. For science use, please consult the

### Parameters Window

The Parameters Window displays metadata which applies to the whole table (rather than that for each column). You can display it using the **Table Parameters** (  ) button when the chosen table is selected in the Control Window's Table List.

In table/database parlance, an item of per-table metadata is often known as a "parameter" of the table. At least the number of rows and columns will be listed. Some table file formats (for instance VOTable and FITS) have provision for storing other table parameters, while others (for instance CSV) do not. In the latter case there may not much of interest displayed in this window.

The top part of the display is a JTable with one row for each parameter. It indicates the parameter's name, its value, the type of item it is (integer, string etc) and other items of interest such as units, dimensionality or UCD if they are defined. If a column of the table has no entries (for instance, the Units column might be empty because none of the parameters has had units defined for it) then that column may be absent from the display - in this case the **Display** menu can be used to reveal it.

You can edit some parameter values and descriptions by double-clicking on them as usual.

The bottom part of the display gives an expanded view of a selected parameter (click on a row in the top part to select one). This is especially useful if the parameter value is too long to show fully in the table display. In most cases you can edit the fields here to change the value and other characteristics of a parameter.

The following items are available in the toolbar:



#### Add Parameter

Pops up a New Parameter Window to allow you to add a new parameter to the table.




#### Remove Parameter

If one or more of the parameters displayed in the JTable in this window have been selected by clicking on the relevant rows, then clicking this button will remove them. Some parameters such as Row Count cannot be removed.





### A.3.3 Columns Window

TOPCAT(3): Table columns						
Window Columns Display Help						
Table columns for 3: 2mass_xsc.fits						
	Visible	Name	\$ID	Class	Units	
3	<input checked="" type="checkbox"/>	designation	\$3	String		
4	<input checked="" type="checkbox"/>	ra	\$4	Double	degrees	J2000.0 Right Ascension bas
5	<input checked="" type="checkbox"/>	dec	\$5	Double	degrees	J2000.0 Declination based o
6	<input checked="" type="checkbox"/>	r_k20fe	\$6	Double	arcsec	20mag/sq." isophotal K fidu
7	<input checked="" type="checkbox"/>	j_m_k20fe	\$7	Double	mag	J magnitude
8	<input type="checkbox"/>	j_msig_k20fe	\$8	Double	mag	J 1-sigma uncertainty
9	<input type="checkbox"/>	j_flg_k20fe	\$9	Integer		J confusion flag (0=no other
10	<input checked="" type="checkbox"/>	h_m_k20fe	\$10	Double	mag	H magnitude
11	<input type="checkbox"/>	h_msig_k20fe	\$11	Double	mag	H 1-sigma uncertainty
12	<input type="checkbox"/>	h_flg_k20fe	\$12	Integer		H confusion flag (0=no other
13	<input checked="" type="checkbox"/>	k_m_k20fe	\$13	Double	mag	K magnitude

## Columns Window

The Columns Window displays a JTable giving all the information (metadata) known about each column in the table. You can display it using the **Column Info** (  ) button when the chosen table is selected in the Control Window's Table List.

The display may take a little bit of getting used to, since each *column* in the main data table is represented by a *row* in the JTable displayed here. The order and widths of the columns of JTable widget can be changed in the same way as those for the Data Window JTable, but this has no effect on the data.

The leftmost column, labelled "Visible", contains a checkbox in each row (one for each column of the data table). Initially, these are all ticked. By clicking on those boxes, you can toggle them between ticked and unticked. When unticked, the column in question will become hidden. The row can still be seen in this window, but the corresponding data column is no longer a part of the Apparent Table, so will not be seen in the Data Window or appear in exported versions of the table. You can tick/untick multiple columns at once by highlighting a set of rows by dragging the mouse over them and then using the **Hide Selected** (  ) or **Reveal Selected** (  ) toolbar buttons or menu items. If you want to hide or reveal all the columns in the table, use the **Hide All** (  ) or **Reveal All** (  ) buttons.

Each column in the displayed JTable corresponds to one piece of information for each of the columns in the data table - column name, description, UCD etc. Tables of different types (e.g. ones read from different input formats) can have different categories of metadata. By default a metadata category is displayed in this JTable if at least one table column has a non-blank value for that metadata category, so for instance if no table columns have a defined UCD then the UCD column will not appear. Categories can be made to appear and disappear however by using the **Display** menu. The metadata items are as follows:

### Visible

Indicates whether the column is part of the Apparent Table. If this box is not filled in, then for most purposes the column will be hidden from display. You can toggle visibility by clicking on this column.

### Name

The name of the column.

### \$ID

A unique and unchanging ID value for each column. These are useful in defining algebraic expressions (see Section 7) since they are guaranteed unique for each column. Although the column Name can be used as well, the Name may not be unique and may not have the correct form for use in an algebraic expression.

### Class

The Java class of the items in that column. You don't have to know very much Java to understand these; they are Float or Double for floating point numbers; Byte, Short, Integer or Long for integer numbers, Boolean for a logical (true/false) flag, or String for a string of ASCII or Unicode characters. There are other possibilities, but these will cover most. The characters '[' after the name of the class indicates that each cell in the column holds an array of the indicated type.

### Shape

Cells of a table can contain arrays as well as scalars. If the column contains an array type, this indicates the shape that it should be interpreted as. It gives the dimensions in column-major

order. The last element may be a '\*' to indicate that the size of the array may be variable. For scalar columns, this item will be blank.

### Units

The units in which quantities in this column are expressed.

### Expression

The algebraic expression defining the values in this column. This will only be filled in if the column in question is a synthetic column which you have added, rather than one present in the data in their original loaded form.

### Description

A textual description of the function of this column.

### UCD

The UCD associated with this column, if one is specified. UCDs are Uniform Content Descriptors, and indicate the semantics of the values in this column.

### UCD Description

If the string in the UCD column is the identifier of a known UCD, the standard description associated with that UCD is shown here.

There may be other items in the list specific to the table in question.

You can edit column names and some other entries in this JTable by double-clicking on them as usual.

The order in which the rows are presented is determined by the table's current Column Set (Section 3.3). You can change the position of a column in the order by dragging the grey number cell at the left of the corresponding row up or down; note this is only possible for currently visible columns (ones with the Visible checkbox ticked). Another way to change this order is to drag the column headers left or right in the Data Window.

The following buttons are available in the toolbar:



### New Synthetic Column

This pops up a Synthetic Column Window which allows you to define a new column in terms of the existing ones by writing an algebraic expression. The new column will be added by default after the last selected column, or at the end if none is selected.



### Add Sky Coordinate Columns

This pops up a Sky Coordinates Window (Appendix A.10.8) which allows you to define a pair of new sky coordinate columns based on an existing pair of sky coordinate columns.



### Replace Column With Synthetic

If a single column is selected, then clicking this button will pop up a Synthetic Column dialogue (Appendix A.10.7) to replace the selected column with a new synthetic one. The dialogue is initialised with the same name, units etc as the selected column, and with an expression that evaluates to its value. You can alter any of these, and the new column will replace the old one, which will be hidden and renamed by appending a suffix like "\_old" to its name.



### Hide Selected Column(s)

If any of the columns are selected, then clicking this button will hide them, that is, remove them from the current Column Set. This has the same effect as deselecting all the checkboxes corresponding to these columns in the **Visible** column.



### Reveal Selected Column(s)

If any of the columns are selected, then clicking this button will make sure they are visible, that is, that they appear in the current Column Set. This has the same effect as selecting all the checkboxes corresponding to these columns in the **Visible** column.



### Hide All Columns

Clicking this button will hide all the columns in the table; the table will have no columns visible in it following this. If you just want to see a few columns, it may be convenient to use this button and then select a few visible ones individually to reveal.



### Reveal All Columns

Clicking this button will ensure that all the table's columns are visible (none are hidden).



### Explode Array Column

If a column is selected which has an array type, clicking this button will replace it with scalar-valued columns containing each of its elements. For instance if a column PMAG contains a 5-element vector of type `float[]` representing magnitudes in 5 different bands, then selecting it and hitting this button will hide PMAG and insert 5 new `Float`-type columns PMAG\_1...PMAG\_5 in its place each containing one of the magnitudes.



### Sort Selected Up

If a single column is selected then the table's current Sort Order will be set to sort ascending on that column. Otherwise this action is not available.



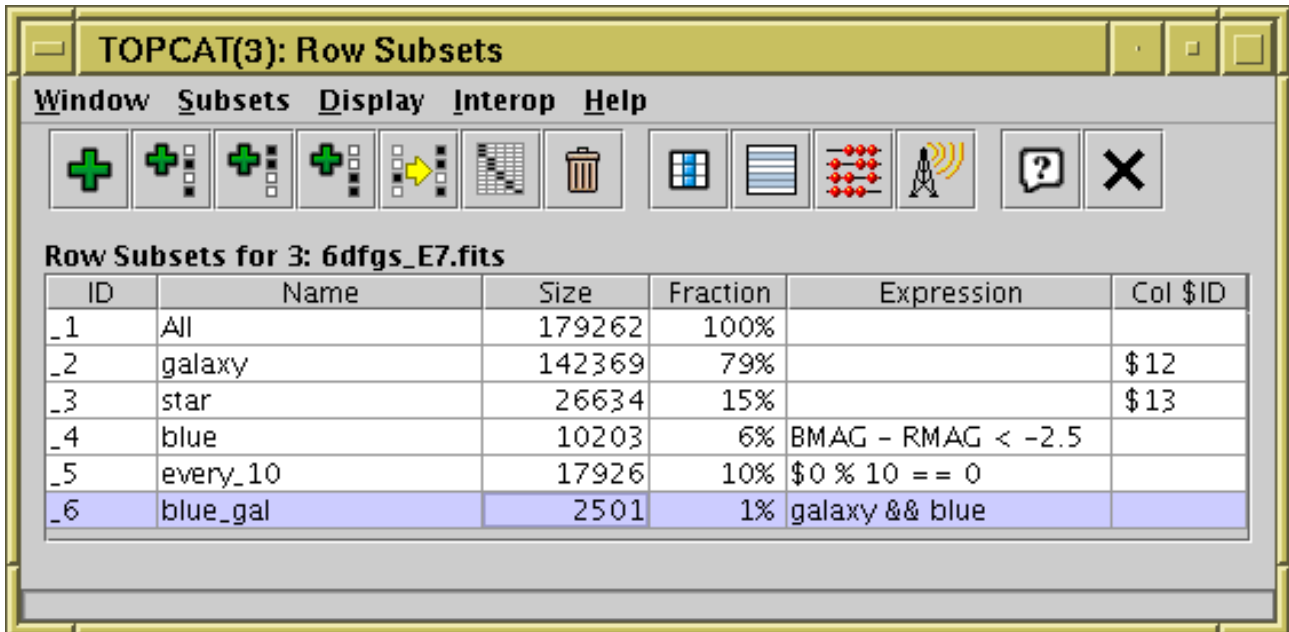
### Sort Selected Down

If a single column is selected then the table's current Sort Order will be set to sort descending on that column. Otherwise this action is not available.

Several of these actions operate on the currently selected column or columns. You can select columns by clicking on the corresponding row in the displayed `JTable` as usual. A side effect of selecting a single column is that the table view in the Data Window will be scrolled sideways so that the selected column is visible in (approximately) the middle of the screen. This can be a boon if you are dealing with a table that contains a large number of columns.


## A.3.4 Subsets Window





ID	Name	Size	Fraction	Expression	Col \$ID
_1	All	179262	100%		
_2	galaxy	142369	79%		\$12
_3	star	26634	15%		\$13
_4	blue	10203	6%	BMAG - RMAG < -2.5	
_5	every_10	17926	10%	\$0 % 10 == 0	
_6	blue_gal	2501	1%	galaxy && blue	

## Subsets Window

The Subsets Window displays the Row Subsets (Section 3.1) which have been defined. You can display it using the **Row Subsets** () button when the chosen table is selected in the Control

Window's Table List.

The subsets are displayed in a JTable widget with a row for each subset. The columns of the JTable are as follows:

### **ID**


A unique and unchanging identifier for the subset, which consists of a "\_" character (underscore) followed by an integer. This can be used to refer to it in expressions (Section 7) for synthetic columns or other subsets.

**Note:** in previous versions of TOPCAT the hash sign ("#") was used instead of the underscore for this purpose; the hash sign no longer has this meaning.

### **Name**

A name used to identify the subset. It is ideally, but not necessarily, unique. It can be edited (double-click on the cell) to change the name.

### **Size**

The number of rows in this subset. This column is usually filled in when the subset is defined, but it is not guaranteed to remain correct if the table data change, since counting may be an expensive process so it is not automatically done with every change. If values in this column are blank or suspected incorrect, a recount can be forced by using the **Count Subsets** ()

button described below.

### **Fraction**

Shows the same information as the preceding **Size** column, but as a percentage of the total number of rows in the table.

### **Expression**

If the subset has been defined by an algebraic expression, this will be here. It can be edited (double-click on the cell) to change the expression.

### **Column \$ID**

If the subset has been defined by equivalence with a boolean-valued column, this will show the \$ID of the column that it came from (see Appendix A.3.3).



Entries in the **Name** and **Expression** columns can be edited by double-clicking on them in the normal way.

The following toolbar buttons are available in this window:



#### New Subset

Pops up the Algebraic Subset Window to allow you to define a new subset algebraically.



#### Add Sample Subset

Pops up a dialogue window to allow you to define a new subset consisting of every  $N$ 'th row of the table.



#### Add Head Subset

Pops up a dialogue window to allow you to define a new subset consisting of the first  $N$  rows of the table.



#### Add Tail Subset

Pops up a dialogue window to allow you to define a new subset consisting of the last  $N$  rows of the table.



#### Invert Subset

Creates a new subset which is the complement of the selected one. The new one will include all the rows which are excluded by the selected one (and vice versa). To use this action, first select a subset by clicking on its row in the JTable.



#### Classify By Column

Pops up the Column Classification Window, which can add a set of mutually exclusive subsets based on the contents of a given column or expression.



#### Remove Subset

Deletes one of the subsets in the list. Once deleted, a subset cannot be recovered. Note that attempts to use its name or `_ID` in algebraic expressions which you add or modify in the future will fail, though its use in existing expressions will continue to work.



#### To Column

If one of the rows in the JTable is selected, this will turn that subset into a new column. It will pop up the Synthetic Column Window, filled in appropriately to add a new boolean column to the table based on the selected subset. You can either accept it as is, or modify some of the fields. To use this action, first select a subset by clicking on its row in the JTable.



#### Highlight Subset

Highlights the contents of this subset by marking the rows visibly in the data window and also ensuring that the rows are visible in any plot windows. This replicates what happens when the subset is first created.



#### Count Rows

Counts how many rows are in each subset and displays this in the **Size** column. This forces a count or recount to fill in or update these values.



### Broadcast Subset

Causes the selected subset to be broadcast using SAMP or PLASTIC to other listening applications. Any other listening application which is displaying the same table is then invited to highlight the selection of rows corresponding to the selected subset. This option and the corresponding **Send Subset to ...** option are also available from the **Interop** menu. See Section 9 for more information about tool interoperability.

The following additional menu items are available:




### Send subset to ...

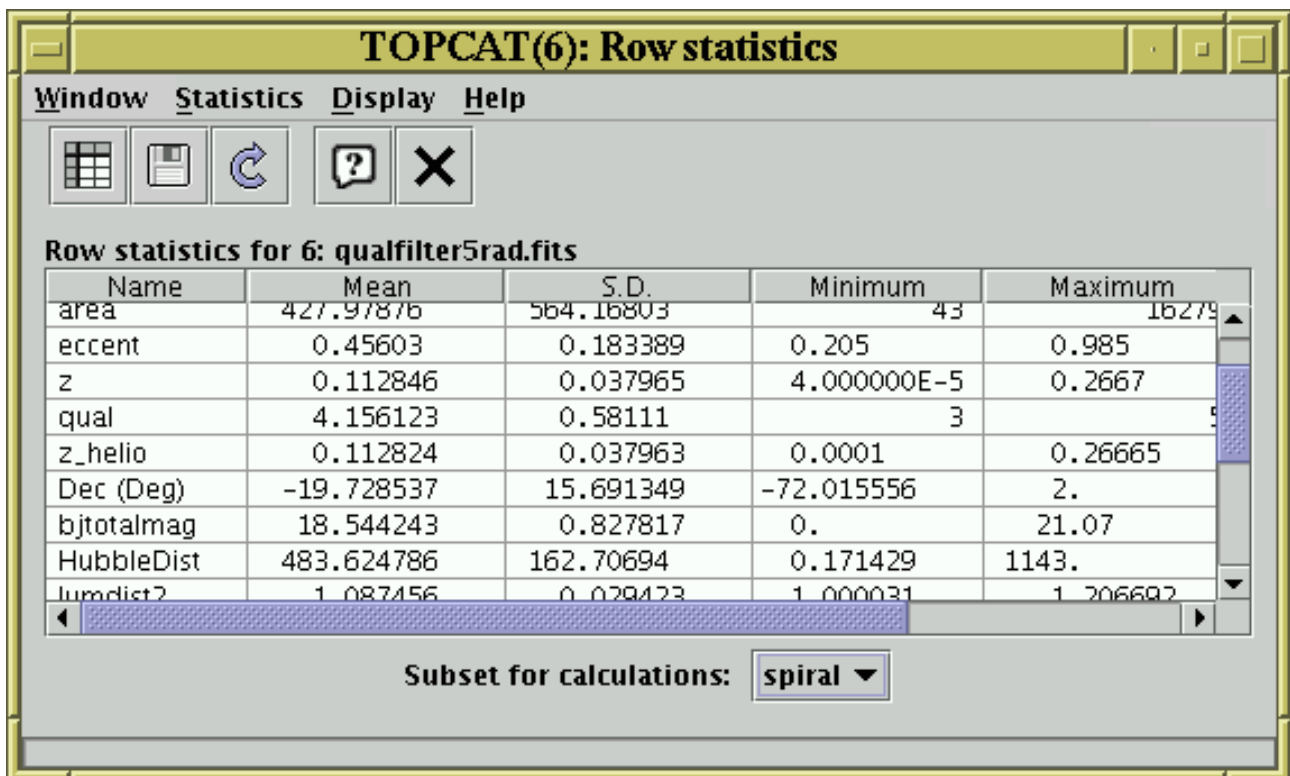
As for the **Broadcast Subset** item above, but sends the subset to only a single selected application. A submenu will give a list of all the currently registered applications which can make sense of a subset. If there are none, this item will be disabled.



### Autocount rows

Normally, the **Size** and **Fraction** columns of the displayed table (see above) are filled in as soon as a new subset is defined. However for very large tables this could be slow - if you want to prevent this autocounting you can deselect this menu item. In this case the Size and Fraction columns will only be filled in when the **Count** (  ) button is hit or if TOPCAT finds out the size as the result of some other operation (such as plotting).

## A.3.5 Statistics Window



### Statistics Window

The Statistics Window shows statistics for the values in each of the table's columns. You can display it using the **Column Statistics** (  $\Sigma$  ) button when the chosen table is selected in the

## Control Window's Table List.

The calculated values are displayed in a JTable widget with a row for each column in the main table, and a column for each of a number of statistical quantities calculated on some or all of the values in the data table column corresponding to that grid row. The following columns are shown by default:

**Name**

The name of the column in the main table represented by this grid row.

**Mean**

The mean value of the good cells. For boolean columns, this is the proportion of good cells which are True.

**SD**

The population standard deviation of the good cells.

**Minimum**

The minimum value. For numeric columns the meaning of this is quite obvious. For other columns, if an ordering can be reasonably defined on them, the 'smallest' value may be shown. For instance string values will show the entry which would be first alphabetically.

**Maximum**

As minimum, but shows the largest values.

**nGood**

The number of non-blank cells.

Several additional items of statistical information are also calculated, but the columns displaying these are hidden by default to avoid clutter. You can reveal these by using the **Display** menu:

**Index**

The index of the column in the table, i.e. the order in which it is displayed.

**\$ID**

The unique identifier label for the column in the main table.

**Sum**

The sum of all the values in the column. For boolean columns this is a count of the number of True values in the column.

**Variance**

The population variance of the good cells.

**Sample SD**

The sample standard deviation of the good cells.

**Sample Variance**

The sample variance of the good cells.

**Median Absolute Deviation**

The median of absolute deviations from the median: `median(abs(x-median(x)))`. This is a robust measure of statistical dispersion.

**Scaled Median Absolute Deviation**

The **Median Absolute Deviation** (see above) multiplied by 1.4826. This is supposed to be a consistent estimator for the standard deviation, on the assumption of a normal distribution.

**Skew**

Gamma 1 measure of skewness of the value distribution.

**Kurtosis**

Gamma 2 measure of peakedness of the value distribution.

**Row of min**

The index of the row in the main table at which the minimum value occurred.

**Row of max**

The index of the row in the main table at which the maximum value occurred.

**nBad**

The number of blank cells; the sum of this value and the Good cells value will be the same for each column.

**Cardinality**

If the column contains a small number of distinct values then that number, the column's *cardinality* will be shown here. Cardinality is the number of distinct values which appear in that column. If the number of values represented is large (currently >50) or a large proportion of the non-bad values (currently >75%) then no value is shown.

In addition, some **quantile** values can be calculated on demand (by selecting their values in the **Display** menu, as for the previous list). The available values are:

**Q001:**

value below which 0.1% of rows fall

**Q01:**

value below which 1% of rows fall (1st percentile)

**Quartile1:**

value below which 25% of rows fall (first quartile)

**Median:**

value below which 50% of rows fall (median)

**Quartile3:**

value below which 75% of rows fall (third quartile)

**Q99:**

value below which 99% of rows fall (99th percentile)

**Q999:**

value below which 99.9% of rows fall

These are considerably more expensive to calculate than the other statistical quantities, and so they are not provided by default. If you attempt to calculate them for large tables, you may get a message saying that there is insufficient memory.

The quantities displayed in this window are not necessarily those for the entire table; they are those for a particular Row Subset (Section 3.1). At the bottom of the window is the **Subset For Calculations** selector, which allows you to choose which subset you want the calculations to be done for. By clicking on this you can calculate the statistics for different subsets. When the window is first opened, or when it is invoked from a menu or the toolbar in the Control Window, the subset will correspond to the current row subset.

The toolbar contains the following extra buttons:



**Save as Table**

Clicking this button will save the quantities displayed in this window to a table on disk. It can be saved in any of the tabular formats which TOPCAT understands.



**Import as Table**

The table of statistical quantities displayed by this window (rows corresponding to input table columns and columns corresponding to statistical quantities) is itself a table. By clicking this

button it can be loaded into TOPCAT as a new table and manipulated in all the usual ways. This has the same effect as saving the statistics to file (see previous button) and then reloading that file.



### Recalculate

Once statistics have been calculated for a given subset they are cached and not normally recalculated again. Use this button if you want to force a recalculation because the data may have changed.







For a large table the calculations may take a little while. While they are being performed you can interact with the window as normal, but a progress bar is shown at the bottom of the window. If you initiate a new calculation (by pushing the Recalculate button or selecting a new subset) or close the window during a calculation, the superceded calculation will be stopped.

## A.4 Plot Windows

*This section describes the plotting windows introduced at TOPCAT version 4. These provide most of the same functionality as the old-style plot windows (Appendix A.5), but additionally much more flexibility, configurability, extensibility and better performance and capabilities for making visual sense of very large data sets.*

TOPCAT has a number of windows for performing data visualisation of various kinds. These share various characteristics which are described in the first subsection below; the specific windows themselves are described in the later subsections.

Six plot windows are currently available (one is experimental):

-  Histogram Plot
-  Plane Plot
-  Sky Plot
-  Cube Plot
-  Sphere Plot
-  Time Plot (experimental)

More may be introduced in future releases.

The rest of this section describes the plotting functions in detail.

### A.4.1 Differences From Old-Style Plot Windows

A brief summary of improvements these windows offer over the old-style plot windows is:

#### New Sky Coordinate Plot

- Choice of projection: Sin (rotatable), Aitoff, Plate Carrée
- Data and view sky coordinate systems selected separately: options are equatorial, galactic, supergalactic, ecliptic
- Sky coordinate grid labelled, configurable and visible at all zooms

#### New data plot options

- Vectors
- Ellipses (with position angle)
- Pair data point links
- Contours
- Variable size markers
- Kernel density estimates

### **Improved interactive response**

- Various mouse wheel and drag actions pan and zoom the plot instantly
- In 3d plots right mouse button recentres cube on selected point
- In 3d plots zooming zooms data in the cube rather than enlarging the cube wireframe itself
- Many controls are sliders which update the plot as you slide

### **Better support for large datasets**

Several features have been introduced to provide more meaningful visualisation of large datasets. Improved density-like plots and contours give you better ways to understand plots containing many more points than there are pixels to plot them on. There is separately some improvement in scalability: you can typically get reasonable interactive performance up to about 10 million points depending on available memory etc and what you're doing (though there are reports of it working with several  $10^8$  points). The intention is to improve this limit further in future.

### **New plot shading modes**

Density colour coding for all plot types, with colour map either absolute or modifying dataset base colour. The result is something that looks like a scatter plot at low densities and a density plot for high densities, and generally means that you can easily make quantitative sense of overcrowded regions of a scatter plot. Flat, transparent and aux colour coding still available as before.

### **Improved axis labelling**

- Choice of font size and style
- Option of LaTeX input for non-ASCII characters etc
- Log axes labelled better
- Minor tick option

### **Analytic function plotting in 2D**

Plot functions of X or Y coordinate using algebraic expression language.

### **Configurability**

Many more configuration options including legend placement, grid display and colour, antialiasing, text label crowding limits etc etc.

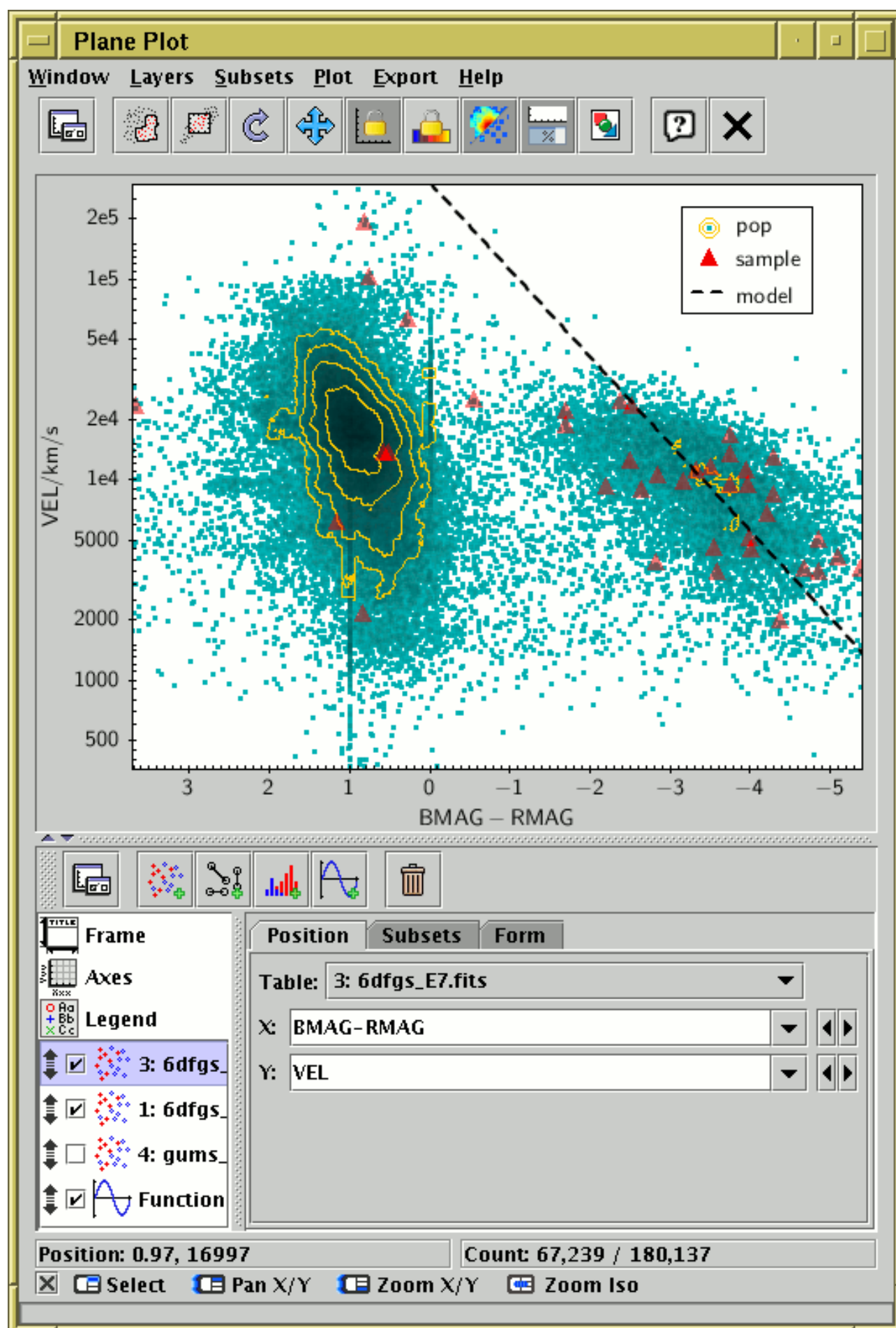
The new windows allow you to assemble a stack of layers representing different plot types of different data sets on the same axes. The user interface for controlling this is quite a bit different than for the old-style plot windows, and is described in the subsequent sections. However, making a simple plot is still simple: select a table, select the columns, and you're off.

Some features from the old-style plots that are *not* currently available in the new-style plots are:

- true RGB density maps (though density shading modes will often do as good or a better job)
- fogging in 3D plots


The intention is to introduce some of all of these in a future release, possibly informed by user demand.

## **A.4.2 Window Overview**

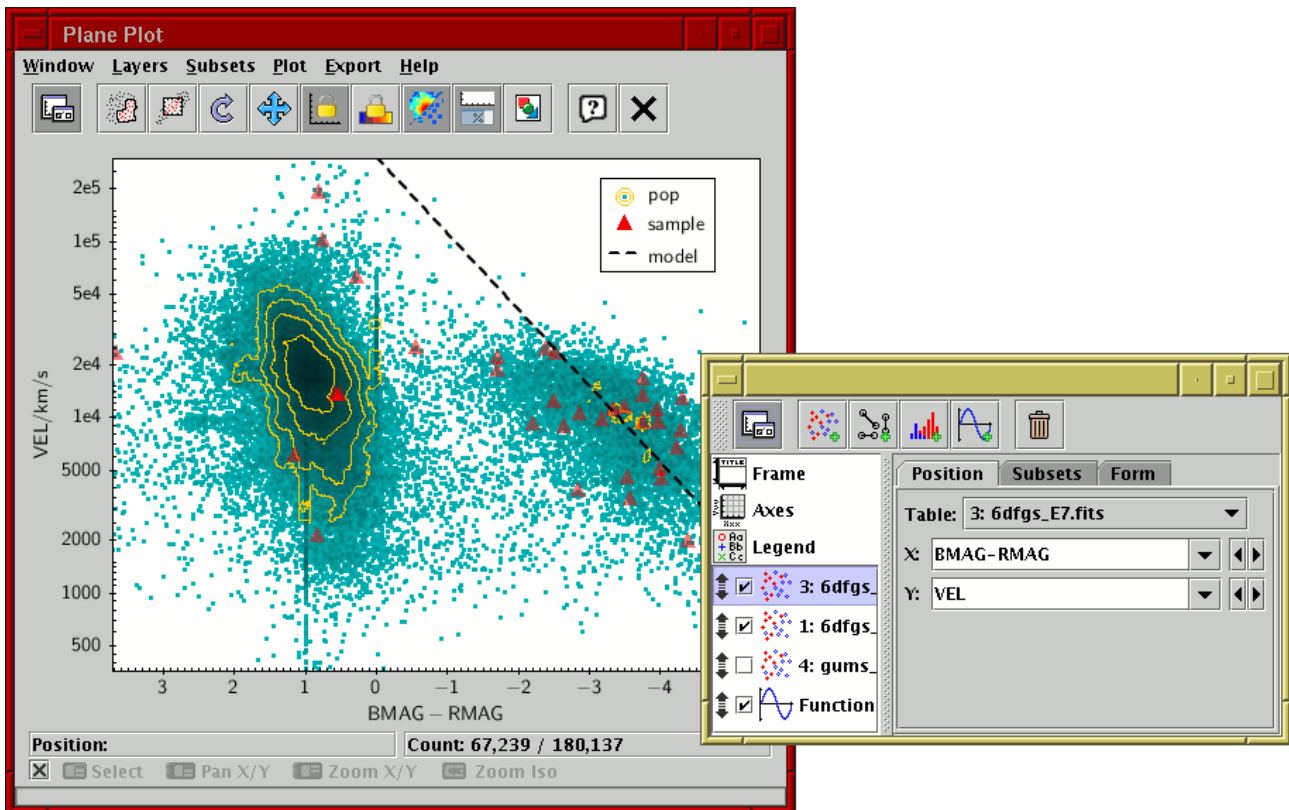




## Plot window

Plot windows consist of two main parts: the **Plot Panel** containing the actual plotted graphics (by default, at the top), and the **Control Panel** (by default, at the bottom). The Control Panel is where you configure what will be plotted. For a simple scatter plot it may just be a case of selecting what columns to plot against each other, but it can get quite detailed. If you want more screen space to play with, it can be helpful to float the control panel into a separate window using the **Float Controls** () toolbar button; you can find this button in both the main and the control panel

toolbars. To unfloat the control panel, either just close the control panel window, or click the Float Controls toolbar button again. With floating controls, the window looks like the following figure.






### Plot window with floated control panel

The **control panel** itself has two main parts: a **control stack** on the left containing currently active controls (represented by names and icons), and a **detail panel** on the right which shows information about the currently selected control. Click on one of the control entries on the left to see its details on the right. Different controls have different detail panels, but in general each one will have multiple tabs for configuring different things. You can select these by clicking on the tab names. A good way to learn about the options is to click on the different controls and their tabs to see what's available and experiment with the various options to see what happens to the plot, but all the panels and tabs are also explained in this manual. The control panel also has a toolbar at the top, used for adding and removing controls from the stack.


The **control list** has two types of entry:

#### Fixed controls:

These control the overall plot appearance. In the above figure, the fixed controls **Frame** () , **Axes** () and **Legend** () are visible.


The different fixed controls are described in Appendix A.4.3.

**Layer controls:**

These determine the actual data that will be plotted and what graphical form it takes; each control contributes a layer or layers to the plot. To add a layer control, use the toolbar just above; each button with a little green "+" adds a control of the corresponding type. To remove a layer control, select it and use the **Remove Current Layer** (  ) button in the toolbar. You

can also use the **Layers** menu at the top of the main window to add and remove layers.

The checkbox beside each control determines whether it is currently active; if unchecked, any plotting instructions it contains will be ignored. You can set them active or inactive by clicking on the checkboxes.

You can also drag each control up or down by dragging with the grab handle (  ). Layers

lower down the list are plotted later (perhaps obscuring earlier one), so you can drag them up and down until you have the layers you want on top.

The different layer controls are described in Appendix A.4.4.

In the foot of the window, there are also four other small panels:


**Position Panel**

When the cursor is positioned over the plot itself, this reports its position in data coordinates. In some cases, such as 3-d plots, this may not be possible, in which case it's blank.


**Count Panel**

Displays the number of points currently plotted and the total number of points represented by the plot. The total (the second figure) is the number of positions in all the plotted data sets, and the current number (the first figure) excludes those in subsets not currently plotted and those outside the bounds of the visible plot.

**Navigation Help**

Shows some reminders for what different mouse gestures do. Little icons are supposed to represent clicking and dragging different mouse buttons and the mouse wheel. Note the content changes according to where the mouse is on the plot, since that affects navigation behaviour. For more information see Appendix A.4.2.1. Clicking on the  button will make this panel go away.

**Progress Bar**

If something slow is happening you may see a progress bar at the bottom of the screen while you wait. Unless you are plotting several million points, you may not see this at all. For slow data loads, this will always be displayed. For actions like actually drawing the plot and turning a blob into a subset selection you can choose whether progress is shown using the **Show Plot Progress** (  ) button in the toolbar. It's nice to know that something's going on, but it can be

distracting; displaying progress also slows the plot down a bit.

The **main toolbar** at the top of the window contains the following actions (repeated in the menus):

**Float Controls**

Puts the Control Panel into a floating window rather than at the bottom of the plot window, as described above. Once floating, the control panel can be joined back to the main window by clicking this button again.

**Draw Subset Region**

Allows you to draw a region on the screen defining a new Row Subset. When you have finished drawing it, click this button again to indicate you're done. See Appendix A.4.2.3 for

more details.



### Subset From Visible

Defines a new **Row Subset** consisting of only the points which are currently visible on the plotting surface. See Appendix A.4.2.3 for more explanation.



### Replot

Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, or by redefining a subset, the plot is not automatically redrawn. Clicking this button redraws the plot taking account of any changes to the table data.



### Rescale

Rescales the axes of the current plot so that it contains all the data points in the currently selected subsets. By default the plot will be initially scaled like this, but it may have changed because of changes in the subset selection or from zooming in or out.



### Lock Axes

Usually, when the data plotted has changed significantly, the axes are automatically rescaled so that all the points are visible. The application makes a guess about when it's a good idea to do this automatic rescaling. If you don't want it to auto-rescale, set this toggle button, and it won't rescale unless it really has to. This is not available for the **Sky Plot**.



### Lock Aux Range

This controls when the Aux data range, sometimes used to colour data points (see the Aux Axis Control), is updated to match the currently visible data. By default, the range is updated dynamically as the plot changes, for instance when you pan or zoom it, so that the data range covered by the aux colour ramp matches the range of the currently visible data. But if this checkbox is checked, then the range is frozen to the current value. Data-sensitive updates to the aux range will then not be performed until it's unchecked again.

It also affects some other dynamic ranging calculations such as **Auto-Scale** sizes for plot forms Size, SizeXY, Vector, SkyVector, Ellipse and SkyEllipse.



### Sketch Frames

If selected causes intermediate "sketch" frames to be drawn when navigating around very large plots. For plots that take a long time (at least a non-negligible fraction of a second) to draw, if this option is selected then when navigating around it will paint intermediate frames based on a subsample of the data rather than painting the whole plot at every step. This can result in a somewhat flickering appearance, but it means that frame updates happen more frequently, so it's a bit more responsive.



### Show Plot Progress

If selected a progress bar at the bottom of the window is active when large (slow) plots are in progress. This can be useful if you are navigating round a very large plot so that you can see something is happening rather than the application apparently just doing nothing. On the other hand a flickering progress bar can be distracting. Updating the progress bar may also slow the plot down a little.




### Export Plot

Allows you to save the plot in a variety of graphics formats using the Plot Export window.


The window **menus** offer an alternative way to perform the actions available from the toolbar described above. They also provide some additional options:



#### Layers menu

This menu repeats the options available in the toolbar from the Control Panel at the bottom of the window; each one adds a new Layer Control of one of the available types to the stack, or

**Removes** (  ) the currently selected control.

#### Export menu

This menu provides the **Export plot** (  ) option which saves the visible plot in an image

format. It also contains two sub-menus that may or may not contain options: **Layer Data Import** (  ) and **Layer Data Save** (  ). Some plotted layer types generate table data as

part of their calculations that can be exported separately. If one of these layers is currently plotted, then options may appear in these menus. The **Import** options retrieve the table and add it to the list of tables currently loaded in TOPCAT; the **Save** options can write the data directly to disk using some table format.

The following subsections explain some other features common to all the plotting windows.

#### A.4.2.1 Navigation

All the plot windows are interactive, and you can navigate around them using the mouse buttons. For most plot types the left and right buttons and the mouse wheel do something; in 3d plots the middle button is used as well.

The details of what mouse actions do what depend on which plot window you are using, and they can also be configured to some extent using the **Navigation** tab in the **Axes** fixed control in the control stack. However, as a rule, the actions when used on the body of the plot are these:

##### **Left drag**

Drags the plot around. Where possible, the same plot position stays under the cursor as you drag. In 2d this pans the plot left/right/up/down, and in 3d it rotates it.

##### **Right drag**

Stretch zoom. Dragging up/down stretches/squashes the plot vertically, and dragging left/right stretches/squashes it horizontally. The zoom is centered on the position where you start the drag from, so that data position stays in the same place on the screen. In 3d, the zoom is along the two plot directions most closely aligned with the plane of the screen.

##### **Middle drag**

Frame zoom. Dragging right-and-down or right-and-up drags out a frame; when the button is released, the plot will be zoomed in to cover area enclosed by the frame. Dragging left (and up or down) does something like the opposite, you can zoom out using a similar (though not quite the same) mechanism.


##### **Mouse wheel**

Spinning the mouse wheel forwards/backwards will zoom in/out. In 2d the zoom is around the current position of the mouse, and in 3d it is (usually) around the center of the view cube.

##### **Left click**





Identifies a point. If there is a plotted point near the cursor, it will plot a marker on it and activate (Section 8) it. If you click on an empty bit of the plot, any existing activated point will be removed, otherwise nothing will happen.

In the 2d plot types, you can pan or zoom in just one direction by using the same actions outside of the plot itself. If you do a pan/zoom action to the left of the Y axis, it will pan/zoom only vertically, and if you do it below the X axis, it will do it only horizontally.

Most of these actions give you some visual feedback on the screen, showing a rectangle or some arrows to give you a clue what you're doing. If you find that distracting, you can turn it off using the **Plot|Show Navigation Graphics** () menu item.



#### Navigation help panel at the bottom of plot windows

In each plot window there is a row at the bottom of the window giving hints on the currently available navigation actions. The little icons are each supposed to represent either a *click* () or a *drag* () with one of the three buttons pressed, or a wheel spin () and are followed by a short description of what it will achieve. Note these hints change according to where the mouse is currently positioned on the screen. Moving the button over this panel will give you some help on the help. Clicking the **X** button will make the line disappear, and you can bring it back with the **Window|Show Navigation Help** () menu item. The **?** button will bring up a help window specific to the navigation in this window.

If you are using a mouse with fewer than three buttons or no wheel, the following **substitute gestures** can usually be used:

##### **Left button**

If you only have one button, this is it.

##### **Right button**

If you only have one button, you can use it with the CTRL key.

##### **Middle button**

If you only have one button, you can use it with the SHIFT key. If you have two buttons, sometimes pressing them both at once will work.

##### **Wheel**

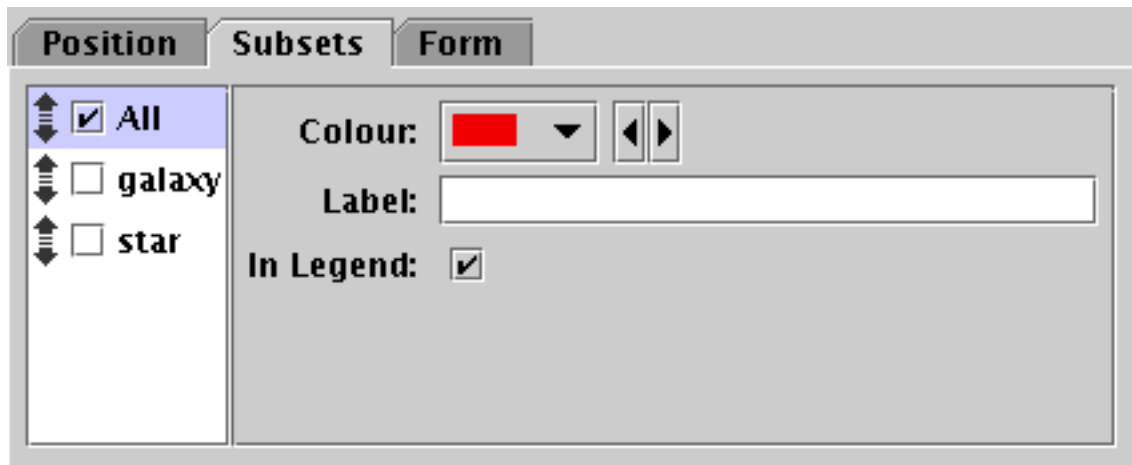
Often a wheel action can be simulated on a trackpad by moving two fingers together up or down.

### A.4.2.2 Table Data Layer Controls

Most, but not all, of the available layer controls (Appendix A.4.4) are concerned with plotting table data in some way or other. In general, these allow you to set up a way to generate a plot layer (some generated graphics) from some selected columns of a given table with a chosen style.

However, if you have more than one Row Subset for a given table, a single control can be used to generate several layers, one for each subset. In most cases it makes sense to have the graphics generated by the different subsets configured similarly, but still visually distinguishable.

These table layer controls generally have three tabs: **Position**, **Subsets** and **Form**. The Position tab is for specifying coordinates, and depends on the specific type of the plot and the layer control. The **Subsets** and **Form** control which row subsets are plotted and how they are displayed, and a description of how they are arranged follows below.



### Subsets tab of a table data layer control

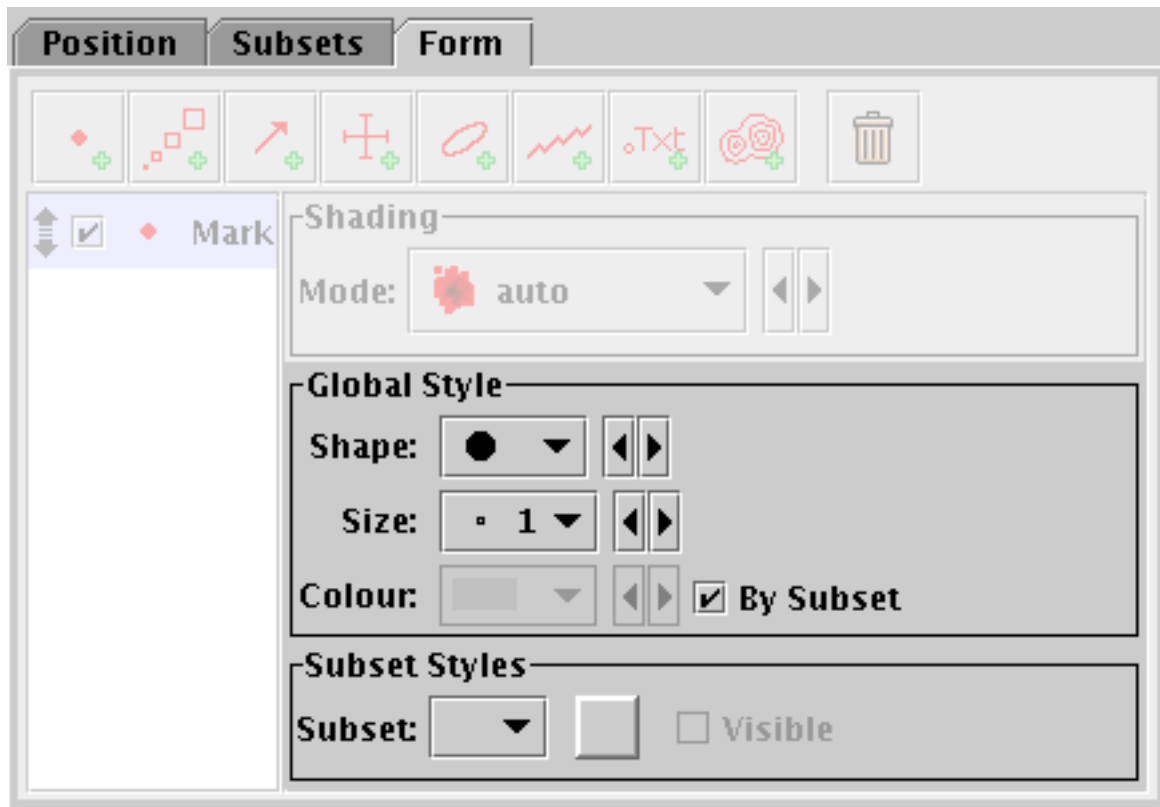
The **Subsets** tab lists the defined Row Subsets for the table you have selected in the **Position** tab. The subset **All** is always present; others may or may not be depending on whether any have been defined. Note that actions you take in the plot (for instance selecting a new subset by region) can result in new entries being added to this list.

The list on the left names the subsets with an activation checkbox and a grab handle; the panel on the right gives the detail for the currently selected subset. Select a subset to see/change its detail by clicking on it.

For each subset you can select:

- Whether it is plotted (using the activation checkbox)
- The plotting sequence (by dragging it up and down using its grab handle)
- The default plotting colour
- The label with which it is annotated in the legend (if visible)
- Whether it appears in the legend at all

Although you can select plotting colours in the **Form** tab as well, it's generally better to do it here since this changes the colour of all the forms plotted for a given subset, rather than one form at a time.



### Style configuration in Form tab of a table data layer control

The **Form** tab allows you to select style information to determine the details of how the plot will look for each layer. The details differ greatly between layer control types, but they generally contain two sub-panels for defining the style details, **Global Style** and **Subset Styles**:

#### Global Style


Controls the style details for the chosen form, for instance marker shape and size. Options here affect all subsets, though by default the colour is taken from the **Subsets** tab. If you want to set the colour the same for all subsets, uncheck the **By Subset** checkbox which will activate the **Colour** selector.

#### Subset Styles

If you want to have different subsets represented with different styles, for instance different shapes for different subsets, you can select a subset here and alter style details for that subset only, overriding the Global settings above. The **Visible** checkbox indicates and controls whether the subset selected for specific configuration is currently visible in the plot.


### A.4.2.3 Defining Subsets by Region

The plot windows can be used to make new Row Subset selections.

If you have panned, zoomed, set the axes manually or otherwise navigated the plot so that only the points currently visible in the plot are the ones you are interested in, you can click the **New Subset From Visible** () toolbar button. This defines a subset consisting of all the points that are

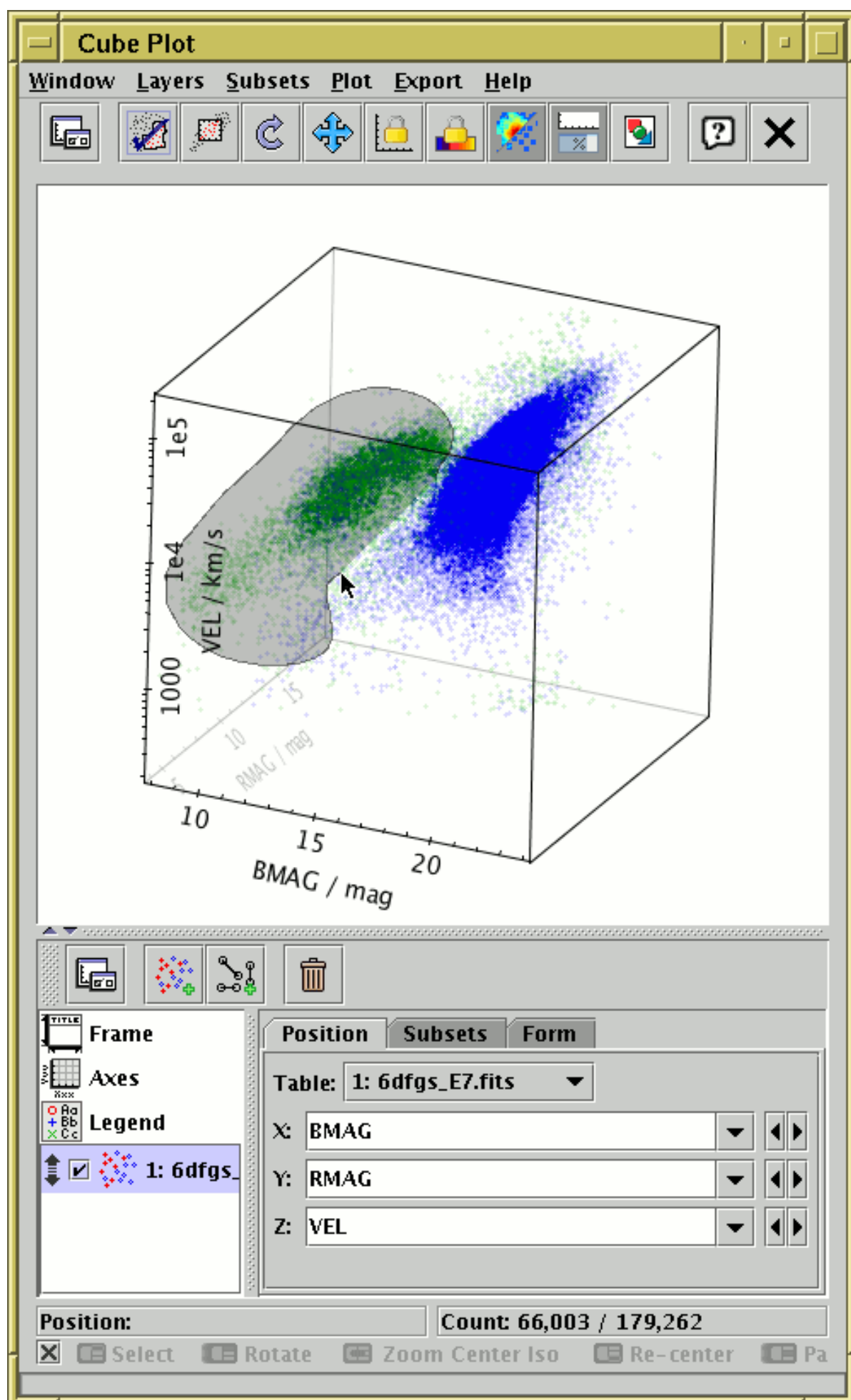
visible on the current plot. This is only useful if the group you are interested in corresponds to a rectangular region in the plotting space.

A more flexible way is to draw a region or regions on the plot which identify the points you are interested in. To do this, hit the **Draw Subset Region** () toolbar button. Having done this, you


can drag the mouse around on the plot (keep the left mouse button down while you move) to encircle the points that you're interested in. As you do so, a translucent grey blob will be left behind - anything inside the blob will end up in the subset. You can draw one or many blobs, which may be overlapping or not. If you make a mistake while drawing a sequence of blobs, you can click the right mouse button, and the most recently added blob will disappear. When you're in this region-drawing mode, you can't zoom or resize the window or change the characteristics of the plot, and the **Draw Subset Region** button appears with a tick over it () to remind you you're in it.

Here's what the plot looks like while you're drawing:



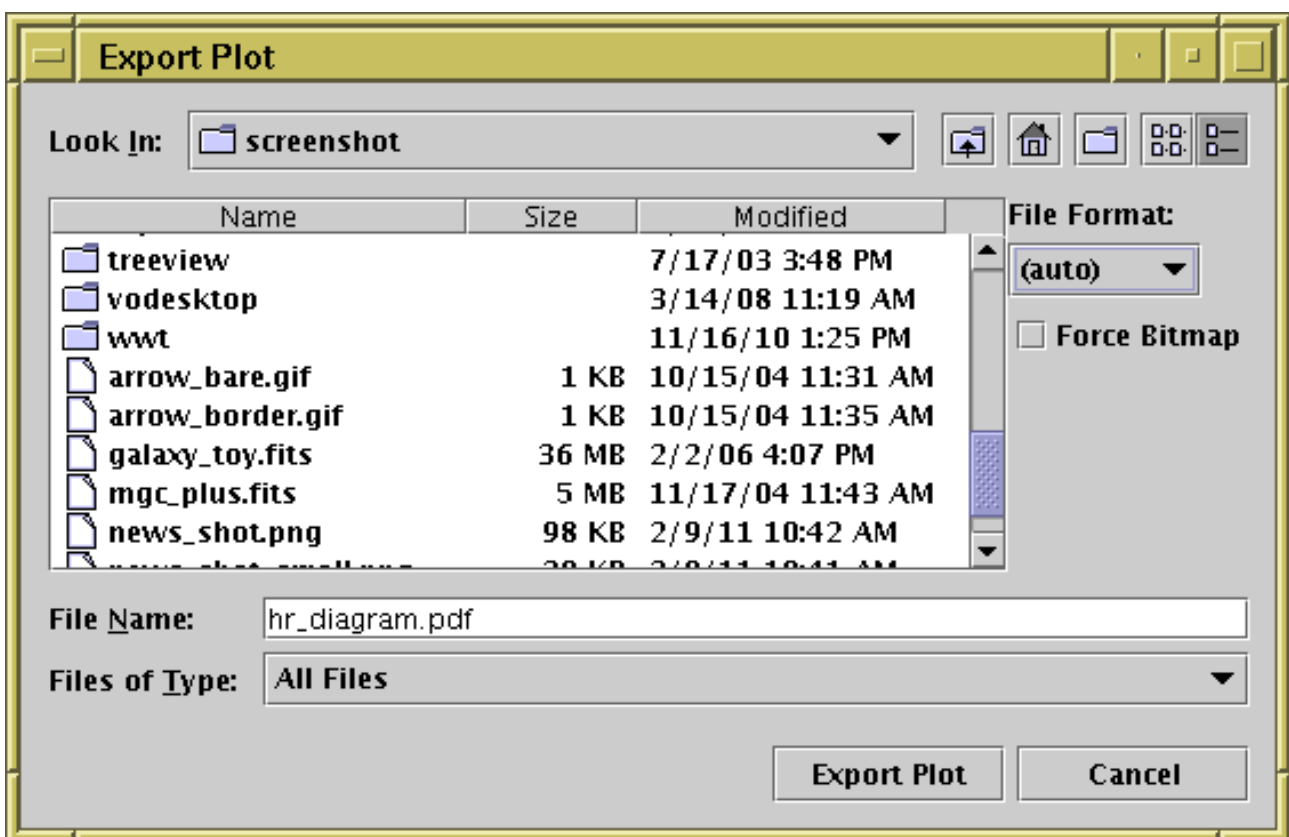


## Region-Drawing Mode


When you're happy with the region you've defined, click the  toolbar button again.

In either case, when you have indicated that you want to define a new row subset, a dialogue box will pop up to ask you its name. As described in Section 3.1.1, it's a good idea to use a name which is just composed of letters, numbers and underscores. You can optionally select a subset name which has been used before from the list, which will overwrite the former contents of that subset. When you enter a name and hit the **OK** button, the new subset will be created and the points in it will be shown straight away on the plot using a new symbol. As usual, you can toggle whether the points in this subset are displayed using the **Subsets** tab in the table layer control.

### A.4.2.4 Plot Export Window



#### Plot export window

The **Plot Export Window** can be reached with the **Export plot to file** () toolbar button in any of the plot windows.

You can select a file in the usual way, and save the plot in one of the following graphics formats:

#### png

PNG bitmap. The background is opaque.

#### png-transp

PNG bitmap with a transparent background. Background pixels that fall outside the plot surface itself (for instance outside the axes for a Plane plot or outside the celestial sphere for a Sky plot) are transparent.

#### gif

GIF bitmap; note the number of colours is limited to 256.

**jpeg**

JPEG bitmap; note that this is a lossy format, better suited to photographs than plots, and colours will be blurred.

**pdf**

Portable Document Format; in most cases this vector format gives pretty good output, in particular text will be rendered properly.

**eps**

Encapsulated PostScript; PostScript cannot handle transparency, which means that in some cases the output will come out wrong. PostScript files can also be very large if there are many data points.

**eps-gzip**

Just like **eps**, but the output is gzipped before output.

There are two additional controls on the right hand side of this window:

**File Format**

Selects the output file format as above. The default setting is **(auto)**, which guesses what format you want to use from the filename, and which usually does the right thing.

**Force Bitmap**

This option only has an effect for vector graphics formats (PDF and PostScript). If selected, it draws the data contents of the plot as a pixel map and embeds that into the output file rather than plotting each point in the output. This may make the output less beautiful (round markers will no longer be perfectly round), but it may result in a much smaller file if there are very many data points. Plot annotations such as axis labels will not be affected - they are still drawn as vector text. Note that in some cases (e.g. use of the auto, density or weighted shading modes) this kind of pixellisation will happen in any case.

Exporting to the pixel-based formats (GIF, JPEG, PNG) is fairly straightforward: each pixel on the screen appears as one pixel in the output file. PNG is generally recommended. GIF works well in most cases, but if there are more than 255 colours some of the colour resolution will be lost. JPEG can preserve a wide range of colours, but does not support transparency and is lossy, so close inspection of image features will reveal blurring.

When exporting to Portable Document Format (PDF) or Encapsulated PostScript (EPS), which are vector graphics formats, there are a few things to consider:

**Positional Quantisation**

Some of the shading modes (Density, Weighted, Auto) are inherently pixellated, and others (Flat, Aux) are not. In the former case you will see pixel boundaries for the plotted points rather than nice rounded edges at high magnifications (though text and axes will always be plotted nicely). In both cases, at present the *positional* resolution is the same as it would be on the screen, so if you have a 400-pixel high plot for instance, there are only 400 possible Y coordinates at which a marker can be plotted, which in general is not obvious by looking at the output plot. In future versions the positional resolution of non-pixellated modes may be improved. In either case, increasing the size of the plot on the screen by resizing the window before performing an export to PDF or EPS will reduce the effect of the positional quantisation. Note it will also have the effect of making the text labels proportionally smaller to the graphics, so you may want to increase the font size too.

**Transparency**

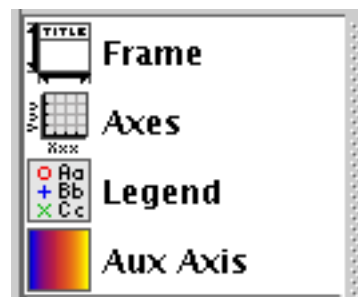
For technical reasons transparent markers cannot easily be rendered when a plot is exported to PostScript. In some cases the plot is done using a bitmap in the PostScript output to permit transparency and in some cases the points are just plotted opaque. PDF does a bit better, but

the compositing of transparent shapes is sometimes a bit different on the screen and rendered to a PDF. It's a good idea to check the output of screen exports by looking at the produced file - if it doesn't look like it should do, setting the **Force Bitmap** option will probably make sure it does, though this will also pixellate the plotted symbols. There is more discussion of this point in the subsections for the various shading modes.

### File Size

In some cases (2D and 3D scatter plots with many thousands of points or more) output EPS files can get extremely large; the size scales with the number of points drawn, currently with a factor of a few hundred bytes per point. In some cases you can work round this by plotting some points as transparent so that the plot is rendered as a bitmap (see the discussion of transparency above) which scales as the number of pixels rather than the number of points. The Gzipped EPS format helps somewhat (though can be slow); PDF output is better still. Even PDF files may be unmanageably large for very many points however.


## A.4.3 Fixed Controls



### Fixed controls in control panel stack

Fixed controls are those controls that appear at the top of the stack in the control panel, and do not have a checkbox or a drag handle. They are used to configure the overall appearance of the plot, independent of any particular plot layer or data set. Those common to all plot types are described in the following subsections, though some plot types may also have their own.

#### A.4.3.1 Frame Control

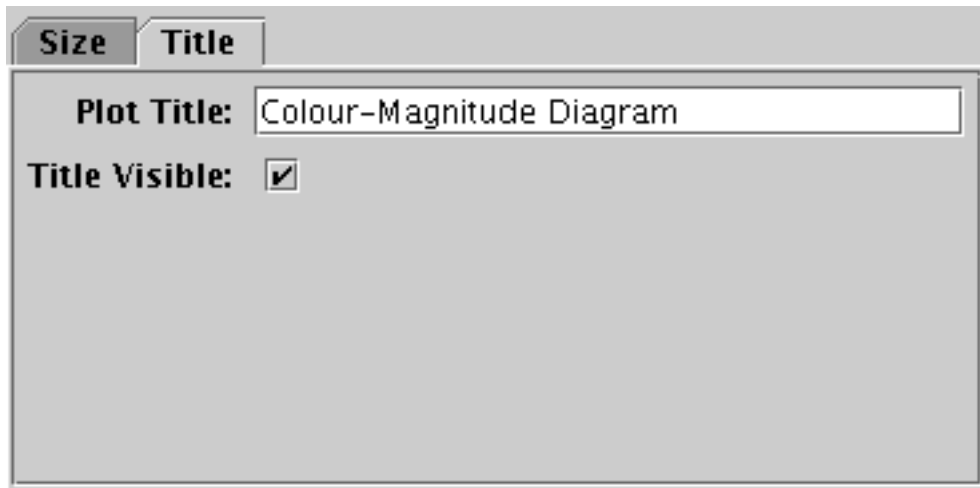
The **Frame** control () controls the graphical area on which the plot graphic is drawn. It contains two tabs, **Size** and **Title**.

Size		Title
Outer width:	800	pixels
Outer height:	600	pixels
Top border:		pixels
Left border:		pixels
Bottom border:		pixels
Right border:		pixels

### Frame control Size tab

The **Size** tab allows you to set the geometry of the generated plot in pixels. By default, the plot is simply resized to fit the current size and shape of the window. However, if you fill in the **Outer width** and/or **Outer height** fields, the size will instead be fixed to the given dimension in pixels. Note if you pick a large value, you may need to manually resize the window to see all of the plot. Similarly, the **border** fields fix the number of pixels surrounding the data region of the plot; if these are left blank, as by default, these borders are sized to accommodate the things that have to fit in them (such as axis annotations and the Aux axis colour ramp).


The main use for this tab is to fix the size and alignment of the generated images if you want to export a series of similar plots.




#### Frame control Title tab

The **Title** tab allows you to give a title for the plot. If text is filled in the **Plot Title** field, and if the **Title Visible** checkbox is checked, then the given text will appear at the top of the plot.

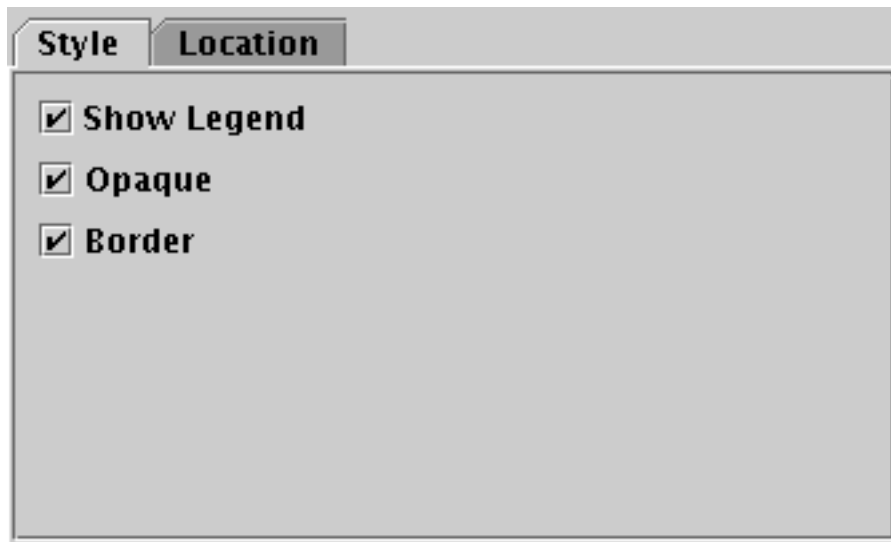
#### A.4.3.2 Axes Control

The **Axes** control () has a number of tabs specific to the particular plot type. These tabs allow you to configure things like the axis ranges, linear/log axis scaling, axis labelling, grid drawing, details of the navigation controls and so on. The different plot types have axis controls that differ significantly from each other; the specific axis controls are described separately along with each plot type.

#### A.4.3.3 Legend Control

The **Legend** control () is always visible in the plot Control Panel. It allows you to configure whether and how the plot legend appears.

This control has two tabs.



### Legend control Style tab

The **Style** tab configures the appearance of the legend panel, and has the following options:

#### Show Legend

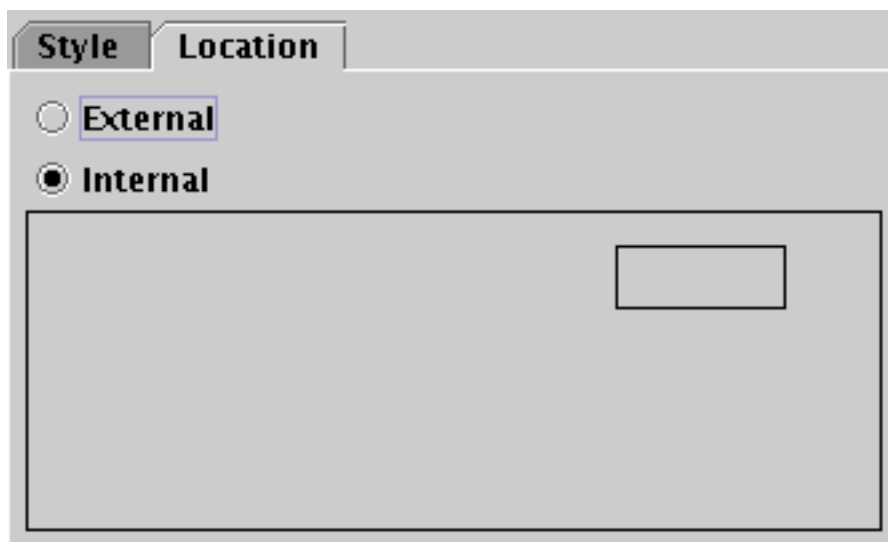
Whether the legend appears in the plot. This is set automatically: if only one data set is present no legend is shown, but once multiple datasets are present the legend is visible. But the option can be overridden manually using this control.

#### Opaque

If true, the legend has an opaque white background. If false, it is transparent, and any plot data underneath it is visible.

#### Border

If true, a black line border is shown around the legend. If false, there is no border.




### Legend control Location tab

The **Location** tab configures where on the plot the legend is drawn (if present). There are two options, **External** and **Internal**. If **Internal** is chosen, then a control is activated showing a small box inside a large box. Drag the small box around with the mouse to change the position of the legend inside the plot bounds.

Note the font used in the legend is currently controlled by the font from the **Axes** selector.

### A.4.3.4 Aux Axis Control

The **Aux Axis** control () is only visible when at least one layer is using the shared colour map.

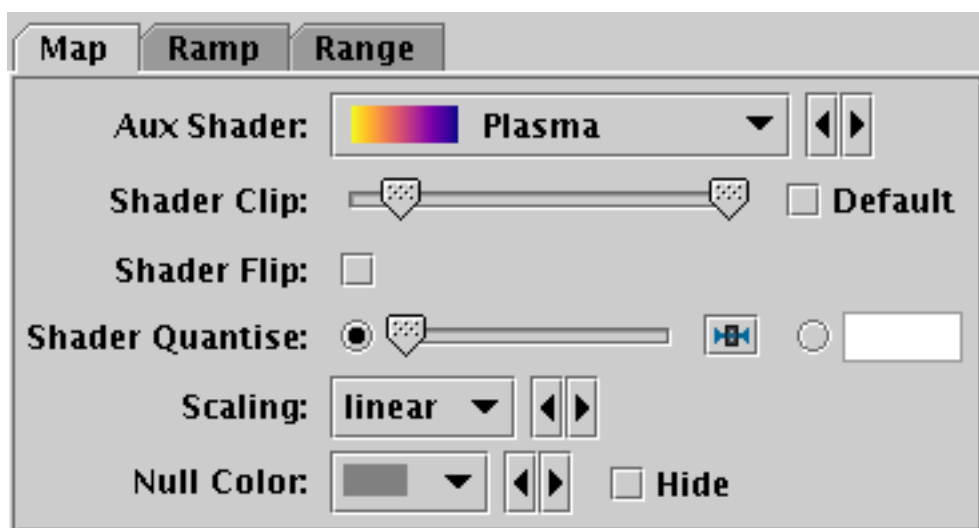
The following plot layers do this:

- Aux colour mode (Appendix A.4.6.6)
- Weighted colour mode (Appendix A.4.6.7)
- Density form (Appendix A.4.5.13)
- Sky Density form (Appendix A.4.5.14)

These all colour parts of the plot in a way that is quantitatively significant, and the Aux Axis control gives you a chance to control the details of the value-colour mapping, and to display the mapping in a colour ramp displayed beside the plot. Some other layer types (e.g. Density mode) also shade the plot according to numeric values, but use their own colour map, and don't display the colour ramp.

If no layer is using the shared colour map, then this control will not appear in the stack.

When present, this control has three following tabs.



#### Aux axis control Map tab

The **Map** tab controls the aux axis colour map. It has the following options:

##### Aux Shader

Select the colour map from a list of options.

##### Shader Clip

Select a sub-range of the full colour map above. If the **Default** checkbox is checked, then all or most of the colour ramp from the **Shader** control is used. If you want to configure the range of colours from the map yourself, uncheck the Default checkbox, and slide the handles in from the end of the slider to choose exactly the range you want.

The default range is clipped at one end for colour maps that fade to white, so that all the plotted colours will be distinguishable against a white background. That is generally a good idea for scatter-type plots, but may not be so good for plots where the whole plotting surface is coloured in, like density maps or spectrograms, so in that case you might want to uncheck Default and leave the handles at the extreme ends of the slider.

##### Shader Flip

Whether the aux scale should map forwards or backwards into the colour map.

##### Shader Quantise

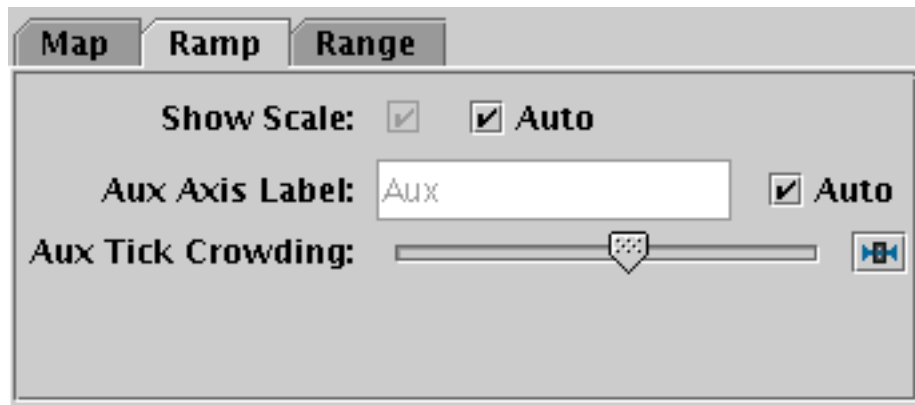
Allows the colour map to be quantised. By default, the colour map is effectively continuous. If you slide the slider to the right, or enter a value in the text field, the map will be split into a decreasing number of discrete colours. This can be used to generate a contour-like effect, and may make it easier to trace the boundaries of regions of interest by eye.

### Scaling

Determines the function used to map the range of aux data values onto the colour map. Options are **linear**, **logarithmic**, **square** and **square root**.

### Null Colour

What colour should be used to represent points with a null value for the aux data coordinate. If the associated **Hide** option is selected, then those points will not appear in the plot at all.



### Aux axis control Ramp tab

The **Ramp** tab controls the display and annotation of the colour ramp that displays the colour map on the plot. It has the following options:

#### Show Scale

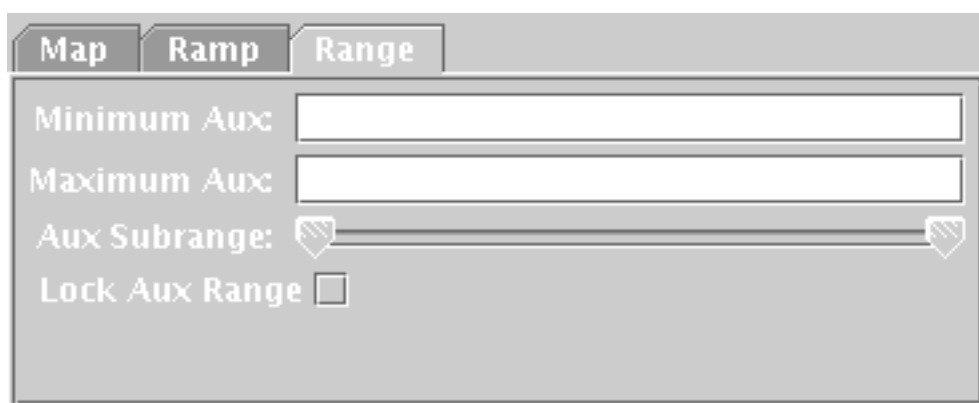
Whether the aux scale ramp is visible; if so an appropriately labelled colour ramp appears at the right of the plot. The associated **Auto** option makes this decision automatically: if any aux data is plotted, the scale will appear, otherwise it won't. Deselect Auto if you want to determine visibility by hand.

#### Aux Axis Label

Selects the axis label to be displayed near the aux colour ramp if it is visible. The associated **Auto** option, if selected, uses the name of one of the coordinates supplying aux data; deselect Auto if you want to enter a label by hand.

#### Aux Tick Crowding


The slider influences how many tick marks are drawn on the colour ramp.



### Aux axis control Range tab



The **Range** tab lets you enter lower and upper values for the aux data range by hand, and provides a double slider to restrict the range within these limits. If either the lower or upper range is left blank, it will be determined from the data.

The **Lock Aux Range** checkbox controls whether the aux range is automatically updated as the plot is adjusted (for instance as you navigate around it with the mouse). If locked, the range will stay the same, but otherwise (the default) it is dynamically updated for the current view of the plot. This checkbox is a duplicate of the  toggle button on the plot window toolbar described in Appendix

A.4.2.


Note the font used for labelling the aux axis is currently controlled by the font from the **Axes** selector.


## A.4.4 Layer Controls

The layer controls are controls in the Control Stack that can be added, removed and moved around to determine what layers go to make up the contents of a plot. You can have zero, one or several of each. Which ones are available is dependent on which plot type you are using (for instance the Spectrogram control is only available for the Time plot). Add instances of each control to the control stack by using the appropriate button from the control panel toolbar (the one in the lower half of the window) or the corresponding item in the **Layers** menu.

The available layer controls are described in the following subsections. More may be added in future releases.

### A.4.4.1 Position Layer Control

The **Position** layer control () is available for all the plot types. Most plots start off with one of

these in the stack by default, and you can add a new instance by using the **Add Position Control** () button in the control panel toolbar, or the corresponding item in the **Layers** menu.

This is the control which is used for most of the data plotting in the plotting windows. Each instance of this control in the stack does plotting for a particular set of positions from a single table. The set of positions is defined in the **Positions** tab as a column name or expression for each plot coordinate (e.g. for X and Y in a plane plot). However, the control can generate multiple layers from these positions; the **Subsets** tab controls which subsets are plotted and how each one is identified, and the **Form** tab provides many options for what graphics will be plotted based on the positions.

This control is a Table Data control as described in Appendix A.4.2.2. That section explains the **Subsets** tab; the **Position** and **Form** tabs are described in more detail below.

The image shows the 'Position' tab of the Position layer control. It has three tabs: 'Position', 'Subsets', and 'Form'. The 'Position' tab is active. It contains a 'Table:' dropdown menu set to '1: 6dfgs\_E7.fits'. Below this are two input fields: 'X:' with the value 'RMAG - BMAG' and 'Y:' with the value 'RMAG'. Each input field has a dropdown arrow and two navigation buttons (left and right arrows).

### Position tab of Position layer control, for Plane plot

In the **Position** tab you enter the base position coordinates for each plotted point. This generally means selecting a table and providing a value (table column or expression) for each positional coordinate. When you first open a plot window, TOPCAT gives you a table layer control by default, and attempts to fill in the positional coordinates with some reasonable values from the table (for instance the first few numeric columns).


Note the details of the **Position** tab will be different for different plot types, for instance the Cube plot has a Z coordinate field alongside X and Y.

The image shows the 'Form' tab of the Position layer control. It has three tabs: 'Position', 'Subsets', and 'Form'. The 'Form' tab is active. At the top is a toolbar with icons for various plot forms: a red diamond with a green cross, a red square with a green cross, a red rectangle with a green cross, a red arrow with a green cross, a red crosshair with a green cross, a red circle with a green cross, a red line with a green cross, a red line with a green cross and a red dot, a red circle with a green cross and a red dot, and a trash can icon. Below the toolbar is a stack of forms on the left: 'Mark' (selected) and 'Vector'. The 'Mark' form is expanded, showing its configuration options. The 'Shading' section has a 'Mode:' dropdown set to 'auto'. The 'Global Style' section has 'Shape:' set to a black circle, 'Size:' set to '1', and 'Colour:' set to a grey box with a 'By Subset' checkbox checked. The 'Subset Styles' section has a 'Subset:' dropdown and a 'Visible' checkbox.

### Form tab of Position layer control

The **Form** tab lets you define how the specified data set is plotted. The stack on the left gives a list of forms currently being plotted, and the panel on the right shows the detailed configuration for the currently selected form.

When first added, the stack contains a single entry, **Mark**, which plots a marker of a given fixed

shape and size. The colour is by default determined by the setting in the Subsets tab. For a simple scatter plot, this is all that you need. However, there are a number of other forms that you can plot as well or instead of the simple markers - vectors, error bars, ellipses, contours, text labels etc. You add a new form to the stack by clicking on one of the buttons in the toolbar in the tab. You can remove a form with the **Remove** () button in the same toolbar. You can also activate/deactivate

them with the checkbox and move them up and down with the drag handle as usual. The list of forms that are available depends on the plot type; the full list is in Appendix A.4.5.

The detail panel of each form depends on the form itself. It is divided into the following panels, though not all forms have all the panels.

### Shading

The shading mode controls how points are shaded based on their chosen colour. The various options are described in Appendix A.4.6. Depending on the mode there may be more settings to fill in here.



### Coordinates


If additional coordinates are required for this form, for instance the size of error bars, you need to enter the column or expression here. Each coordinate effectively adds another dimension to the plot. Many forms, like Mark, do not require any additional coordinates.

### Global/Subset Styles

Controls the style details for the chosen form; see Appendix A.4.2.2.

## A.4.4.2 Pair Position Layer Control

The **Pair Position** layer control () allows you to plot symbols linking two positions in the plot space from the same table. You can add one of these controls to the stack by using the **Add Pair Control** () button in the control panel toolbar, or the corresponding item in the **Layers** menu.

This control is particularly useful for visualising the results of a crossmatch, and it is used automatically by the **Plot Result** () option offered by some of the Match Window results. If you want to plot pairs of points from different input tables, you first have to create a joined table by using one of the crossmatch functions.

It works in a similar way to the Single Position Layer Control, but the **Position** tab has fields for not one but two sets of coordinates to fill in.

This control is a Table Data control as described in Appendix A.4.2.2. That section explains the **Subsets** tab; the **Position** and **Form** tabs are described in more detail below.

The image shows the 'Position' tab of the 'Pair Position layer control' dialog. It has three tabs: 'Position', 'Subsets', and 'Form'. The 'Position' tab is active. It contains the following fields:

- Table:** A dropdown menu showing '7: match(5,6)'.
- Data Sky System:** A dropdown menu showing 'Equatorial' with left and right arrow buttons.
- Lon1:** A text field containing 'alpha' with a dropdown arrow and left/right arrow buttons.
- Lat1:** A text field containing 'delta' with a dropdown arrow and left/right arrow buttons.
- Lon2:** A text field containing 'RAJ2000' with a dropdown arrow and left/right arrow buttons.
- Lat2:** A text field containing 'DEJ2000' with a dropdown arrow and left/right arrow buttons.

### Position tab of Pair Position layer control, for Sky plot

In the **Position** tab you enter the pair of position coordinates for each plotted pair. This generally means selecting a table and providing a value (table column or expression) for the first and second sets of positional coordinates.



Note the details of the **Position** tab will be different for different plot types, for instance the Sky plot has **Lon1**, **Lat1**, **Lon2** and **Lat2** fields for the first and second longitude/latitude pairs, while a Plane plot has **X1**, **Y1**, **X2** and **Y2**.

The image shows the 'Form' tab of the 'Pair Position layer control' dialog. It has three tabs: 'Position', 'Subsets', and 'Form'. The 'Form' tab is active. It contains the following elements:

- Top Bar:** Three icons: a red line with green dots, a red circle with a green dot, and a trash can.
- Left Panel:** A list of forms with checkboxes and icons. The 'Link' form is selected and highlighted in blue.
  - Marks:** A checkbox with a double-headed vertical arrow and a red circle icon.
  - Link:** A checkbox with a double-headed vertical arrow and a red line icon.
- Right Panel:** Configuration options for the selected form.
  - Shading:** A section with a 'Mode' dropdown set to 'flat' and a red star icon.
  - Global Style:** A section with a 'Colour' dropdown and a 'By Subset' checkbox which is checked.
  - Subset Styles:** A section with a 'Subset' dropdown, a small square icon, and a 'Visible' checkbox.

### Form tab of Pair Position layer control

The **Form** tab lets you define how the specified data set is plotted. The list on the left gives a list of forms currently being plotted, and the panel on the right shows the detailed configuration for the currently selected form.

 ) which draws the points at both ends of the pair, and Link (  ) which draws a line linking them. You can configure these, and select them on or off, separately. The detail panel for these forms are dependent on the form itself and are described in more detail in Appendix A.4.5, but the detail panels are divided into these parts:


#### Shading

The shading mode controls how points are shaded based on their chosen colour. The various options are described in Appendix A.4.6. Depending on the mode there may be more settings to fill in here.

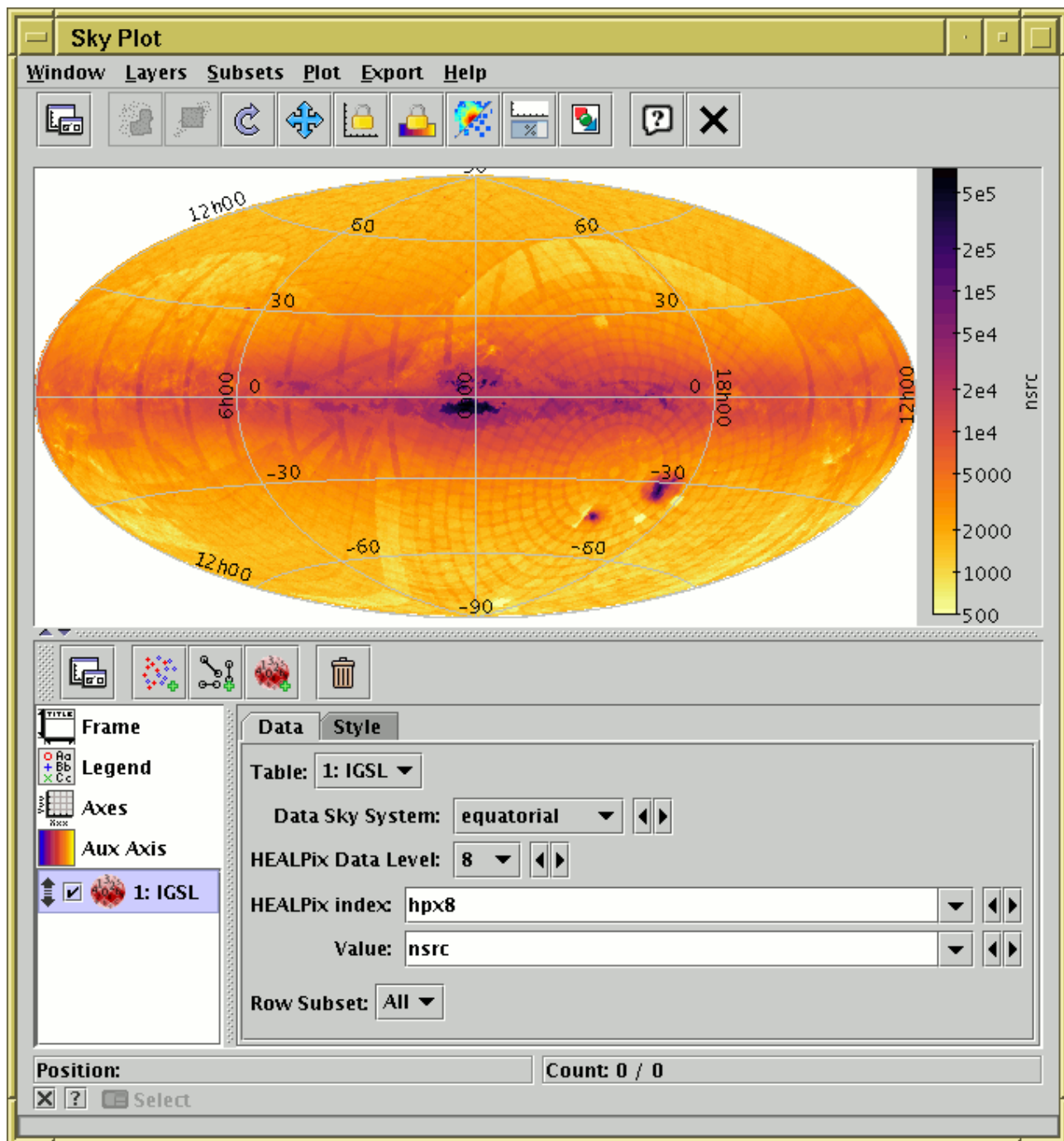
#### Global/Subset Styles

Controls the style details for the chosen form; see Appendix A.4.2.2.

### A.4.4.3 Healpix Layer Control

The **Healpix** layer control (  ) is available from the Sky Plot Window for plotting tables that represent HEALPix maps on the celestial sphere. Each row in the table must represent a single healpix tile, and a value from that row is used to colour the corresponding region of the sky plot. The resolution (healpix order) of the input table is supplied or may be guessed in order to do the plot, but the plot may be drawn at a degraded order (bigger pixels) if required.

Note this is different from the SkyDensity form, which takes a table containing sky positions, and represents that as a density map on the healpix grid by gathering the rows up into bins. The control described here works on a table for which that binning has already been done, for instance as a prepared sky map data product, or by executing a suitable (GROUP BY healpix\_index) database query.



**Sky Plot Window with a Healpix layer**

This control has two tabs, **Data** and **Style**.

The screenshot shows a software window with two tabs: 'Data' and 'Style'. The 'Data' tab is active. It contains the following controls:

- Table:** A dropdown menu showing '1: IGSL'.
- Data Sky System:** A dropdown menu showing 'equatorial' with left and right arrow buttons.
- HEALPix Data Level:** A dropdown menu showing '8' with left and right arrow buttons.
- HEALPix index:** A text input field containing 'hpx8' with a dropdown arrow and left/right arrow buttons.
- Value:** A text input field containing 'nsrc' with a dropdown arrow and left/right arrow buttons.
- Row Subset:** A dropdown menu showing 'All'.

### Healpix control Data tab

The **Data** tab lets you specify the table and columns containing the healpix tile data. It has the following fields:

#### Table

The table supplying the data.

#### Data Sky System

The sky coordinate system of the grid on which the pixels in the input table are laid out.

#### HEALPix Data Level

HEALPix level of the (implicit or explicit) tile indices. Only values up to 13 are currently supported. This must be the value assumed by the data in the input table. If **-1** is selected, an attempt is made to determine the correct level from the data; this may or may not be successful. Using the wrong value here will result in a nonsensical plot, or no plot at all.

#### HEALPix index

Column in the input table giving HEALPix index value. If this is left blank, then the row index is used, i.e. pixel #0 is assumed to be in the first row etc. This only works if the input table has full sky coverage and the rows are in the correct sequence.

#### Value

The column, or other expression, giving the colour to plot for the pixel at each row.

#### Row Subset

May be used to restrict the plot to one of the defined row subsets.

To control the colour map used to colour the sky tiles, use the Aux  fixed control.

The screenshot shows the 'Style' tab of the same software window. It contains the following controls:

- Degrade:** A numeric input field showing '1' with up and down arrow buttons.
- Combine:** A dropdown menu showing 'sum' with left and right arrow buttons.
- Transparency:** A slider control with a dotted pattern icon on the left and a zoom icon on the right.

## Healpix control Style tab

The **Style** tab lets you configure additional details of the plot's appearance. It has the following fields:

### Degrade

Controls the resolution at which the pixel grid is actually plotted. If the value is zero, then the grid plotted is the same as that of the input data, but if a positive value is supplied then the healpix order will be reduced by that many. Each increment means that 4 pixels from the previous order will be combined into one. The way that combination is done is controlled by the **Combine** option.


### Combine


Determines how pixels are combined together when the resolution is **Degraded**. Options are currently **sum**, **mean**, **min**, **max**.

### Transparency

Adjusts the transparency of the filled area.

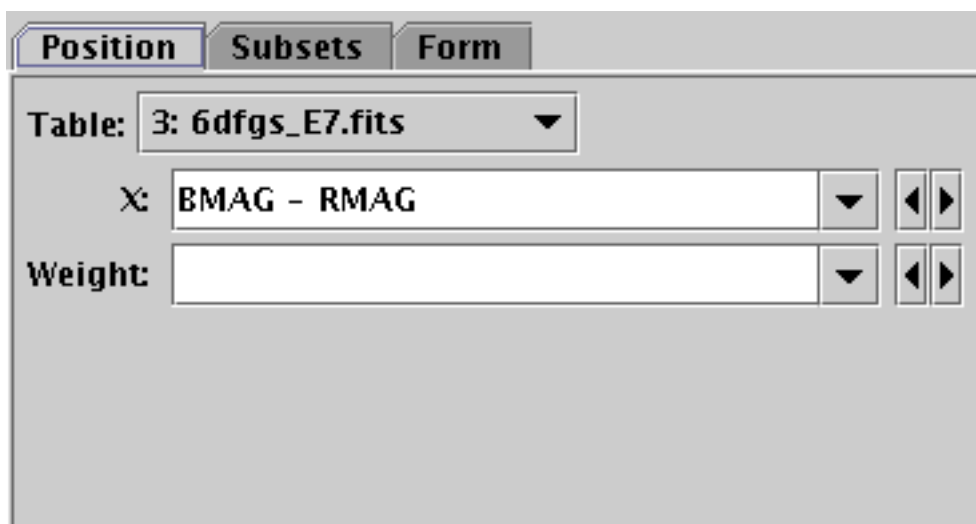
## A.4.4.4 Histogram Layer Control

The **Histogram** layer control (  ), allows you to make a 1-dimensional histogram-like plots

(weighted and unweighted histograms, various kinds of Kernel Density Estimates). It is the main control in the Histogram plot window, but is also available in the Plane plot window, so you can overplot a histogram on a scatter plot, if you can think of some reason to do that. The Histogram window starts off with one of these controls in the stack by default, but as usual you can add more by using the **Add Histogram Control** (  ) button in the control panel toolbar, or the corresponding item in the **Layers** menu.

Each instance of this control in the stack can make histogram-like plots of a particular quantity from a particular table. The value to be histogrammed, and optionally an associated weighting term, are defined in the **Positions** tab using column names or expressions. However, the control can generate multiple layers from these coordinates; the **Subsets** tab controls which subsets are plotted and how each one is identified, and the **Form** tab provides many options for what graphics will be plotted based on the sample values specified in the Position tab.

This control is a Table Data control as described in Appendix A.4.2.2. That section explains the **Subsets** tab; the **Position** and **Form** tabs are described in more detail below.

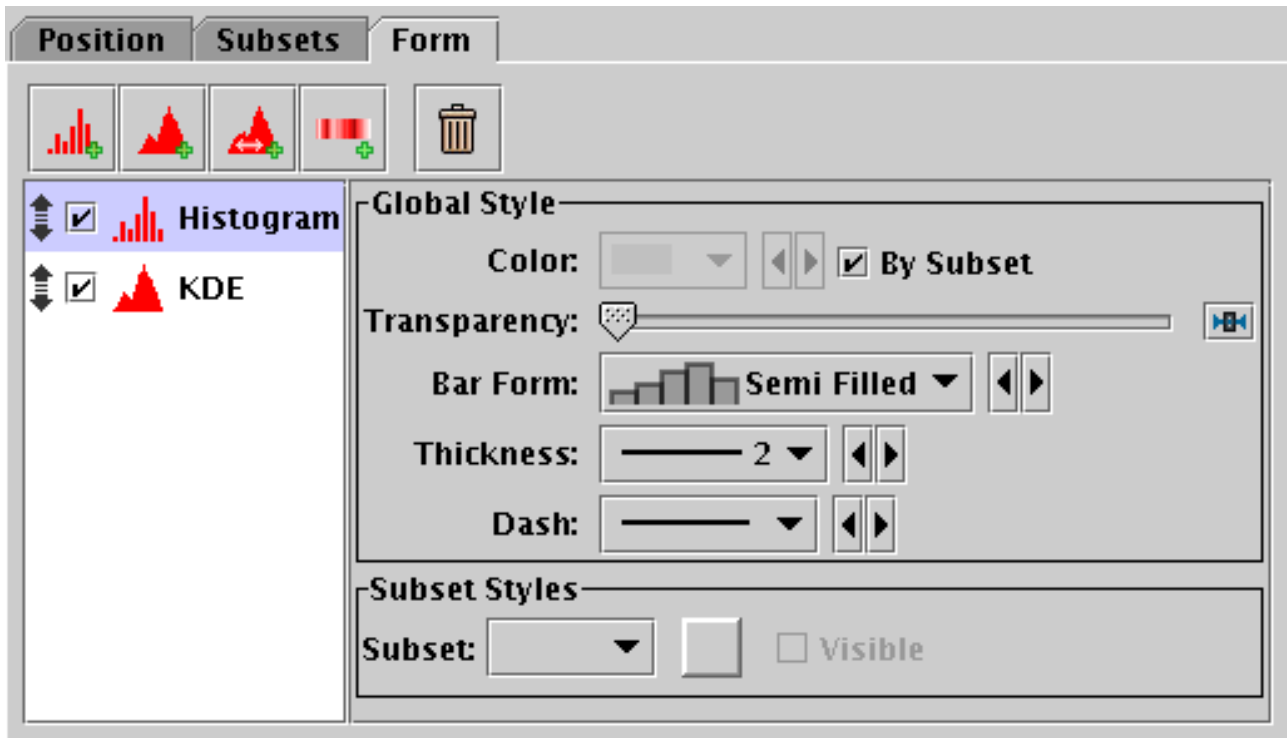


The screenshot shows the 'Position' tab of the Histogram Layer Control dialog. It has three tabs: 'Position', 'Subsets', and 'Form'. The 'Position' tab is active. It contains a 'Table' dropdown menu with the value '3: 6dfgs\_E7.fits'. Below it, there is a label 'X:' followed by a text input field containing 'BMAG - RMAG'. To the right of this field are three buttons: a dropdown arrow, a left arrow, and a right arrow. Below the 'X' field, there is a label 'Weight:' followed by an empty text input field. To the right of this field are three buttons: a dropdown arrow, a left arrow, and a right arrow.




### Position tab of Histogram layer control



In the **Position** tab you enter the value of the **X** value, the one whose values along the X axis will be used to generate the plots. This may be a table column name or an expression. You may also optionally fill in a column name or expression in the **Weight** field. This weights the values as they are counted; each table row contributes a height value of the weighting coordinate to the bin into which its X coordinate falls. If Weight is not filled in, a value of 1 is assumed.






### Form tab of Histogram layer control


The **Form** tab lets you define how the specified data set is plotted. The stack on the left gives a list of forms currently included in the plot, and the panel on the right shows the detailed configuration for the currently selected form.

When first added, the stack contains a single entry, **Histogram**, which plots a simple histogram. The colour is by default determined by the setting in the **Subsets** tab. This may be all you need, but if you want to use other histogram-like plots such as Kernel Density Estimates, you can add new forms to the stack by clicking on one of the buttons in the toolbar in the tab. You can remove a form with the **Remove** (  ) button in the same toolbar. You can also activate/deactivate them with the

checkbox and move them up and down with the drag handle as usual. The available forms are currently the  Histogram and some variants on the idea of a Kernel Density Estimate: 

KDE,  KDE,  KNN and  Densogram.

The **Global Style** and **Subset Style** sub-panels control shared and per-subset configuration specific to the selected form as described in Appendix A.4.2.2.


Histograms have some additional configuration items, as described in the Bins (  ) fixed control.

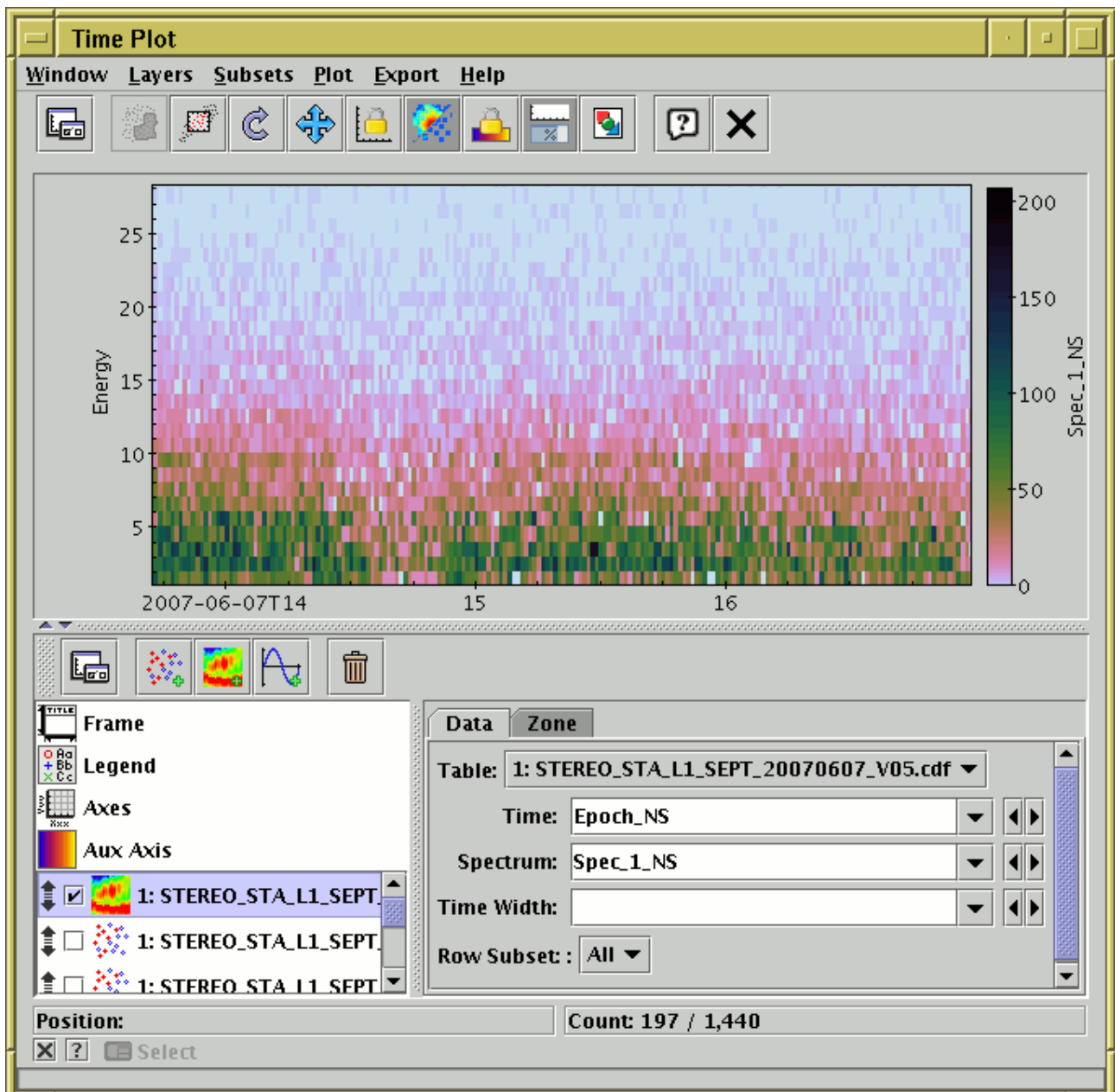
When a histogram layer control is used in the Histogram window, those configuration items are found in the **Bins** fixed control, where they control the settings for all the histogram layers in the plot. However, if you use a histogram layer in a Plane plot, those items will appear as additional

items in the **Form** tab and can be controlled separately for each histogram.

#### A.4.4.5 Spectrogram Layer Control

*This control is experimental. It is missing some important features, and the interface may be changed in a future version.*

The **Spectrogram Layer Control**  plots a spectrum at successive (usually, but not necessarily, regularly-spaced) points in a time series. It is only available for the *experimental* Time Plot Window.



**Time Plot window with a Spectrogram layer**

This control has only one layer-specific tab, **Data**, described below. The **Zone** tab is described in Appendix A.4.12.1.

### Spectrogram control Data tab

The **Data** tab allows you to specify which values from a table will generate a spectrogram. It has the following fields:

#### Table

The table supplying the data.

#### Time

A table column or expression giving the epoch coordinate at which spectra are located. This should normally be a time-typed column; if it is simply of numeric type it will be interpreted as seconds since 1 Jan 1970.

#### Spectrum


An array-valued table column giving the spectral data.

#### TimeWidth


A table column or expression (variable or constant) giving the temporal coverage of a plotted spectrum. If not filled in, it is assumed to be the most common (median) difference between time points.

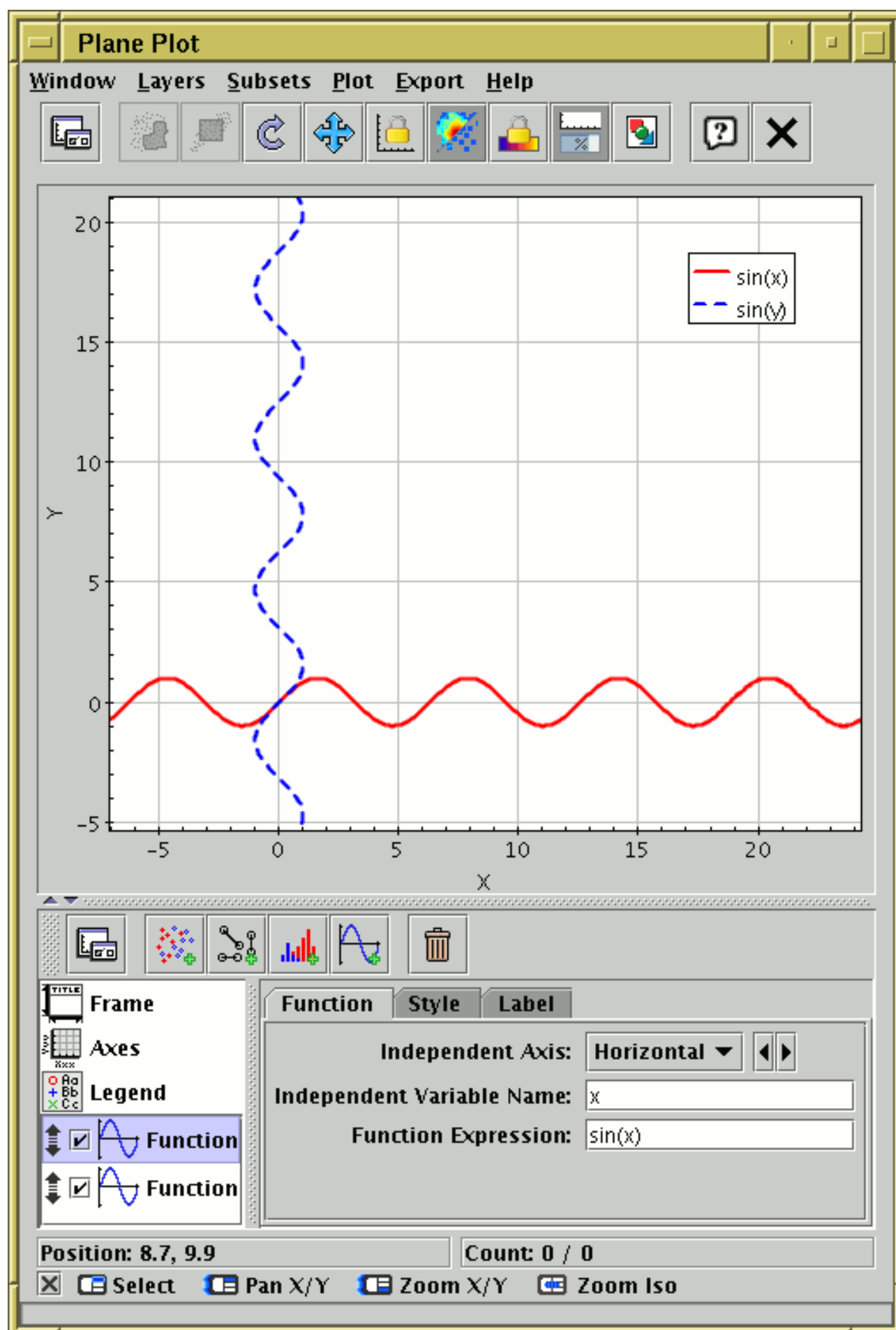
#### Row Subset

The subset for which the spectrum should be plotted. To plot multiple subsets (not usually useful with this kind of plot) you would need multiple spectrogram layer controls in the stack.

To control the colour map used to represent the spectral values, use the Aux  fixed control.

### A.4.4.6 Function Layer Control

The **Function** layer control () is only available for the Plane and Histogram plots. Use the **Layers** menu item or corresponding toolbar button to add an instance to the control stack.



## Plane plot with two function layers plotted, one as a function of Horizontal axis value, and one as a function of Vertical axis value

This control has three tabs, **Function**, **Style** and **Label**, described below.

The screenshot shows the 'Function' tab of a control interface. It has three tabs: 'Function', 'Style', and 'Label'. The 'Function' tab is selected. It contains three fields: 'Independent Axis' with a dropdown menu set to 'Vertical' and navigation buttons; 'Independent Variable Name' with a text input field containing 'y'; and 'Function Expression' with a text input field containing 'sin(y)'.

### Function control Function tab

The **Function** tab defines the function to be plotted and has the following fields:

#### Independent Axis

Which axis the independent variable varies along; options are currently **Horizontal** and **Vertical**.

#### Independent Variable Name

Name of the independent variable. This is typically  $x$  for a horizontal independent variable and  $y$  for a vertical independent variable, but any string that is a legal expression language identifier (starts with a letter, continues with letters, numbers, underscores) can be used.

#### Function Expression

An expression using TOPCAT's expression language in terms of the independent variable that defines the function. This expression must be standalone - it cannot reference any tables.

The screenshot shows the 'Style' tab of the control interface. It has three tabs: 'Function', 'Style', and 'Label'. The 'Style' tab is selected. It contains four fields: 'Colour' with a dropdown menu showing a red color swatch and navigation buttons; 'Thickness' with a dropdown menu showing a line segment and the value '3' and navigation buttons; 'Dash' with a dropdown menu showing a dashed line pattern and navigation buttons; and 'Antialiasing' with a checked checkbox.

### Function control Style tab

The **Style** tab configures the plotting style. Options are:

#### Colour

Colour of the line.

#### Thickness

Thickness of the line in pixels.

#### Dash

Dash pattern of the line - solid is the default, but various options are available.

**Antialiasing**

If true, lines are antialiased, which makes them look smoother on the screen or bitmapped export images. Has no effect on vector export images (PDF, EPS).

**Function control Label tab**

The **Label** tab allows you to choose the text that appears in the legend. Options are:

**Label**

Gives the label that will appear in the legend. By default the function expression is used, but if you want to override this you can deselect the associated **Auto** checkbox and enter your own value.

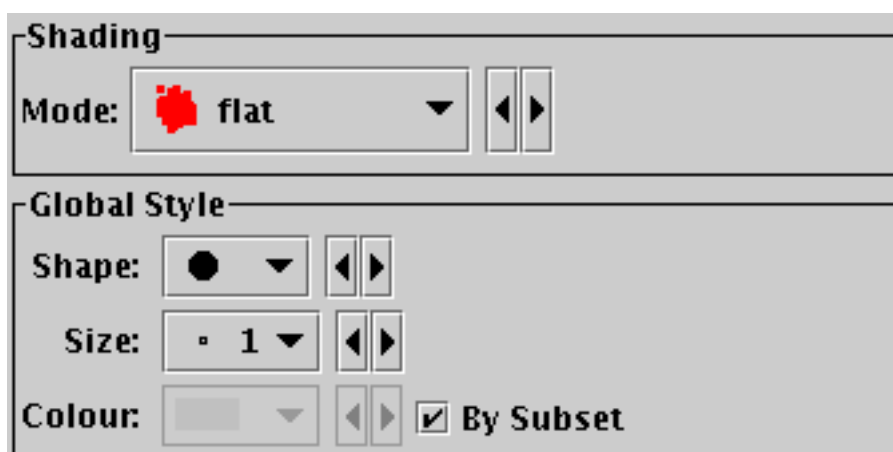
**In Legend**

If true, an entry for this function appears in the legend, if false it does not. Note the setting of this value does not affect whether the legend itself appears, which is controlled by the Legend control.


**A.4.5 Plot Forms**

Plot "forms" are the instructions for what shapes to actually paint on the plotting surface given some positional data from a table. The most obvious (Mark) is simply to plot a marker with a fixed size and shape at each data position. However, there is a range of other things you might want to do, including error bars, vectors, contours, text labels, ...

The options provided are described, with the additional configuration information required, in the following sections. Not all are available for every plot type.

**A.4.5.1 Mark Form**

### Mark form configuration panel

The **Mark** form (  ) simply plots markers with a fixed shape and size.

Configuration options are:

#### Shading Mode

See Appendix A.4.6.

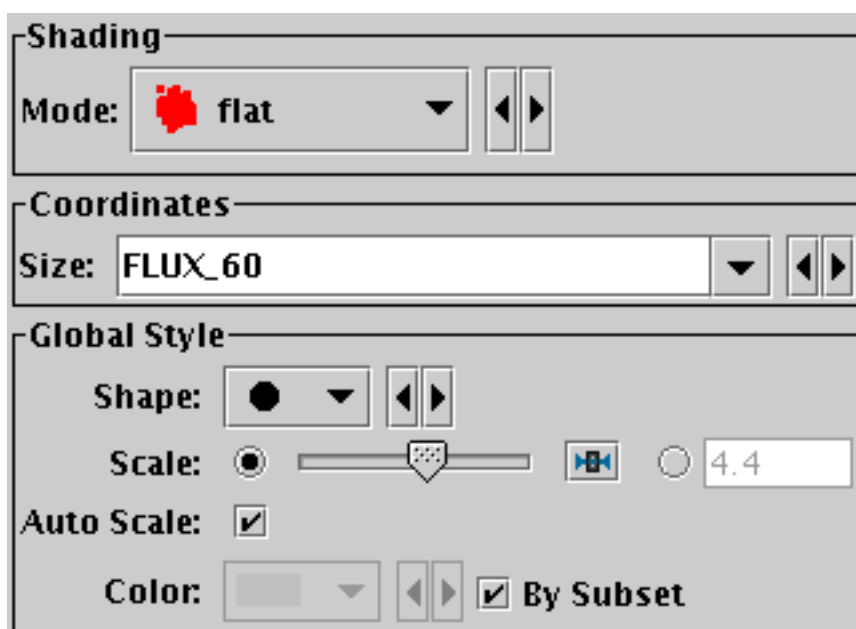
#### Shape

Marker shape from a list of options.

#### Size


Marker size in pixels.

### A.4.5.2 Size Form



The screenshot shows the 'Size form' configuration panel with three main sections: 'Shading', 'Coordinates', and 'Global Style'.  
 - **Shading**: 'Mode' is set to 'flat' with a red plus marker icon.  
 - **Coordinates**: 'Size' is set to 'FLUX\_60'.  
 - **Global Style**:  
 - 'Shape' is set to a solid black circle.  
 - 'Scale' has a radio button selected, a slider, and a numeric input field showing '4.4'.  
 - 'Auto Scale' is checked.  
 - 'Color' has a color selection box and a 'By Subset' checkbox which is checked.

### Size form configuration panel

The **Size** form (  ) plots markers of a fixed shape whose size varies according to a supplied data coordinate. The marker size thus represents an additional dimension of the plot.

The actual size of the markers depends on the setting of the **Auto Scale** option. If autoscaling is off, then the basic size of each marker is the input data value in units of pixels. If autoscaling is on, then the data values are gathered for all the currently visible points, and a scaling factor is applied so that the largest ones will be a sensible size (a few tens of pixels). This basic size can be further adjusted with the **Scale** slider.

Currently data values of zero always correspond to marker size of zero, negative data values are not represented, and the mapping is linear. An absolute maximum size on markers is also imposed. Other options may be introduced in future.

Note the scaling to size is in terms of screen dimensions (pixels). For sizes that correspond to actual data values, the Error form may be more appropriate.

The configuration options are:

**Shading Mode**

See Appendix A.4.6.

**Size**

The size coordinate data values. Fill this in with a column name or expression from the table just like for the positional coordinates. The units are either pixels or arbitrary, according to the **Auto Scale** setting.

**Shape**

Marker shape from a list of options.

**Scale**

Scales the size of variable-sized markers. Adjusting the slider will make all markers larger or smaller. Alternatively you can enter a scale factor in the text field.

**Auto Scale**


Determines whether the basic size of variable sized markers is automatically scaled to have a sensible size. If checked, then the sizes of all the plotted markers are examined, and some dynamically calculated factor is applied to them all to make them a sensible size (by default, the largest ones will be a few tens of pixels). If unchecked, the sizes will be the actual input values in units of pixels.

If auto-scaling is off, then markers will keep exactly the same screen size during pan and zoom operations; if it's on, then the visible sizes will change according to what other points are currently plotted.

**A.4.5.3 SizeXY Form**

The screenshot shows a configuration panel for the SizeXY form. It is organized into three main sections: **Shading**, **Coordinates**, and **Global Style**. In the **Shading** section, the 'Mode' is set to 'flat' with a red square icon. The **Coordinates** section contains two text boxes: 'X Size' with the value 'velX\*velX' and 'Y Size' with the value 'velY\*velY'. The **Global Style** section includes a 'Shape' dropdown (black square icon), a 'Scale' slider (set to 1), an 'Auto Scale' checkbox (checked), and a 'Color' dropdown (gray square icon) with a 'By Subset' checkbox (checked).

**SizeXY form configuration panel**

The **SizeXY** form (  ) plots a shaped marker whose width and height vary independently according to two supplied data coordinates. The marker shape can thus encode two additional dimensions of the plot.

The actual size of the markers depends on the setting of the **Auto Scale** option. If autoscaling is off, the basic dimensions of each marker are given by the input data values in units of pixels. If



autoscaling is on, the data values are gathered for all the currently visible points, and scaling factors are applied so that the largest ones will be a sensible size (a few tens of pixels). This autoscaling happens independently for the X and Y directions. The basic sizes can be further adjusted with the **Scale** slider.

Currently data values of zero always correspond to marker extent of zero, negative data values are not represented, and the mapping is linear. An absolute maximum size on markers is also imposed. Other options may be introduced in future.

Note the scaling to size is in terms of screen dimensions (pixels). For sizes that correspond to actual data values, the Error form may be more appropriate.

The configuration options are:

**Shading Mode**

See Appendix A.4.6.

**X/Y Size**

For width and height respectively, the size coordinate data values. Fill in with a column name or expression from the table just like for the positional coordinates. The units are either pixels or arbitrary, according to the **Auto Scale** setting.

**Shape**

Marker shape from a list of options.

**Scale**

Scales the size of variable-sized markers. Adjusting the slider will make all markers larger or smaller. Alternatively you can enter a scale factor in the text field.

**Auto Scale**

Determines whether the basic size of variable sized markers is automatically scaled to have a sensible size. If checked, then the sizes of all the plotted markers are examined, and some dynamically calculated factor is applied to them all to make them a sensible size (by default, the largest ones will be a few tens of pixels). If unchecked, the sizes will be the actual input values in units of pixels.


If auto-scaling is off, then markers will keep exactly the same screen size during pan and zoom operations; if it's on, then the visible sizes will change according to what other points are currently plotted.

#### A.4.5.4 Vector Form

The image shows a configuration panel for vector plots, divided into three sections:

- Shading:** Contains a 'Mode' dropdown menu set to 'flat' with a red square icon to its left. To the right of the dropdown are two small square buttons with left and right arrows.
- Coordinates:** Contains two text input fields. The first is labeled 'DeltaX:' and contains the text 'velX'. The second is labeled 'DeltaY:' and contains the text 'velY'. Each field has a dropdown arrow on its right and two small square buttons with left and right arrows.
- Global Style:** Contains several controls:
  - Arrow:** A dropdown menu showing a right-pointing arrow, with two small square buttons with left and right arrows.
  - Scale:** A control consisting of a radio button (selected), a horizontal slider with a diamond-shaped handle, a small square button with a magnifying glass, an unselected radio button, and a text field containing the number '3'.
  - Auto Scale:** A checkbox that is checked.
  - Colour:** A color selection dropdown menu, two small square buttons with left and right arrows, and a checked checkbox labeled 'By Subset'.

### Vector form configuration panel

The **Vector** form (  ) plots directed lines from the data position given delta values for the coordinates. The plotted markers are typically little arrows, but there are other options.

In some cases such delta values may be the actual magnitude required for the plot, but more often the vector data represents a value which has a different magnitude or is in different units to the positional data. As a convenience for this case, the plotter can optionally (and does by default) scale the magnitudes of all the vectors to make them a sensible size (so the largest ones are a few tens of pixels long). This scaling can be adjusted or turned off using the **Scale** and **Auto Scale** options below.

The configuration options are:

#### Shading Mode

See Appendix A.4.6.

#### DeltaX, DeltaY, ...

The coordinates of the changes in each coordinate which gives the vector. The coordinates here match the coordinates of the plot.

#### Arrow

Arrow shape selected from a range of options.

#### Scale

Changes the factor by which all vector sizes are scaled. If the arrows are too small, slide it right, if they are too big, slide it left. The slider scale is logarithmic. Alternatively, enter a fixed value in the text field.

#### Auto Scale


If selected, this option will determine the default arrow scale size from the data - it will fix it so that the largest arrows are a few tens of pixels long by default. That scaling can then be adjusted using the **Scale** slider. If unselected, then the default position of the Scale slider corresponds to the actual positions given by the submitted delta coordinates.

#### A.4.5.5 SkyVector Form

The screenshot shows the SkyVector form configuration panel. It is organized into three main sections:

- Shading:** Contains a 'Mode' dropdown menu currently set to 'auto', accompanied by a red arrow icon and navigation buttons.
- Coordinates:** Contains two text input fields. 'Delta Longitude' is set to 'pmRA' and 'Delta Latitude' is set to 'pmDE'. Both fields have dropdown arrows and navigation buttons.
- Global Style:** Contains several controls:
  - 'Arrow': A dropdown menu showing a right-pointing arrow.
  - 'Scale': A slider control with a central marker and a text field set to '1'.
  - 'Auto Scale': A checked checkbox.
  - 'Color': A color selection dropdown and a 'By Subset' checkbox.

#### SkyVector form configuration panel

The **SkyVector** form (  ) plots directed lines from the data position given relative offsets to longitude.cos(latitude) and latitude on the celestial sphere. The plotted markers are typically little arrows, but there are other options.

In some cases such delta values may be the actual magnitude required for the plot, but more often the vector data represents a value which has a different magnitude or is in different units to the positional data (for instance proper motions). As a convenience for this case, the plotter can optionally (and does by default) scale the magnitudes of all the vectors to make them a sensible size (so the largest ones are a few tens of pixels long). This scaling can be adjusted or turned off using the **Scale** and **Auto Scale** options below.

The configuration options are:

##### Shading Mode

See Appendix A.4.6.

##### Delta Longitude

Change in the longitude coordinate represented by the plotted vector. The supplied value is (if not auto-scaled) an angle in degrees, and is considered to be premultiplied by  $\cos(\text{Latitude})$ .

##### Delta Latitude

Change in the latitude coordinate represented by the plotted vector. The supplied value is (if not auto-scaled) an angle in degrees.

##### Arrow

Arrow shape selected from a range of options.

##### Scale


Changes the factor by which all vector sizes are scaled. If the arrows are too small, slide it right, if they are too big, slide it left. The slider scale is logarithmic. Alternatively, enter a fixed value in the text field.

### Auto Scale

If selected, this option will determine the default arrow scale size from the data - it will fix it so that the largest arrows are a few tens of pixels long by default. That scaling can then be adjusted using the **Scale** slider. If unselected, then the default position of the Scale slider corresponds to the actual values in degrees given by the submitted Delta Lon/Lat coordinates.

### A.4.5.6 Error Bars Form

#### Error Bars form configuration panel

The **Error Bars** form (  ) draws symmetric or asymmetric error bars in some or all of the dimensions of a 1-, 2- or 3-dimensional Cartesian plot. The shape of the error "bars" is quite configurable, including (for 2- and 3-d errors) ellipses, rectangles etc, aligned with the axes (for rotated ellipses and rectangles, see the XYEllipse form).

The configuration options are:

#### Shading Mode

See Appendix A.4.6.

#### X Positive Error, X Negative Error, ...


For each dimension of the plot, you can enter positive and/or negative error values as columns or expressions from the table. All of these values must be positive; any negative values will be ignored. If both positive and negative values are filled in for an axis, the errors will be asymmetric. If the negative value is blank (either because the coordinate is not filled in, or because its value is blank for that row), the error bars will be symmetric, i.e. the negative error bar will be the same size as the positive one. If you want to ensure only a positive error bar is plotted, enter "0" for the corresponding negative error.

#### Error Bar

Error bar shape from a list of options. Exact appearance will also depend on the dimensionality of the supplied errors.

### A.4.5.7 Ellipse Form

**Shading**

Mode:  flat


**Coordinates**




Major radius: EXTENSION\_FWHM\_MAJ

Minor radius: EXTENSION\_FWHM\_MIN


Orientation: POS\_POSANG

**Global Style**


Ellipse: 

Scale:    20

Auto Scale: ☒

Colour:  ☒ By Subset

#### Ellipse form configuration panel

The **Ellipse** form (  ) draws an ellipse (or rectangle, triangle, or other similar figure) defined by major and minor radii and an optional rotation angle.

The configuration options are:

#### Shading Mode

See Appendix A.4.6.

#### Primary Radius, Secondary Radius

The two principal radii for the ellipse, in the same units as the position coordinates. It doesn't matter whether the primary is larger than the secondary. If the Secondary Radius field is left blank, it is assumed to equal the Primary Radius field, i.e. the ellipses are circles.

#### Orientation

Position angle in degrees. This is the angle anticlockwise from the X axis to the primary radius. If the value is missing (either this field not filled in or blank in the data) it is considered to be zero.

#### Ellipse

Ellipse graphical representation, selected from a range of options (includes also rectangles, crosses etc).

#### Scale

Changes the factor by which all ellipse sizes are scaled. If the ellipses are too small, slide it right, if they are too big, slide it left. The slider scale is logarithmic. Alternatively, enter a fixed value in the text field.

#### Auto Scale

If selected, this option will determine the default ellipse scale size from the data - it will fix it so that the largest ellipses are a few tens of pixels long by default. That scaling can then be adjusted using the **Scale** slider. If unselected, then the default position of the Scale slider

corresponds to the actual positions given by the submitted ellipse radii coordinates.

#### A.4.5.8 SkyEllipse Form

**Shading**

Mode: flat

**Coordinates**

Primary Radius: EXTENSION\_FWHM\_MAJ/ 3600.

Secondary Radius: EXTENSION\_FWHM\_MIN/ 3600.

Position Angle: POS\_POSANG

**Global Style**

Ellipse:

Scale: 1

Auto Scale: ☐

Color: ☒ By Subset

#### SkyEllipse form configuration panel

The **SkyEllipse** ( ) draws an ellipse (or rectangle, triangle, or other similar figure) on the celestial sphere defined by major and minor radii and an optional rotation angle.

The configuration options are:

##### Primary Radius, Secondary Radius

The two principal radii for the ellipse. If auto-scaling is off, these are in units of **degrees**. It doesn't matter whether the primary is larger than the secondary. If the Secondary Radius field is left blank, it is assumed to equal the Primary Radius field, i.e. the ellipses are circles.

##### Position Angle

Orientation of the ellipse. This is the angle in degrees from the North pole to the primary axis of the ellipse in the direction of increasing longitude. If the value is missing (either this field not filled in or blank in the data) it is considered to be zero.

##### Ellipse

Ellipse graphical representation, selected from a range of options (includes also rectangles, crosses etc).

##### Scale

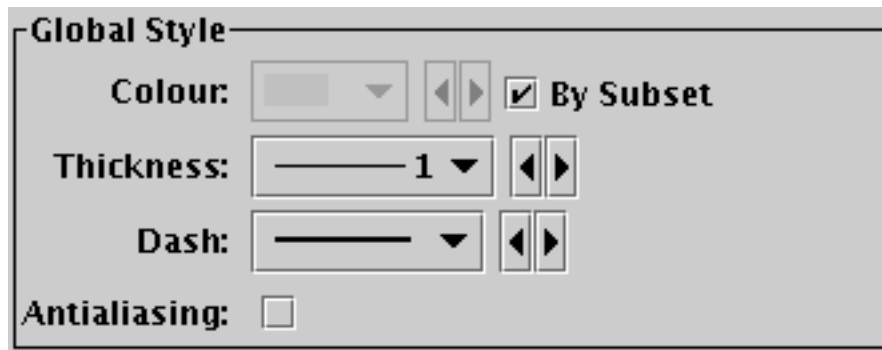
Changes the factor by which all ellipse sizes are scaled. If the ellipses are too small, slide it right, if they are too big, slide it left. The slider scale is logarithmic. Alternatively, enter a fixed value in the text field.

##### Auto Scale


If selected, this option will determine the default ellipse scale size from the data - it will fix it so that the largest ellipses are a few tens of pixels long by default. That scaling can then be

adjusted using the **Scale** slider. If unselected, then the default position of the Scale slider corresponds to the actual values in degrees given by the submitted ellipse radii coordinates.

#### A.4.5.9 Line Form



##### Line form configuration panel

The **Line** form (  ) draws a point-to-point line joining up the points making up the data set. Note that for a large and unordered data set this can lead to a big scribble on the screen.

The configuration options are:

##### **Thickness**

Line thickness in pixels.

##### **Dash**

Dash pattern. The line is solid by default.

##### **Antialiasing**

If true, pixels are blurred to give the line a smoother appearance.

#### A.4.5.10 Linear Fit Form

**Coordinates**

Weight:

**Global Style**

Color:

Thickness:

Dash:

Antialiasing: ☐

**Report**

Subset:


Equation:

c:

m:

Correlation:

### Linear Fit configuration panel

The **Linear Fit** form (  ) determines a least-squares best fit of a line to the data points, optionally weighted by a given value, and plots the corresponding line.

The configuration options are:

#### Weight

Weighting to apply to each point. If left blank, no weighting (equivalent to unit weighting) is applied. If independent sampling errors  $\epsilon$  are available for each point, a suitable value for this may be  $1. / (\epsilon * \epsilon)$  (equivalently  $\text{pow}(\epsilon, -2)$ ).

#### Thickness

Line thickness in pixels.

#### Dash

Dash pattern. The line is solid by default.

#### Antialiasing

If true, pixels are blurred to give the line a smoother appearance.

As well as drawing the line onto the plot, the calculated fitting coefficients are displayed at the bottom of the form configuration panel, under the heading **Report**. Note the coefficients are calculated by subset, and are only displayed for one subset at a time. To see the calculated values, select the subset of interest in the **Subset** selector. The reported items are:

#### Equation

The equation of the line. The basic linear equation is  $y = m * x + c$ , but if one or both of the axes are plotted logarithmically, the fit is performed to take account of this, and the equation is displayed accordingly.

**c**



The constant intercept of the line on the Y axis.

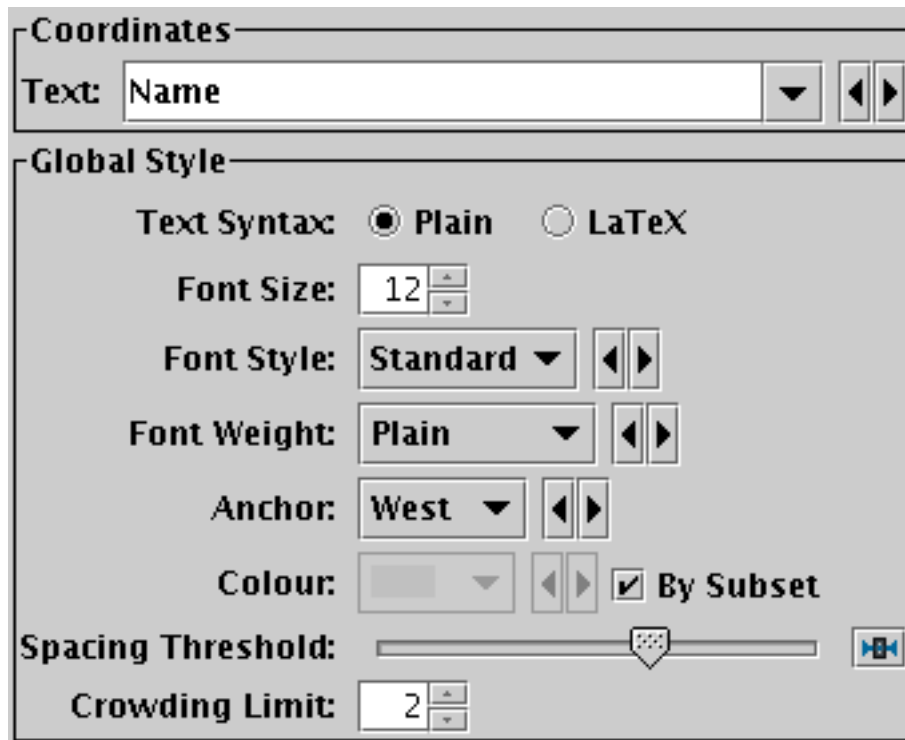
**m**

The line gradient.

### Correlation


Pearson's product moment correlation coefficient.

#### A.4.5.11 Text Label Form



The image shows a configuration panel for a text label form. It is divided into two main sections: 'Coordinates' and 'Global Style'. The 'Coordinates' section has a 'Text' field containing the word 'Name' and navigation buttons. The 'Global Style' section contains several settings: 'Text Syntax' with radio buttons for 'Plain' (selected) and 'LaTeX'; 'Font Size' set to 12; 'Font Style' set to 'Standard'; 'Font Weight' set to 'Plain'; 'Anchor' set to 'West'; 'Colour' with a color picker and a 'By Subset' checkbox; 'Spacing Threshold' with a slider; and 'Crowding Limit' set to 2.

#### Text Label form configuration panel

The **Text Label** form (  ) draws a text label by each point position.

The configuration options are:

#### Text

A column or expression from the table supplying the text to write on the plot. Any data type (string or numeric) is permitted.

#### Text Syntax

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at face value, but **Antialias** smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

#### Font Size

Size of the font in points.

#### Font Style

Style of the font - standard, serif or monospaced.

#### Font Weight

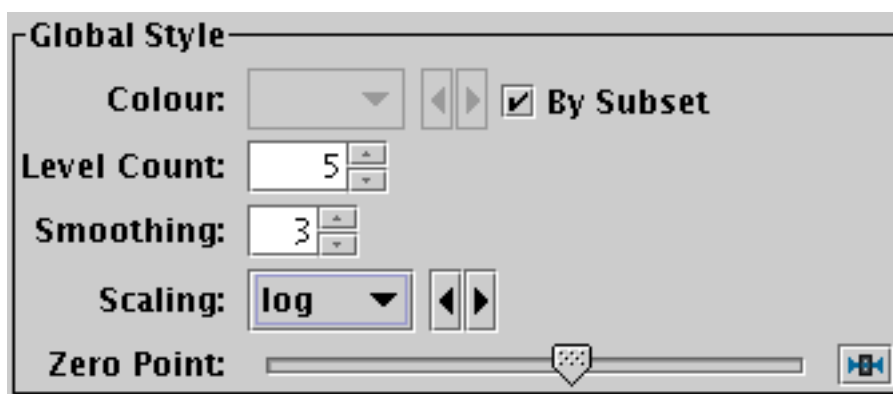
Whether the font is plain, bold or italic.


**Anchor**

The position of the text relative to the data position.

**Spacing Threshold****Crowding Limit**

These two options control how closely spaced labels can be. Labels which are too closely crowded together will simply not be shown, since overplotting many labels together ends up with them being illegible. The **Spacing Threshold** slider controls the smallest area that a group of labels can have to themselves - if there are too many in the same area, none will be drawn. Sliding it left allows more crowding and right allows less. The **Crowding Limit** controls the largest number of labels that can be in a group. Setting it to 2 for instance is useful if you want to see pairs of labels, even if the pair is close.

**A.4.5.12 Contour Form****Contour form configuration panel**

The **Contour** form (  ) plots position density contours. These provide another way (alongside the Auto, Density and Weighted shading modes) to visualise the characteristics of overdense regions in a crowded plot.

The contours are currently drawn as pixels rather than lines, so they don't look very beautiful in exported vector output formats (PDF, PostScript).

The configuration options are:

**Level Count**

The number of contours drawn. In fact this is an upper limit; if there is not enough variation in the plot's density, then fewer contour lines will be drawn.

**Smoothing**

The size of the smoothing kernel applied to the density before performing the contour determination. If set too low the contours will be too crinkly, and if too high they will lose definition.

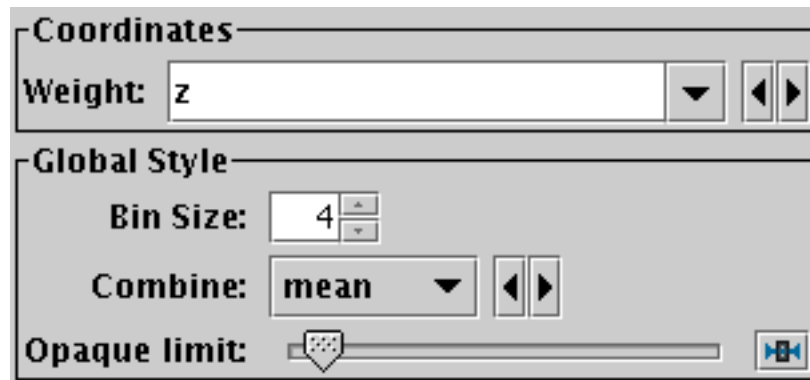
**Scaling**

How the smoothed density is treated before the contours levels are determined. Options are linear, logarithmic and equal area.

**Zero Point**


Determines the level at which the first contour (and hence all the others, which are separated from it by a fixed amount) are drawn. By sliding this from side to side you can sweep the contours over the density range and get a good idea of where interesting features lie.

### A.4.5.13 Density Form



The image shows a configuration panel for the Density form. It is divided into two main sections: 'Coordinates' and 'Global Style'. In the 'Coordinates' section, there is a 'Weight:' label followed by a text input field containing the letter 'z', and a small dropdown arrow to its right. The 'Global Style' section contains three controls: 'Bin Size' with a numeric input field set to '4' and small up/down arrows; 'Combine:' with a dropdown menu showing 'mean' and small left/right arrows; and 'Opaque limit:' with a horizontal slider bar and a small icon on the right.

#### Density form configuration panel

The **Density** form (  ) plots a density map (2-d histogram) on the pixel grid of the plot surface, coarsened by a configurable factor. You can optionally use a weighting for the data values, and you can configure how the values are combined to produce the output bin values.

This form is somewhat experimental and may be modified in the future.

The shading is done using the shared colour map. This colour map is used by all currently visible **Density**, SkyDensity, Aux and Weighted layers. When at least one such layer is being plotted, the Aux Axis control (Appendix A.4.3.4) is visible in the control panel, which allows you to configure the colour map, range, ramp display etc.

The configuration options are:

#### Weight

The weight value applied to each plotted point. Fill this in with a column name or expression from the table just like for a positional coordinate. The exact way this quantity is used depends on the setting of the **Combine** control below. If it's left blank, the weighting is considered to be unity (all values are 1); this makes sense for some combination types (e.g. sum) but not others (e.g. mean).

#### Bin Size

Determines the dimension of grid bins in pixels. Bins are square in pixel dimensions and this value gives the extent in pixels along each side. Currently, only integer values are allowed.

#### Combine

Determines how the weight values associated with markers plotted covering a given bin are combined to produce the numeric value used for that bin's colour.

The following options (some are more useful than others) are currently available:

- **sum**: the sum of all weights
- **mean**: the mean of all weights
- **median**: the median of all weights; (note, this is expensive to calculate and may be slow)
- **min**: the minimum weight
- **max**: the maximum weight
- **stdev**: the sample standard deviation of all weights
- **count**: the number of points plotted (weight value is ignored, this is like Density mode)
- **hit**: one if any point is plotted, blank otherwise (weight value is ignored, this is like Flat mode)

#### Opaque Limit

Determines transparency of the points. By default, they are fully opaque, but if you slide the


slider to the right, they will become progressively more transparent.

#### A.4.5.14 SkyDensity Form

The screenshot shows the SkyDensity form configuration panel. It is organized into four main sections:

- Coordinates:** Contains a 'Weight' field with the text 'ln(Fnu\_60)' and a dropdown arrow.
- Global Style:** Contains 'HEALPix Level' with radio buttons for 'Abs' (selected) and 'Rel'. The 'Abs' value is '5'. The 'Rel' value is '-2' with a range of '5'. Below this is a 'Combine' dropdown set to 'mean'. At the bottom is a 'Transparency' slider.
- Report:** Contains a 'Subset' dropdown set to 'All' and a 'Tile size/sq.deg' field displaying '3.3571746'.
- HEALPix Map:** Contains two buttons: 'Import' and 'Save', each with a red circular icon.

#### SkyDensity form configuration panel

The **SkyDensity** form (  ) plots a density map on the sky using a HEALPix grid with a configurable resolution. You can optionally use a weighting for the data values to accumulate within each HEALPix tile, and you can configure how the weighted values are combined to generate the eventual pixel values (and hence colours). HEALPix is a tiling scheme for the sky which uses square-ish pixels of equal area to cover the celestial sphere.

The shading is done using the shared colour map. This colour map is used by all currently visible **SkyDensity**, Density, Aux and Weighted layers. When at least one such layer is being plotted, the Aux Axis control (Appendix A.4.3.4) is visible in the control panel, which allows you to configure the colour map, range, ramp display etc.

The configuration options are:

##### Weight

The weight value applied to each plotted point. Fill this in with a column name or expression from the table just like for a positional coordinate. The exact way this quantity is used depends on the setting of the **Combine** control below. If it's left blank, the weighting is considered to be unity (all values are 1); this makes sense for some combination types (e.g. sum) but not others (e.g. mean).

##### HEALPix Level

This allows you to control the resolution of the HEALPix grid onto which the weighted values are resampled. According to the radio buttons, you can configure this using either a **Absolute** or **Relative** value. In Absolute mode you specify the HEALPix level ( $k$ ) directly; the number of pixels on the sky is  $12 \cdot 2^{2k}$ . In Relative mode the level is set so that a HEALPix tile has approximately the given number of screen pixels along a side, and hence the absolute level will change if you zoom in and out. In either case, you can see the absolute level at the right hand side of the control.

### Combine

Determines how the weight values associated with markers plotted covering a given HEALPix tile are combined to produce the numeric value used for that tile's colour.

The following options (some are more useful than others) are currently available:

- **sum**: the sum of all weights
- **mean**: the mean of all weights
- **median**: the median of all weights; (note, this is expensive to calculate and may be slow)
- **min**: the minimum weight
- **max**: the maximum weight
- **stdev**: the sample standard deviation of all weights
- **count**: the number of points plotted (weight value is ignored, this is like Density mode)
- **hit**: one if any point is plotted, blank otherwise (weight value is ignored, this is like Flat mode)

### Opaque Limit


Determines transparency of the points. By default, they are fully opaque, but if you slide the slider to the right, they will become progressively more transparent.

The **Report** panel provides information calculated by the plot:

#### Tile size/sq.deg

Reports the size of each pixel in square degrees. This is simply a function of the actual HEALPix level used for the plot (which is reported at the right of the **HEALPix Level** control).

#### HEALPix Map

This allows you to export the healpix map that you can see in the plot in the form of a table. The table has a row for each plotted healpix pixel, with two columns: one giving the healpix pixel index, and the other giving the pixel value (corresponding to the colour in the plot). The **Import** (  ) option loads the data as a new table in the TOPCAT application, and the **Save** (




) option lets you save it directly to disk in one of the available table formats.

#### A.4.5.15 Fill Form



#### Fill form configuration panel

If a two-dimensional dataset represents a single-valued function, the **Fill** form (  ) will fill the

area underneath the function's curve with a solid colour. Parts of the surface which would only be partially covered (because of rapid function variation within the width of a single pixel) are represented using appropriate alpha-blending. The filled area may alternatively be that above the curve or (in some plot types) to its left or right.

One example of its use is to reconstruct the appearance of a histogram plot from a set of histogram bins. For X,Y data which is not single-valued, the result may not be very useful.

This form may be used in the Plane or Time plot windows.

The configuration options are:

### Transparency

Adjusts the transparency of the filled area.

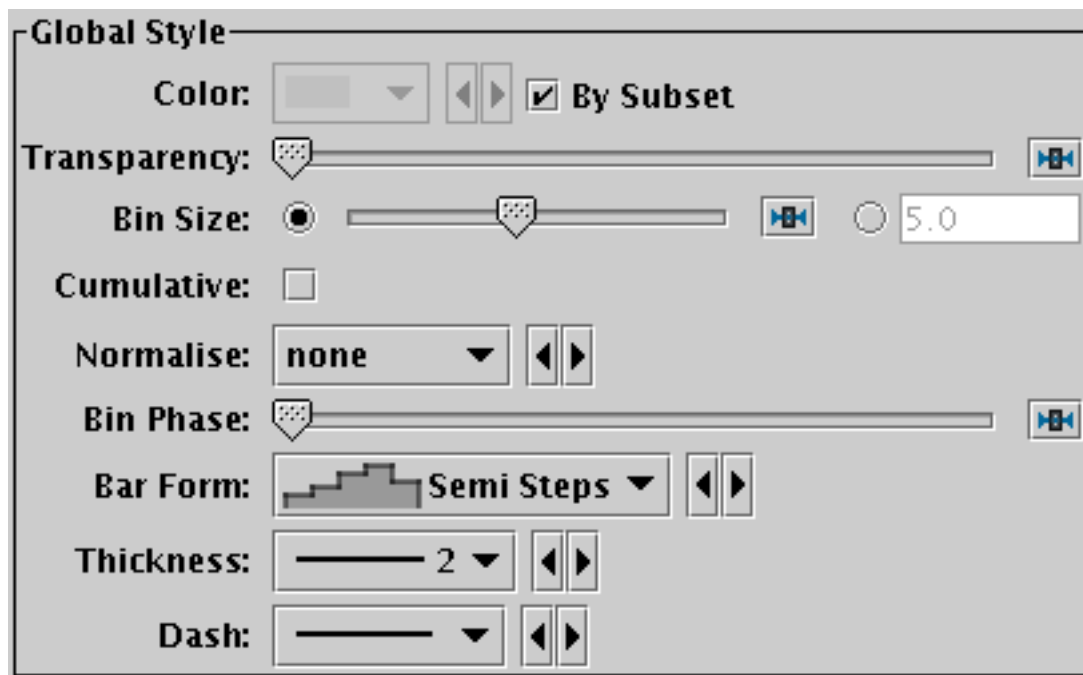
### Horizontal

Determines whether the filling is vertical (suitable for functions of the horizontal variable) or horizontal (suitable for functions of the vertical variable). In the Time plot, the fill is always vertical, so this option is not provided.

### Positive

Determines the directional sense of the filling. If false, the fill is between the data points and negative infinity along the relevant axis (e.g. down from the data points to the bottom of the plot). If true, the fill is in the other direction.


## A.4.5.16 Histogram Form



The image shows a configuration panel for the Histogram form. It is titled 'Global Style' and contains several settings:

- Color:** A color selection box with a dropdown arrow and a 'By Subset' checkbox.
- Transparency:** A slider control with a 'By Subset' checkbox.
- Bin Size:** A slider control with a 'By Subset' checkbox and a numeric input field set to 5.0.
- Cumulative:** A checkbox.
- Normalise:** A dropdown menu set to 'none' with a 'By Subset' checkbox.
- Bin Phase:** A slider control with a 'By Subset' checkbox.
- Bar Form:** A dropdown menu set to 'Semi Steps' with a 'By Subset' checkbox.
- Thickness:** A dropdown menu set to 2 with a 'By Subset' checkbox.
- Dash:** A dropdown menu with a 'By Subset' checkbox.

### Histogram form configuration panel

The **Histogram** form (  ) plots a histogram generated by binning samples along the X axis.

This form may be used in either the Histogram or Plane plot windows.

These options always appear in the form configuration panel:

### Colour

Selects the basic colour for the dataset.

### Transparency

Adjusts the transparency of the bars or line. For bar styles which are already partially transparent, this fades them further.

**Bar Form**


Type of histogram bars: filled, open, steps, spikes etc.

**Thickness**

Controls line thickness where applicable. This is only relevant for bar styles that draw a finite thickness line, so has no effect for solid bars.

**Dash**

Controls the dash pattern of the line (solid, dots, dashes etc). This is only relevant for bar styles that draw a finite thickness line, so has no effect for solid bars.

And these options appear in the form configuration panel for the Plane window, or the Bins control (  ) for the Histogram window:

**Bin Size**

A scale for the width of bins that are shown on the screen. There are two ways to specify this. If the left-hand radio button is selected, the adjacent slider will adjust the bin size, which is also affected by the actual width of the plotting window in pixels. Slide the slider left to get narrower bins or right to get wider ones. If the right-hand radio button is selected, you can enter a numeric value giving the actual width in data units of each bar (for a logarithmic X axis this value is a factor).

**Bin Phase**

Controls where the horizontal zero point for binning is set. For instance if your bin size is 1, it controls whether bins boundaries are at 0, 1, 2, .. or 0.5, 1.5, 2.5, ... etc. If the slider is at either end of the scale, there will be a bin boundary at  $X=0$  (linear X axis) or  $X=1$  (logarithmic X axis).

**Cumulative**

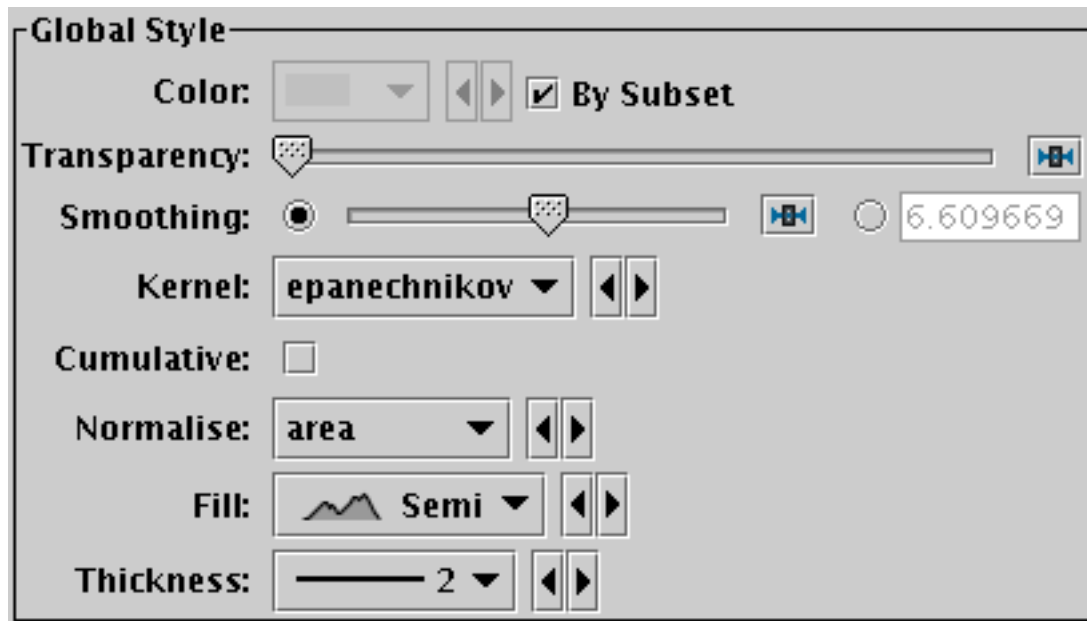
If true, the histogram bars plotted are calculated cumulatively; each bin includes the counts from all previous bins.

**Normalise**


Defines how, if at all, the bars are normalised. The available options are:

- **none**: No normalisation is performed.
- **area**: The total area of histogram bars is normalised to unity. For cumulative plots, this behaves like **height**.
- **maximum**: The height of the tallest histogram bar is normalised to unity. For cumulative plots, this behaves like **height**.
- **height**: The total height of histogram bars is normalised to unity.

**A.4.5.17 KDE Form**



### KDE form configuration panel

The **KDE** form (  ) plots a discrete Kernel Density Estimate giving a smoothed frequency of data values along the horizontal axis, using a fixed-width smoothing kernel. (for a variable-bandwidth kernel, see KNN). This is a generalisation of a histogram in which the bins are always 1 pixel wide, and a smoothing kernel is applied to each bin. The width and shape of the kernel may be varied.

This is suitable for cases where the division into discrete bins done by a normal histogram is unnecessary or troublesome.

Note this is not a true Kernel Density Estimate, since, for performance reasons, the smoothing is applied to the (pixel-width) bins rather than to each data sample. The deviation from a true KDE caused by this quantisation will be at the pixel level, hence in most cases not visually apparent.

This form may be used in either the Histogram or Plane plot windows.

These options always appear in the form configuration panel:

#### Colour

Selects the basic colour for the dataset.

#### Transparency

Adjusts the transparency of the bars or line. For bar styles which are already partially transparent, this fades them further.

#### Fill

Determines how the density function is represented. The options are:


- **solid**: area between level and axis is filled with solid colour
- **line**: level is marked by a wiggly line
- **semi**: level is marked by a wiggly line, and area below it is filled with a transparent colour

#### Thickness

Controls line thickness where applicable. This is only relevant for bar styles that draw a finite thickness line, so has no effect for **solid** filling.

And these options appear in the form configuration panel for the Plane window, or the Bins control



(  ) for the Histogram window:

### Smoothing

Configures the smoothing width for kernel density estimation. This is the characteristic width of the kernel function to be convolved with the density to produce the visible plot.

Sliding the slider to the right makes the kernel width larger. The width in data units is shown in the text field on the right (if the X axis is logarithmic, this is a factor). Alternatively you can click the radio button near the text field, and enter the width in data units directly.

### Kernel

The functional form of the smoothing kernel. The functions listed refer to the unscaled shape; all kernels are normalised to give a total area of unity.

The available options are:

- **square**: Uniform value:  $f(x)=1$ ,  $|x|=0..1$
- **linear**: Triangle:  $f(x)=1-|x|$ ,  $|x|=0..1$
- **epanechnikov**: Parabola:  $f(x)=1-x^2$ ,  $|x|=0..1$
- **cos**: Cosine:  $f(x)=\cos(x\pi/2)$ ,  $|x|=0..1$
- **cos2**: Cosine squared:  $f(x)=\cos^2(x\pi/2)$ ,  $|x|=0..1$
- **gauss3**: Gaussian truncated at 3.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..3$
- **gauss6**: Gaussian truncated at 6.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..6$

### Cumulative

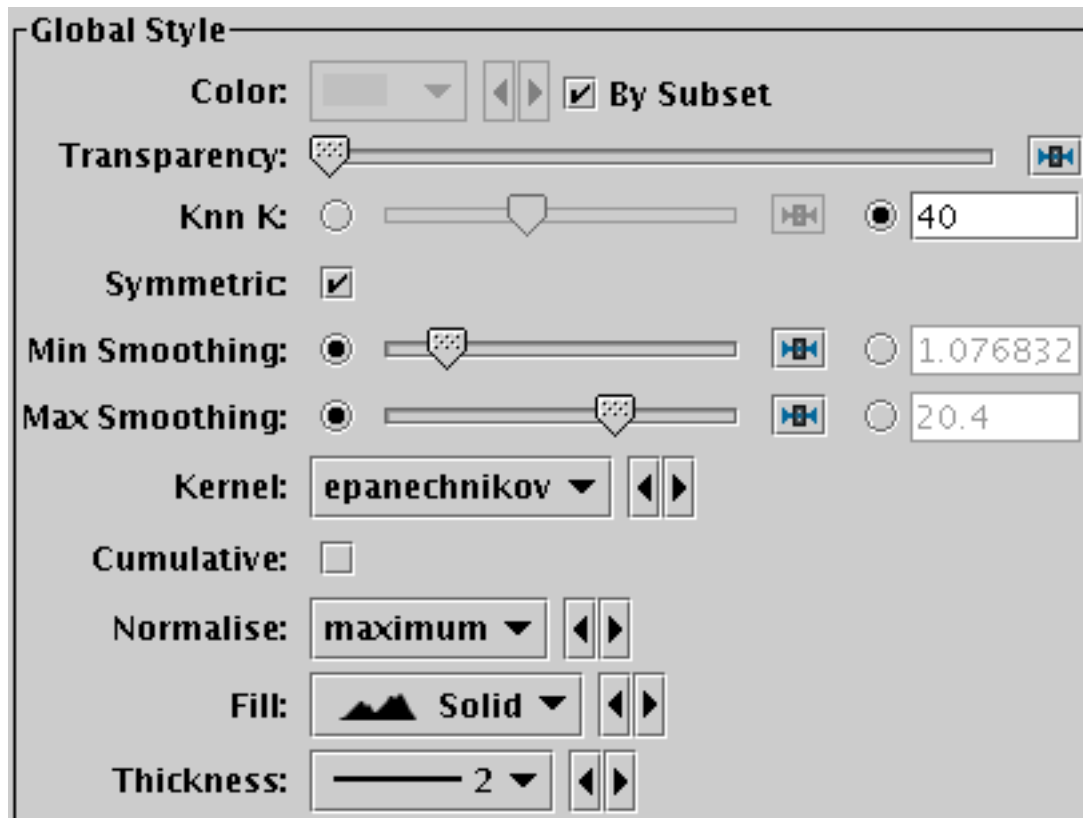
If true, the histogram bars plotted are calculated cumulatively; each bin includes the counts from all previous bins.

### Normalise

Defines how, if at all, the bars are normalised. The available options are:

- **none**: No normalisation is performed.
- **area**: The total area of histogram bars is normalised to unity. For cumulative plots, this behaves like **height**.
- **maximum**: The height of the tallest histogram bar is normalised to unity. For cumulative plots, this behaves like **height**.
- **height**: The total height of histogram bars is normalised to unity.

#### A.4.5.18 KNN Form



**Global Style**

Color:    ☒ By Subset

Transparency:

Knn K:    ☒ 40

Symmetric: ☒

Min Smoothing: ☒    1.076832

Max Smoothing: ☒    20.4

Kernel:


Cumulative: ☐

Normalise:

Fill:

Thickness:

### KNN form configuration panel

The **KNN** form (  ) plots a discrete Kernel Density Estimate giving a smoothed frequency of data values along the horizontal axis, using an adaptive (K-Nearest-Neighbours) smoothing kernel. This is a generalisation of a histogram in which the bins are always 1 pixel wide, and a variable-bandwidth smoothing kernel is applied to each bin (for a fixed-bandwidth kernel, see KDE).

The K-Nearest-Neighbour figure gives the number of points in each direction to determine the width of the smoothing kernel for smoothing each bin. Upper and lower limits for the kernel width are also supplied; if the upper and lower limits are equal, this is equivalent to a fixed-width kernel.

Note this is not a true Kernel Density Estimate, since, for performance reasons, the smoothing is applied to the (pixel-width) bins rather than to each data sample. The deviation from a true KDE caused by this quantisation will be at the pixel level, hence in most cases not visually apparent.

This form may be used in either the Histogram or Plane plot windows.

These options always appear in the form configuration panel:

#### Colour

Selects the basic colour for the dataset.

#### Transparency

Adjusts the transparency of the bars or line. For bar styles which are already partially transparent, this fades them further.

#### Knn K

Sets the number of nearest neighbours to count away from a sample point to determine the width of the smoothing kernel at that point. For the symmetric case this is the total of nearest neighbours summed over both directions, and for the asymmetric case it is the number in a single direction. The threshold is actually the weighted total of samples; for unweighted

(weight=1) bins that is equivalent to the number of samples.

### Symmetric

If checked, the nearest neighbour search is carried out in both directions, and the kernel is symmetric. If unchecked, the nearest neighbour search is carried out separately in the positive and negative directions, and the kernel width is accordingly different in the positive and negative directions.

### Min Smoothing

Fixes the minimum size of the smoothing kernel. This functions as a lower limit on the distance that is otherwise determined by searching for the K nearest neighbours at each sample point.

You can either use the slider, or check the radio button on the right and enter the value in data units directly.

### Max Smoothing

Fixes the maximum size of the smoothing kernel. This functions as an upper limit on the distance that is otherwise determined by searching for the K nearest neighbours at each sample point.

You can either use the slider, or check the radio button on the right and enter the value in data units directly.


### Fill

Determines how the density function is represented. The options are:

- **solid:** area between level and axis is filled with solid colour
- **line:** level is marked by a wiggly line
- **semi:** level is marked by a wiggly line, and area below it is filled with a transparent colour

### Thickness

Controls line thickness where applicable. This is only relevant for bar styles that draw a finite thickness line, so has no effect for **solid** filling.

And these options appear in the form configuration panel for the Plane window, or the Bins control (  ) for the Histogram window:

### Kernel

The functional form of the smoothing kernel. The functions listed refer to the unscaled shape; all kernels are normalised to give a total area of unity.

The available options are:

- **square:** Uniform value:  $f(x)=1$ ,  $|x|=0..1$
- **linear:** Triangle:  $f(x)=1-|x|$ ,  $|x|=0..1$
- **epanechnikov:** Parabola:  $f(x)=1-x^2$ ,  $|x|=0..1$
- **cos:** Cosine:  $f(x)=\cos(x\pi/2)$ ,  $|x|=0..1$
- **cos2:** Cosine squared:  $f(x)=\cos^2(x\pi/2)$ ,  $|x|=0..1$
- **gauss3:** Gaussian truncated at 3.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..3$
- **gauss6:** Gaussian truncated at 6.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..6$

### Cumulative

If true, the histogram bars plotted are calculated cumulatively; each bin includes the counts from all previous bins.

### Normalise

Defines how, if at all, the bars are normalised. The available options are:

- **none:** No normalisation is performed.


- **area**: The total area of histogram bars is normalised to unity. For cumulative plots, this behaves like **height**.
- **maximum**: The height of the tallest histogram bar is normalised to unity. For cumulative plots, this behaves like **height**.
- **height**: The total height of histogram bars is normalised to unity.

#### A.4.5.19 Densogram Form

The image shows a configuration panel titled "Global Style" for a Densogram. It contains the following settings:

- Color**: A color selection box with a dropdown arrow and a "By Subset" checkbox (checked).
- Smoothing**: A radio button, a slider with a shield icon, a zoom icon, and a text box containing "0.91877".
- Kernel**: A dropdown menu showing "epanechnikov" with a zoom icon.
- Density Shader**: A color gradient bar labeled "Heat" with a dropdown arrow and a zoom icon.
- Shader Clip**: A slider with shield icons at both ends and a "Default" checkbox.
- Shader Flip**: A checked checkbox.
- Shader Quantise**: A radio button, a slider with a shield icon, a zoom icon, and an empty text box.
- Scaling**: A dropdown menu showing "linear" with a zoom icon.
- Density Subrange**: A slider with shield icons at both ends.
- Cumulative**: An unchecked checkbox.
- size**: A slider with a shield icon and a zoom icon.
- pos**: A slider with a shield icon and a zoom icon.

#### Densogram form configuration panel

The **Densogram** form (  ) represents smoothed density of data values along the horizontal axis

using a colourmap. This is like a Kernel Density Estimate (smoothed histogram with bins 1 pixel wide), but instead of representing the data extent vertically as bars or a line, values are represented by a fixed-size pixel-width column of a colour from a colour map. A smoothing kernel, whose width and shape may be varied, is applied to each data point.

This is a rather unconventional way to represent density data, and this plotting mode is probably not very useful. But hey, nobody's forcing you to use it.

This form may be used in either the Histogram or Plane plot windows.

These options always appear in the form configuration panel:

##### Colour

The base colour for the data set. Only certain shader colour maps respect this value, for others it is ignored.

##### Density Shader

The colour map for displaying density values. There are two types, relative and absolute. Relative maps have names marked by a star ("\*"), and alter the basic dataset colour, for

instance by darkening or lightening it, while absolute maps (the rest) ignore the basic dataset colour altogether. For a single-dataset plot, the absolute maps are best, but for multiple subsets it may be less confusing to use a relative one.

### Shader Clip

Select a sub-range of the full colour map above. If the **Default** checkbox is checked, then all or most of the colour ramp from the **Shader** control is used. If you want to configure the range of colours from the map yourself, uncheck the Default checkbox, and slide the handles in from the end of the slider to choose exactly the range you want.

The default range is clipped at one end for colour maps that fade to white, so that all the plotted colours will be distinguishable against a white background. If you don't want that, you can uncheck Default and leave the handles at the extreme ends of the slider.

### Shader Flip

Whether the density scale should map forwards or backwards into the colour map.

### Shader Quantise

Allows the colour map to be quantised. By default, the colour map is effectively continuous. If you slide the slider to the right, or enter a value in the text field, the map will be split into a decreasing number of discrete colours. This can be used to generate a contour-like effect, and may make it easier to trace the boundaries of regions of interest by eye.

### Scaling

Determines the function used to map the range of density values onto the colour map. Options are **linear**, **logarithmic**, **square** and **square root**.

### Density Subrange


Adjusts the density range over which the colour map is applied. By default the colour map is scaled using limits found from the data density in the plot (the most dense few pixels are ignored), but you can restrict the range using this slider.

### Size

Height of the density bar in pixels.

### Position

Determines where on the plot region the density bar appears. The value should be in the range 0..1; zero corresponds to the bottom of the plot and one to the top.

And these options appear in the form configuration panel for the Plane window, or the Bins control (  ) for the Histogram window:

### Smoothing

Configures the smoothing width for kernel density estimation. This is the characteristic width of the kernel function to be convolved with the density to produce the visible plot.

Sliding the slider to the right makes the kernel width larger. The width in data units is shown in the text field on the right (if the X axis is logarithmic, this is a factor). Alternatively you can click the radio button near the text field, and enter the width in data units directly.

### Kernel

The functional form of the smoothing kernel. The functions listed refer to the unscaled shape; all kernels are normalised to give a total area of unity.

The available options are:

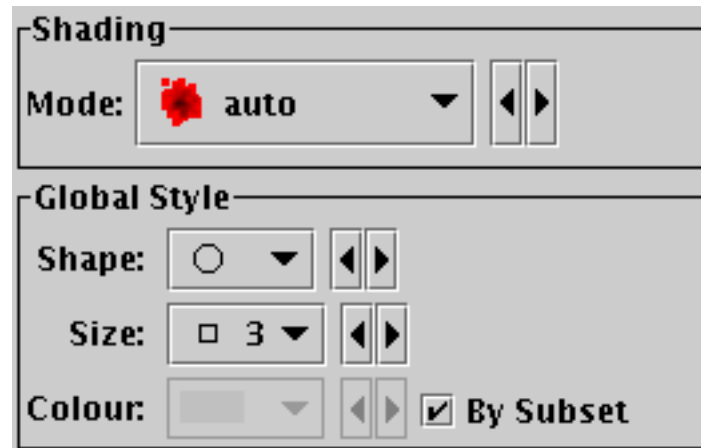
- **square**: Uniform value:  $f(x)=1$ ,  $|x|=0..1$
- **linear**: Triangle:  $f(x)=1-|x|$ ,  $|x|=0..1$
- **epanechnikov**: Parabola:  $f(x)=1-x^2$ ,  $|x|=0..1$
- **cos**: Cosine:  $f(x)=\cos(x*\pi/2)$ ,  $|x|=0..1$
- **cos2**: Cosine squared:  $f(x)=\cos^2(x*\pi/2)$ ,  $|x|=0..1$

- **gauss3**: Gaussian truncated at 3.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..3$
- **gauss6**: Gaussian truncated at 6.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..6$

### Cumulative


If true, the histogram bars plotted are calculated cumulatively; each bin includes the counts from all previous bins.

#### A.4.5.20 Pair Marks Form



The configuration panel for the Pair Marks form is divided into two sections: **Shading** and **Global Style**. The **Shading** section contains a **Mode** dropdown menu set to **auto**, accompanied by a red flower-shaped icon and navigation arrows. The **Global Style** section contains three controls: **Shape** (a circle icon), **Size** (a square icon with the value 3), and **Colour** (a grey square). Each of these three controls has navigation arrows. Additionally, there is a checkbox labeled **By Subset** which is checked.

#### Pair Mark form configuration panel

The **Pair Mark** form (  ) is available from the pair position layer control, and plots a marker of configurable shape and size at both points in a position pair. The same marker is used for both ends; if you want different points you can use two single Mark forms.

Configuration options are:

#### Shading Mode

See Appendix A.4.6.

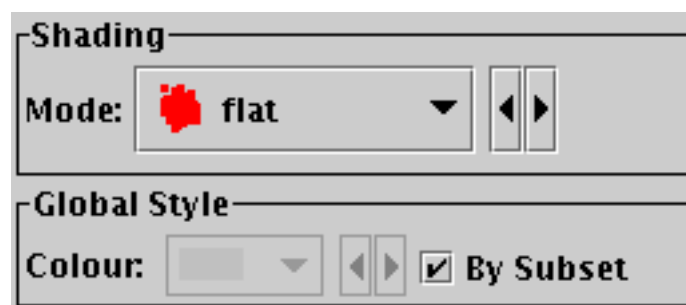
#### Shape

Marker shape from a list of options.

#### Size


Marker size in pixels.

#### A.4.5.21 Pair Link Form



The configuration panel for the Pair Link form is divided into two sections: **Shading** and **Global Style**. The **Shading** section contains a **Mode** dropdown menu set to **flat**, accompanied by a red flower-shaped icon and navigation arrows. The **Global Style** section contains a **Colour** dropdown menu (showing a grey square) and a checked checkbox labeled **By Subset**. Navigation arrows are present for the Colour dropdown.

#### Pair Link form configuration panel

The **Pair Link** form (  ) is available from the pair position layer control, and plots a line linking

the two points in a position pair.

Configuration options are:

### Shading Mode

See Appendix A.4.6.

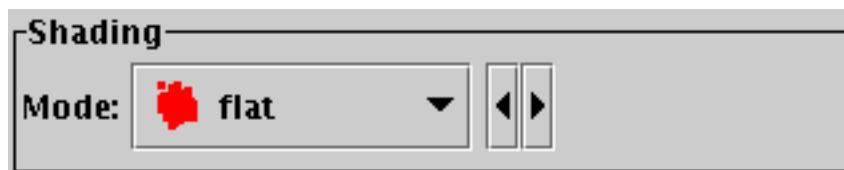
## A.4.6 Shading Modes

Most of the Plot Forms (Appendix A.4.5) have a style colour associated with them for each data set. This defines the basic colour used to plot the shape at each data point. However, many of the forms also ask you to select a **Shading Mode**, which determines the actual colour displayed in the plot for the plotted points. The shaded colour is based on the selected style colour, but may also be influenced by the number of points plotted there, some extra data coordinate, or other configuration information.


When exporting plots to an external vector graphics format (EPS or PDF), some of the shading modes may not behave the same as in a bitmap (on the screen, or to a bitmapped format such as GIF or PNG). Any such anomalies are noted in the in the mode descriptions below.

The different mode shading mode options are described in the following subsections.

### A.4.6.1 Flat Mode

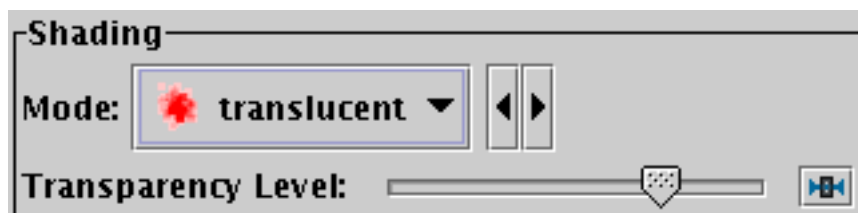


#### Flat shading mode selection


The **Flat** shading mode (  ) simply colours points in the colour selected by their style. It has no additional parameters or coordinates.

**Exporting:** This mode works without problem for both bitmapped and vector output.

### A.4.6.2 Translucent Mode



#### Translucent shading mode selection


The **Translucent** shading mode (  ) colours shapes in a transparent version of the colour selected by their style. The degree of transparency is determined by how many points are currently being plotted on top of each other, and by the **Transparency Level** slider; as you slide further to the right, points get more transparent. Unlike transparent mode, the transparency varies according to the current point density, so you can usually leave the setting the same as you zoom in and out.

**Exporting:** When the points are opaque, this mode works without problem for both bitmapped and vector output, but when the transparency is set there may be anomalies. Transparent points are rendered in PDF output, though the transparency levels may not be exactly the same as on the screen. This can be fixed by using the **Force Bitmap** option in the Plot Export (Appendix A.4.2.4) dialogue. For PostScript, transparent points are rendered as opaque. You can use **Force Bitmap** with PostScript which will get transparency right for this layer, but then any earlier layers will be completely obscured.

#### A.4.6.3 Transparent Mode

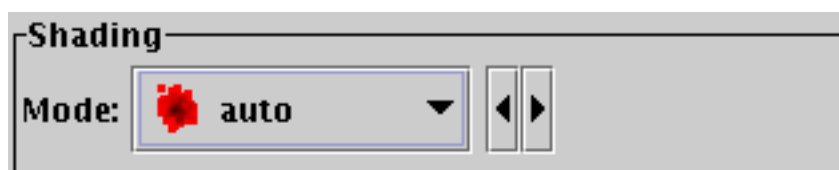


##### Transparent shading mode selection


The **Transparent** shading mode (  ) colours shapes in a transparent version of the colour selected by their style. The degree of transparency is determined by the **Opaque Limit** slider - at the left end, points are fully opaque and this is equivalent to Flat mode, and as you slide further to the right, the points get more transparent. The higher the opaque limit, the more points have to be plotted on top of each other to reach colour that fully obscures the background. Unlike translucent mode, transparency of each colour is fixed by the opaque limit, rather than adjusting depending on the density of points currently plotted.

**Exporting:** When the points are opaque, this mode works without problem for both bitmapped and vector output, but when the transparency is set there may be anomalies. Transparent points are rendered in PDF output, though the transparency levels may not be exactly the same as on the screen. This can be fixed by using the **Force Bitmap** option in the Plot Export (Appendix A.4.2.4) dialogue. For PostScript, transparent points are rendered as opaque. You can use **Force Bitmap** with PostScript which will get transparency right for this layer, but then any earlier layers will be completely obscured.

#### A.4.6.4 Auto Mode



##### Auto shading mode selection

The **Auto** shading mode (  ) colours isolated points in their selected colour, but where multiple points *from the same data set* overlap it adjusts the colour by darkening it. This means for that isolated points (most or all points in a non-crowded plot, or outliers in a crowded plot) it behaves just like Flat mode, but it's easy to see where overdense regions lie.

This is like Density mode, but with no user-configurable options.

This is the default mode for 2d plots, since it gives you a good first idea of what the data is doing. For 3d plots it can be used, and it works well for single dataset plots, but in the case of multiple

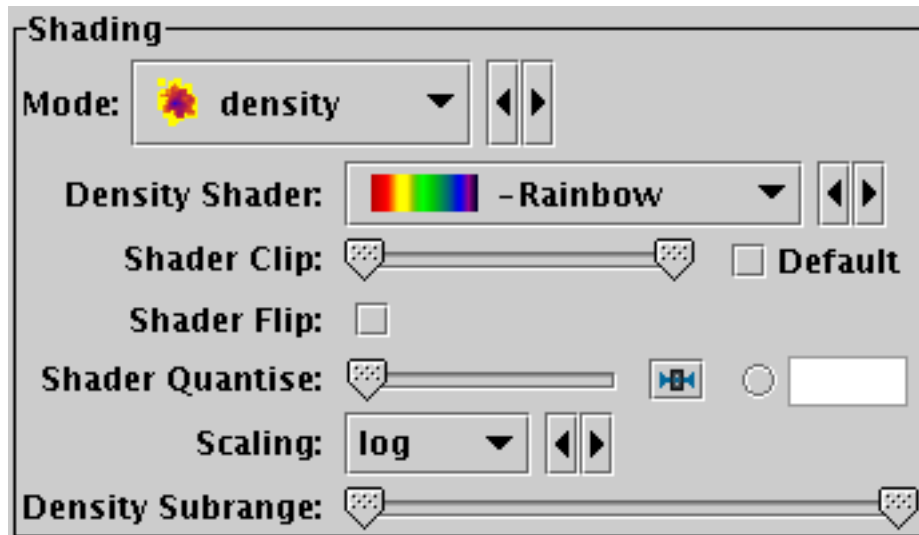


datasets it can be misleading since the coloured pixels can't be placed sensibly in the 3d space.


The colour darkening is based on the asinh function; the intention is that two points overlaid should be just enough different in colour for the difference to be visible, and the mapping is scaled so that if there are very dense regions they will come out nearly black.

**Exporting:** When exported to vector formats, the output is automatically forced to a bitmap for Auto-mode layers. In the case of PostScript, this completely obscures any previous layers.

#### A.4.6.5 Density Mode



#### Density mode selection

The **Density** shading mode (  ) uses a configurable colour map to indicate how many points are plotted over each other. Specifically, it colours each pixel according to how many times that pixel has been covered by one of the shapes plotted by the layer in question. To put it another way, it generates a false-colour density map with pixel granularity using a smoothing kernel of the form of the shapes plotted by the layer. The upshot is that you can see the plot density of points or other shapes plotted.

This is like Auto mode, but with more user-configurable options. The options are:

##### Density Shader

The colour map for displaying density values. There are two types, relative and absolute. Relative maps have names marked by a star ("\*"), and alter the basic dataset colour, for instance by darkening or lightening it, while absolute maps (the rest) ignore the basic dataset colour altogether. For a single-dataset plot, the absolute maps are best, but for multiple subsets it may be less confusing to use a relative one.

##### Shader Clip

Select a sub-range of the full colour map above. If the **Default** checkbox is checked, then all or most of the colour ramp from the **Shader** control is used. If you want to configure the range of colours from the map yourself, uncheck the Default checkbox, and slide the handles in from the end of the slider to choose exactly the range you want.

The default range is clipped at one end for colour maps that fade to white, so that all the plotted colours will be distinguishable against a white background. If you don't want that, you can uncheck Default and leave the handles at the extreme ends of the slider.

##### Shader Flip

Whether the density scale should map forwards or backwards into the colour map.

**Shader Quantise**

Allows the colour map to be quantised. By default, the colour map is effectively continuous. If you slide the slider to the right, or enter a value in the text field, the map will be split into a decreasing number of discrete colours. This can be used to generate a contour-like effect, and may make it easier to trace the boundaries of regions of interest by eye.

**Scaling**

Determines the function used to map the range of density values onto the colour map. Options are **linear**, **logarithmic**, **square** and **square root**.


**Density Subrange**

Adjusts the density range over which the colour map is applied. By default the colour map is scaled using limits found from the data density in the plot (the most dense few pixels are ignored), but you can restrict the range using this slider.

Although these options give you quite some control over how densities are mapped to colours, this mode does not display the colour mapping in a way that shows you quantitatively which colours correspond to which numeric density values. If you want that kind of visual feedback, you should use the Weighted shading mode, which can be configured to display point densities (as well as other quantities), and also causes a colour ramp to be displayed under control of the Aux Axis control.

**Exporting:** When exported to vector formats, the output is automatically forced to a bitmap for Density-mode layers. In the case of PostScript, this completely obscures any previous layers.

**A.4.6.6 Aux Mode****Aux mode selection**

The **Aux** shading mode (  ) colours each point according to the value of an additional data coordinate. The point colours then represent an additional dimension of the plot. There is an additional option to draw the points with a fixed transparency.

The shading is done using the shared colour map. This colour map is used by all currently visible **Aux**, **Weighted**, **Density** and **SkyDensity** layers. When at least one such layer is being plotted, the Aux Axis control (Appendix A.4.3.4) is visible in the control panel, which allows you to configure the colour map, range, ramp display etc.

The options are:

**Aux**

The auxiliary coordinate data values. Fill this in with a column name or expression from the table just like for a positional coordinate.

**Opaque Limit**


Determines transparency of the points. By default, they are fully opaque, but if you slide the slider to the right, they will become progressively more transparent.

**Exporting:** Transparent points are rendered in PDF output, though the transparency levels may not be exactly the same as on the screen. This can be fixed by using the **Force Bitmap** option in the Plot Export (Appendix A.4.2.4) dialogue. For PostScript, transparent points are rendered as opaque. You can use **Force Bitmap** with PostScript which will get transparency right for this layer, but then any earlier layers will be completely obscured.

#### A.4.6.7 Weighted Mode



##### Weighted mode selection

The **Weighted** shading mode (  ) paints markers using colours indicating density at each pixel like the Density mode, but with an optional weighting coordinate. You can configure how the weighted coordinates are combined at each pixel to give the final weighted result. Depending on the configuration and actual point density, this can behave in some ways like Aux mode and in some ways like Density mode, but allows you to do things that are not possible in either.

The shading is done using the shared colour map. This colour map is used by all currently visible **Weighted**, Aux, Density and SkyDensity layers. When at least one such layer is being plotted, the Aux Axis control (Appendix A.4.3.4) is visible in the control panel, which allows you to configure the colour map, range, ramp display etc.

The options are:

##### Weight

The weight value applied to each plotted point. Fill this in with a column name or expression from the table just like for a positional coordinate. The exact way this quantity is used depends on the setting of the **Combine** control below. If it's left blank, the weighting is considered to be unity (all values are 1); this makes sense for some combination types (e.g. sum) but not others (e.g. mean).

##### Combine

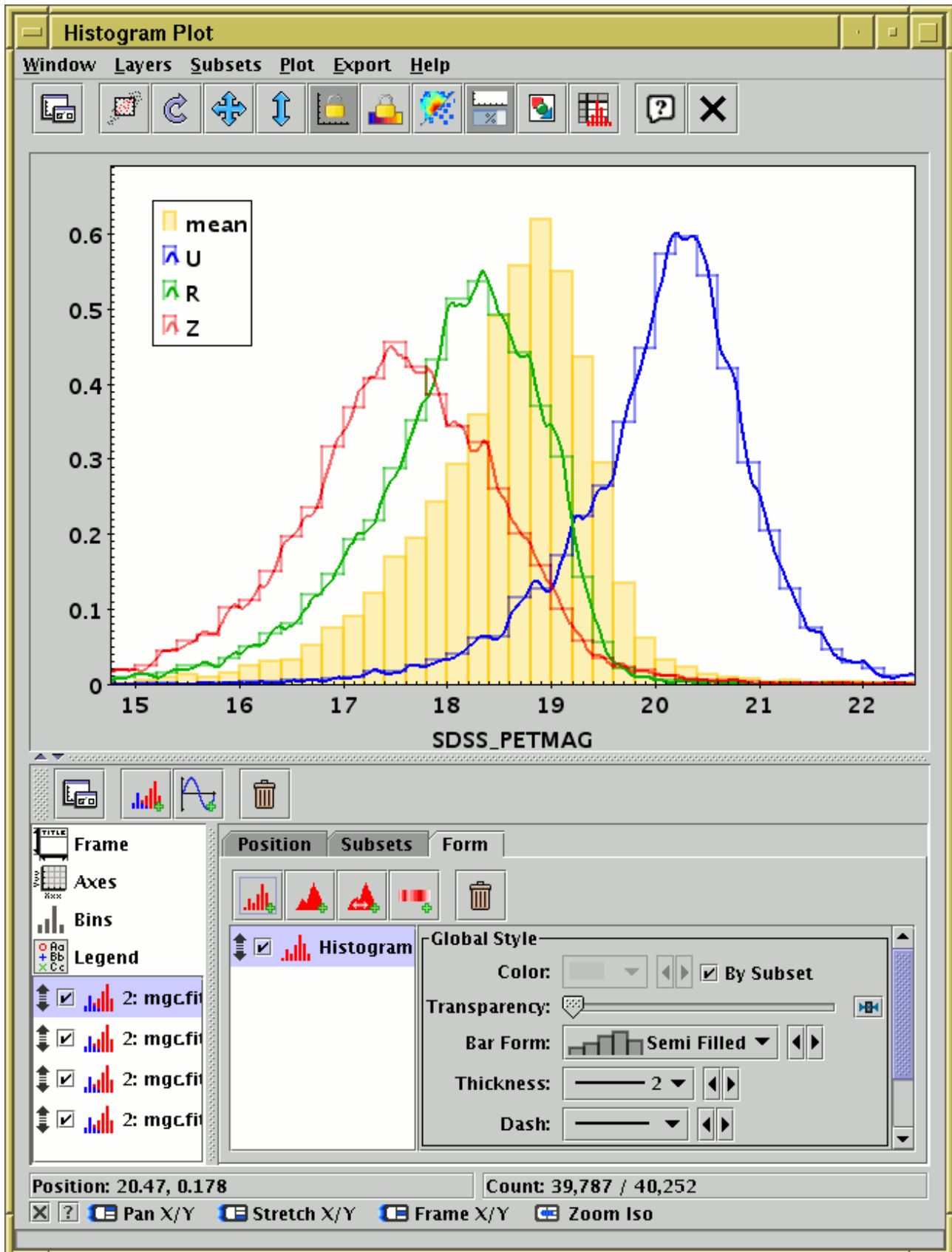
Determines how the weight values associated with markers plotted covering a given screen pixel are combined to produce the numeric value used for that pixel's colour.

The following options (some are more useful than others) are currently available:


- **sum:** the sum of all weights
- **mean:** the mean of all weights
- **median:** the median of all weights; (note, this is expensive to calculate and may be slow)
- **min:** the minimum weight
- **max:** the maximum weight
- **stdev:** the sample standard deviation of all weights
- **count:** the number of points plotted (weight value is ignored, this is like Density mode)
- **hit:** one if any point is plotted, blank otherwise (weight value is ignored, this is like Flat mode)

**Exporting:** When exported to vector formats, the output is automatically forced to a bitmap for Density-mode layers. In the case of PostScript, this completely obscures any previous layers.

#### A.4.7 Histogram Plot Window



## Histogram Plot Window

The **Histogram Plot** () plots 1-dimensional histograms and some variants on the idea of a 1-dimensional Kernel Density Estimate. In many respects it works like the Plane Plot, but it has a restricted set of plot types and an additional fixed control Bins, and the scrolling works a bit differently.

See the Window Overview (Appendix A.4.2) for features common to all plotting windows.

The following non-standard action is available from the toolbar as well as the **Plot** menu:



### Rescale Y

Adjusts the vertical range of the visible data region to accommodate all the histogram bars in the currently visible horizontal region of the plot. The horizontal range is not changed.

In addition, the histogram window lets you export the binned data as a new table, either saving it or loading it directly into TOPCAT's table list. The following actions are available in the **Export** menu; note they only apply to histograms proper, not to KDEs:



### Save as Table







The bin counts/sums corresponding to the currently plotted histogram will be written to disk in tabular form. The first two columns give the lower and upper bounds of each bin, and the subsequent columns give the occupancies of each bin for each plotted data set. If only one dataset is plotted, there will only be three columns.



### Import as Table

Assembles a table as per the **Save** option above, but rather than writing it to disk imports it directly into TOPCAT, where it can be manipulated in all the usual ways.

The Histogram Plot offers the following plot controls:

-  Histogram Layer Control, with these form options:
  -  Histogram
  -  KDE
  -  KNN
  -  Densogram
-  Function Layer Control

The following subsections describe navigation, axis configuration and bin configuration.

### A.4.7.1 Histogram Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

The navigation actions for this window are:

#### **Left drag**

*Pan.* On the body of the plot this moves it around up/down/left/right. To move it only vertically, drag on the left of the Y axis. To move it only horizontally, drag below the X axis. You can configure dragging on the body of the plot to be only vertical or horizontal by setting the **Pan/Zoom Axes** in the **Navigation** axis configuration tab.

#### **Right drag (CTRL-drag)**

*Stretch zoom.* On the body of the plot, dragging up/down stretches/squashes the plot vertically, and dragging left/right stretches/squashes it horizontally. To zoom only vertically, drag on the left of the Y axis. To zoom only horizontally, drag below the X axis.

Zooming horizontally will normally adjust the width of the histogram bars appropriately.

Normally, the zoom will be centered horizontally at the mouse position and vertically on the X axis, so the zoom will not move the bottom of the histogram. You can adjust that with the **Anchor X/Y Axis** options in the **Navigation** axis configuration tab.

#### **Wheel**


*Isotropic zoom.* Spinning the mouse wheel forwards/backwards will zoom in/out just like dragging with the right button up-and-right or down-and-left.

You can also manually fix the plot bounds using the **Range** tab of the Axes control.

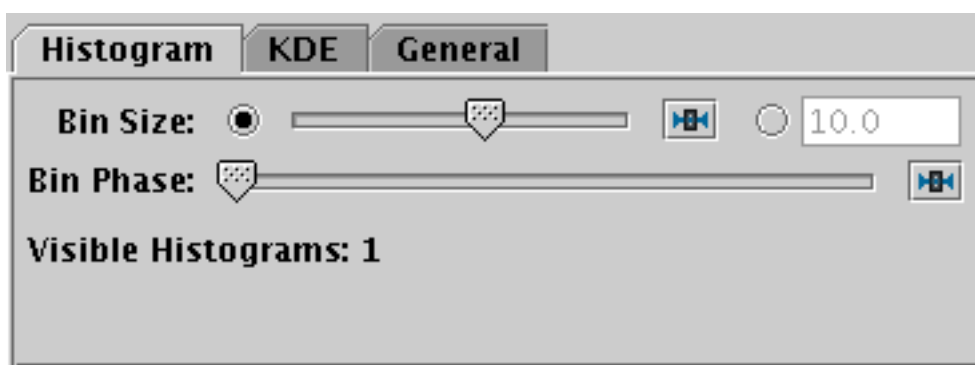
### A.4.7.2 Histogram Axes Control

The Histogram Plot axis control is very similar to the Plane Plot axis control described in Appendix A.4.8.2. The only difference is in the **Navigation** tab. For the histogram, the **Anchor X Axis** option is set **on** by default, since you will usually want the Y=0 line to stay anchored to the bottom of the plot when zooming.

### A.4.7.3 Bins Control

The **Bins** control (  ) is found in the control stack of the Histogram window. It configures the common placement and calculation of some options for all the histogram-like layers displayed. Being able to set these values in common for all displayed layers of a similar type can be convenient, but if you need to use different parameters for different datasets, you can plot the same layer forms in the Plane window (which has no fixed Bins control) instead.

There are three tabs: **Histogram**, **KDE** and **General**.



#### **Histogram tab of histogram window Bins fixed control**

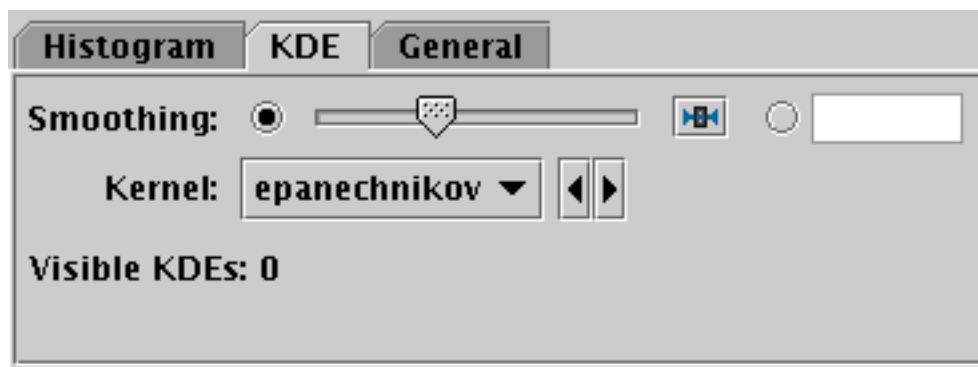
The **Histogram** tab affects all Histogram layers, and has the following controls:

### Bin Size

A scale for the width of bins that are shown on the screen. There are two ways to specify this. If the left-hand radio button is selected, the adjacent slider will adjust the bin size, which is also affected by the actual width of the plotting window in pixels. Slide the slider left to get narrower bins or right to get wider ones. If the right-hand radio button is selected, you can enter a numeric value giving the actual width in data units of each bar (for a logarithmic X axis this value is a factor).

### Bin Phase

Controls where the horizontal zero point for binning is set. For instance if your bin size is 1, it controls whether bins boundaries are at 0, 1, 2, .. or 0.5, 1.5, 2.5, ... etc. If the slider is at either end of the scale, there will be a bin boundary at  $X=0$  (linear X axis) or  $X=1$  (logarithmic X axis).



### KDE tab of histogram window Bins fixed control

The **KDE** (Kernel Density Estimate) tab affects all KDE, KNN and Densogram layers, and has the following controls:

#### Smoothing

Configures the smoothing width for kernel density estimation. This is the characteristic width of the kernel function to be convolved with the density to produce the visible plot.

Sliding the slider to the right makes the kernel width larger. The width in data units is shown in the text field on the right (if the X axis is logarithmic, this is a factor). Alternatively you can click the radio button near the text field, and enter the width in data units directly.

Note this affects KDE and Densogram layers, but *not* KNN layers, which have their own smoothing controls.

#### Kernel

The functional form of the smoothing kernel. The functions listed refer to the unscaled shape; all kernels are normalised to give a total area of unity.

The available options are:

- **square**: Uniform value:  $f(x)=1$ ,  $|x|=0..1$
- **linear**: Triangle:  $f(x)=1-|x|$ ,  $|x|=0..1$
- **epanechnikov**: Parabola:  $f(x)=1-x^2$ ,  $|x|=0..1$
- **cos**: Cosine:  $f(x)=\cos(x*\pi/2)$ ,  $|x|=0..1$
- **cos2**: Cosine squared:  $f(x)=\cos^2(x*\pi/2)$ ,  $|x|=0..1$
- **gauss3**: Gaussian truncated at 3.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..3$
- **gauss6**: Gaussian truncated at 6.0 sigma:  $f(x)=\exp(-x^2/2)$ ,  $|x|=0..6$



### General tab of histogram window Bins fixed control

The **General** tab affects all histogram-like layers, and has the following controls:

#### Cumulative

If true, the histogram bars plotted are calculated cumulatively; each bin includes the counts from all previous bins.

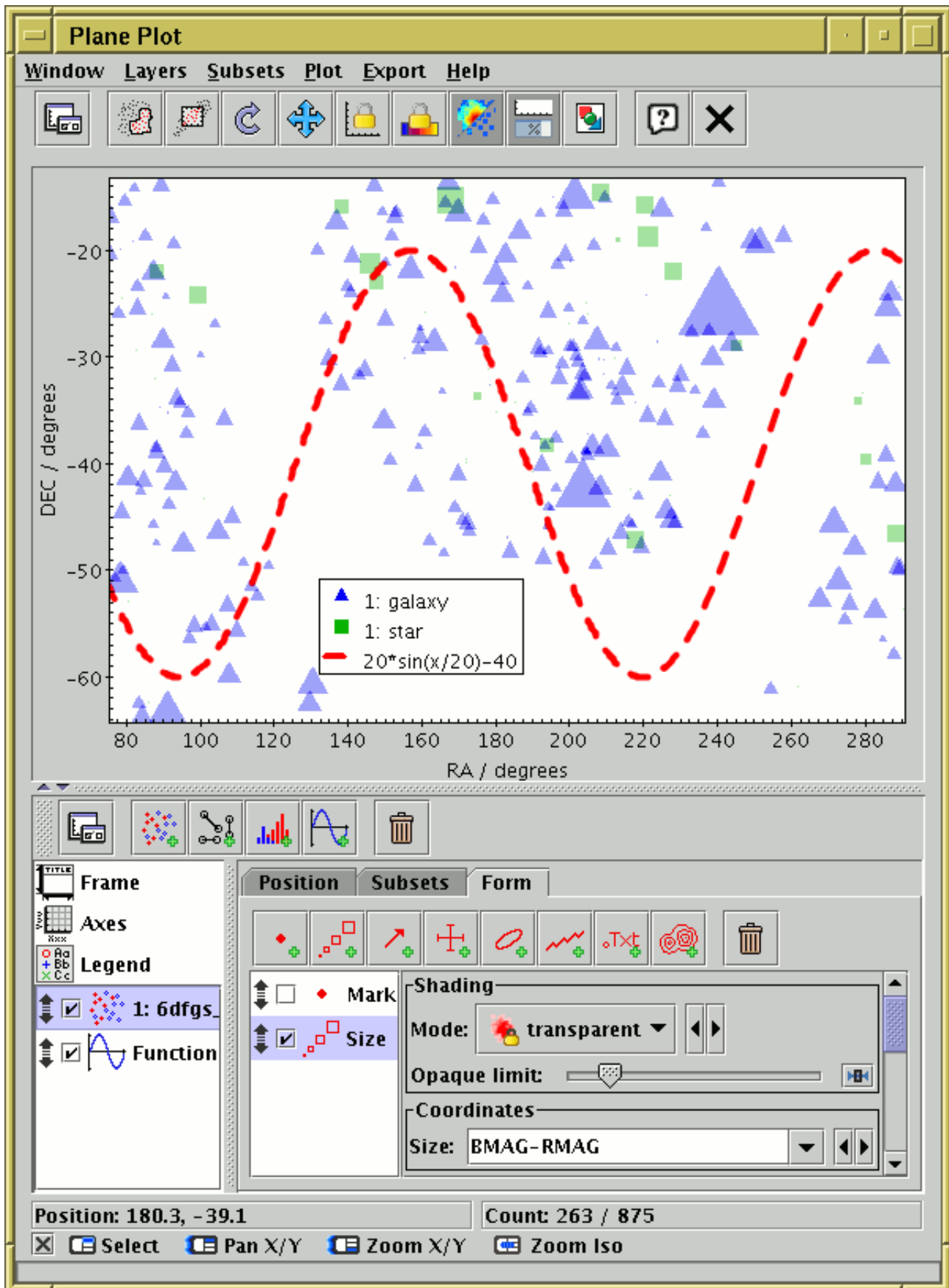
#### Normalise

Defines how, if at all, the bars are normalised. The available options are:


- **none**: No normalisation is performed.
- **area**: The total area of histogram bars is normalised to unity. For cumulative plots, this behaves like **height**.
- **maximum**: The height of the tallest histogram bar is normalised to unity. For cumulative plots, this behaves like **height**.
- **height**: The total height of histogram bars is normalised to unity.

### A.4.8 Plane Plot Window























**Plane Plot Window**

The **Plane Plot** () plots 2-dimensional Cartesian positions on a plane. The positional coordinates are **X** and **Y**. To control the direction and linear/log scaling of the axes, see the **Coords**

tab of the Axes control.

The Plane Plot offers the following plot controls:

-  Position Layer Control, with these form options:
  -  Mark
  -  Size
  -  SizeXY
  -  Vector
  -  Error Bars
  -  XYEllipse
  -  Line
  -  Linear Fit
  -  Text Label
  -  Contour
  -  Density
  -  Fill
-  Pair Layer Control, with the following plot controls:
  -  Marks
  -  Link
-  Histogram Layer Control
-  Function Layer Control

See the Window Overview (Appendix A.4.2) for features common to all plotting windows. The following subsections describe navigation and axis configuration.

#### A.4.8.1 Plane Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

The navigation actions for this window are:

##### Left drag

*Pan.* On the body of the plot this moves it around up/down/left/right. To move it only vertically, drag on the left of the Y axis. To move it only horizontally, drag below the X axis.

You can configure dragging on the body of the plot to be only vertical or horizontal by setting the **Pan/Zoom Axes** in the **Navigation** axis configuration tab.

#### **Right drag (CTRL-drag)**

*Stretch zoom.* On the body of the plot, dragging up/down stretches/squashes the plot vertically, and dragging left/right stretches/squashes it horizontally. Zooming is around the mouse position at the start of the drag. To zoom only vertically, drag on the left of the Y axis. To zoom only horizontally, drag below the X axis.

#### **Center drag (SHIFT-drag)**

*Frame zoom.* On the body of the plot, dragging right-and-down or right-and-up drags out a frame; when the button is released, the plot will be zoomed in to cover the area enclosed by the frame. Dragging left (and up or down) does something like the opposite, you can zoom out using a similar (though not quite the same) mechanism. To zoom in/out only horizontally, drag right/left below the X axis. To zoom in/out only vertically, drag up/down on the left of the Y axis.

#### **Wheel**


*Isotropic zoom.* Spinning the mouse wheel forwards/backwards will zoom in/out around the mouse position, just like dragging with the right button up-and-right or down-and-left.

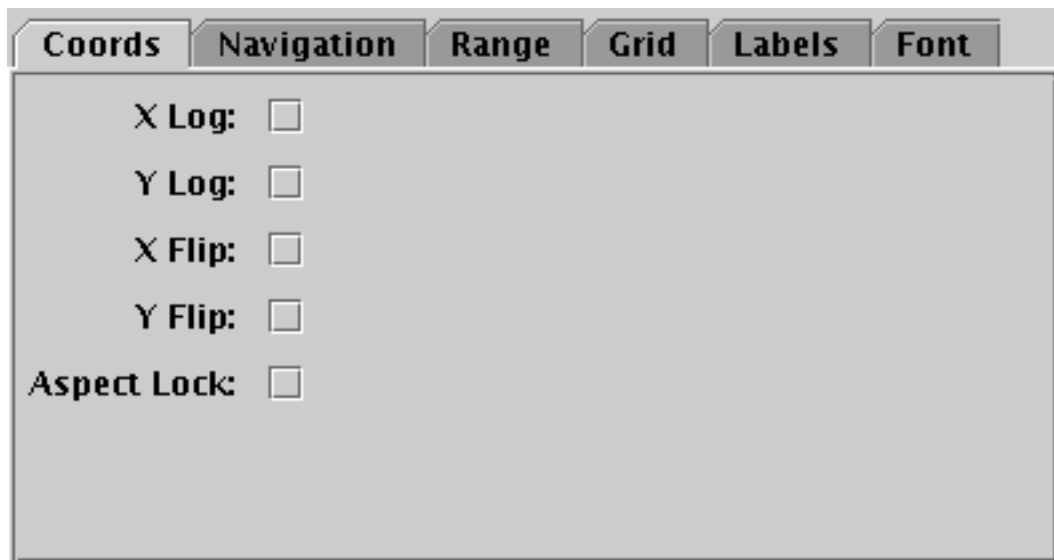
#### **Left click**

*Select.* If there is a plotted point near the cursor, it will plot a marker on it and activate (Section 8) it.

You can also manually fix the plot bounds using the **Range** tab of the Axes control.

### A.4.8.2 Plane Axes Control

The **Axes** control () for the plane plot window has the following tabs:



#### **Coords tab of plane Axes control**

The **Coords** tab controls the axis coordinates. It has the following options:

##### **X/Y Log**

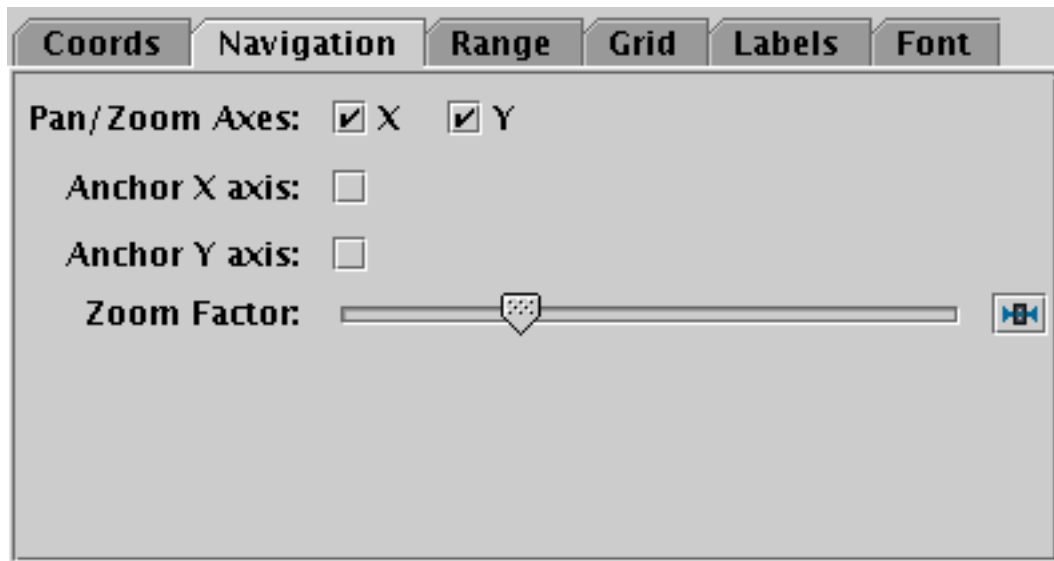
If selected, horizontal/vertical axis coordinates are logarithmic, otherwise they are linear.

##### **X/Y Flip**

If selected, horizontal/vertical axis coordinate axes run in the opposite direction to normal.

**Aspect Lock**

If selected, the number of pixels per unit is always the same on both axes, i.e. the unit square is always a square. Otherwise, there is no constraint on the relative sizes of the X and Y axis units.

**Navigation tab of plane Axes control**

The **Navigation** tab controls details of how the navigation works. It has the following options:

**Pan/Zoom Axes**

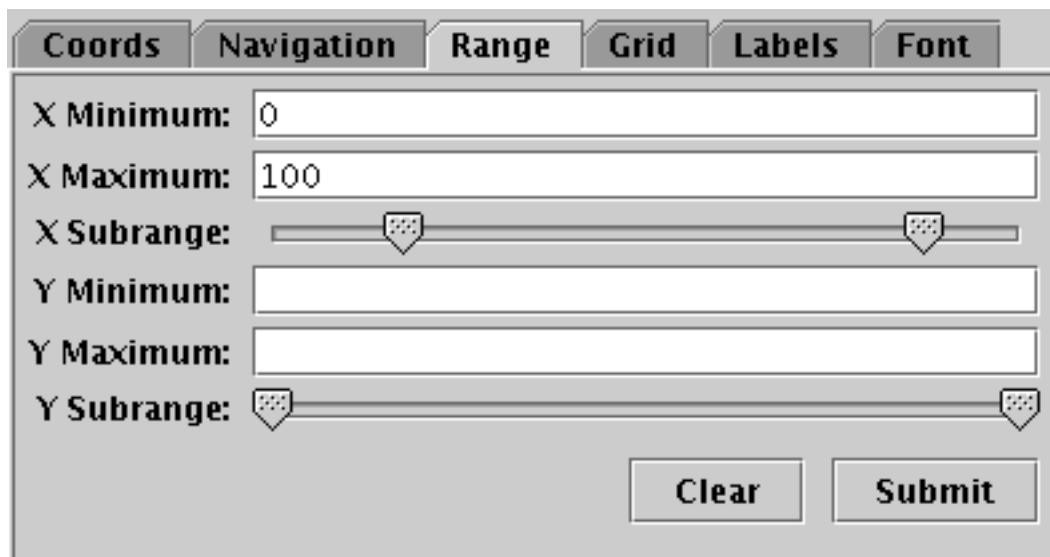
Normally, dragging with the left/right mouse buttons or using the mouse wheel on the main part of the plot will pan/zoom it in both X and Y directions. By unchecking the X or Y checkbox here, you can prevent pan/zoom in the corresponding direction, so if the Y box is unchecked, pan/zoom will only affect the vertical direction (note the same effect can be achieved by dragging to the left of the Y axis).

**Anchor X/Y axis**

Normally, zoom operations zoom around the position of the mouse at the start of the wheel/drag gesture. Checking these boxes fixes the X/Y reference coordinate for zooms to be the Y=0 or X=0 lines. This can be useful if you want the X or Y axis to stay put (e.g. at the edge of the plot) during zoom actions.

**Zoom Factor**

Controls the factor by which each zoom action zooms the plot. Moving this slider to the left/right makes the mouse more/less sensitive (one wheel click or dragging a fixed distance has more/less zoom effect).



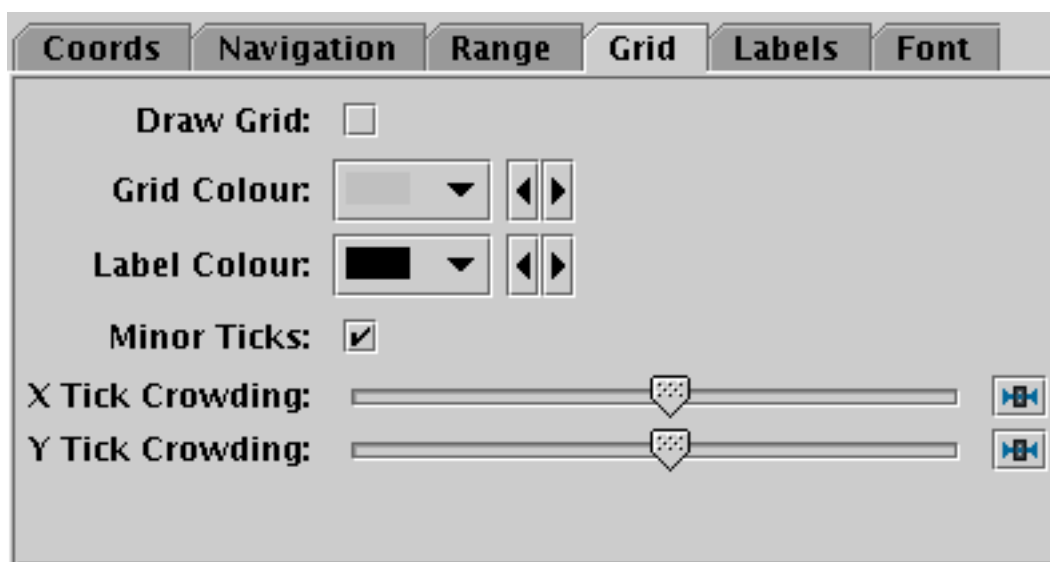
The image shows the 'Range' tab of the 'plane Axes control' dialog. It features six tabs at the top: 'Coords', 'Navigation', 'Range' (selected), 'Grid', 'Labels', and 'Font'. The 'Range' tab contains input fields for 'X Minimum' (0), 'X Maximum' (100), 'Y Minimum' (empty), and 'Y Maximum' (empty). Below these are 'X Subrange' and 'Y Subrange' double-slider controls. At the bottom right are 'Clear' and 'Submit' buttons.

### Range tab of plane Axes control

The **Range** tab provides manual configuration of the visible range of the plot. Making changes to this tab will reset the visible plot range, but not vice versa - zooming and panning in the usual way will not change the settings of this panel.

Filling in the **Minimum/Maximum** fields for either or both axes will constrain the corresponding range of the visible data. The limits corresponding to any of those fields that are left blank will initially be worked out from the data. The **Subrange** double-sliders restrict the ranges within the (explicit or automatic) min/max ranges. Note you can move both sliders at once by grabbing a position between the two.

The **Clear** button resets all the fields. The **Submit** button updates the plot with the current values in this tab, as does making any changes to it.



The image shows the 'Grid' tab of the 'plane Axes control' dialog. It features the same six tabs at the top: 'Coords', 'Navigation', 'Range', 'Grid' (selected), 'Labels', and 'Font'. The 'Grid' tab contains a 'Draw Grid' checkbox (unchecked), 'Grid Colour' and 'Label Colour' dropdown menus (both set to black), and 'Minor Ticks' checkbox (checked). At the bottom are 'X Tick Crowding' and 'Y Tick Crowding' double-slider controls. At the bottom right are 'Clear' and 'Submit' buttons.

### Grid tab of plane Axes control

The **Grid** tab configures the appearance of the axis grid. It has the following options:

#### Draw Grid

If true, grid lines will be drawn across the plot for every tick mark.

#### Grid Colour

Selects the colour with which grid lines will be drawn.

**Label Colour**

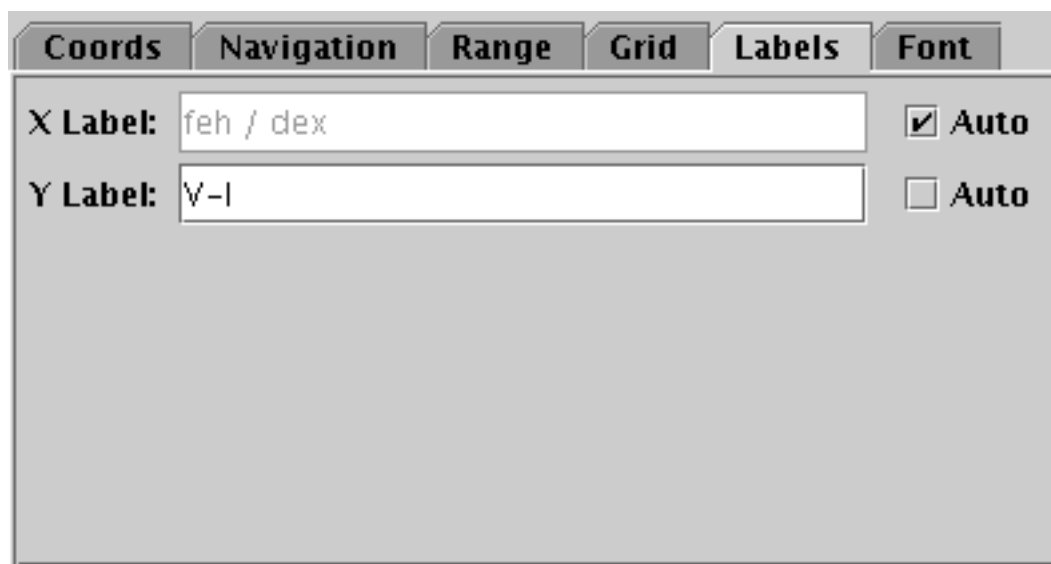
Selects the colour in which axis label text will be written.

**Minor Ticks**

If set, minor (unlabelled) tick marks will be drawn between the major (labelled) ones.

**X/Y Tick Crowding**

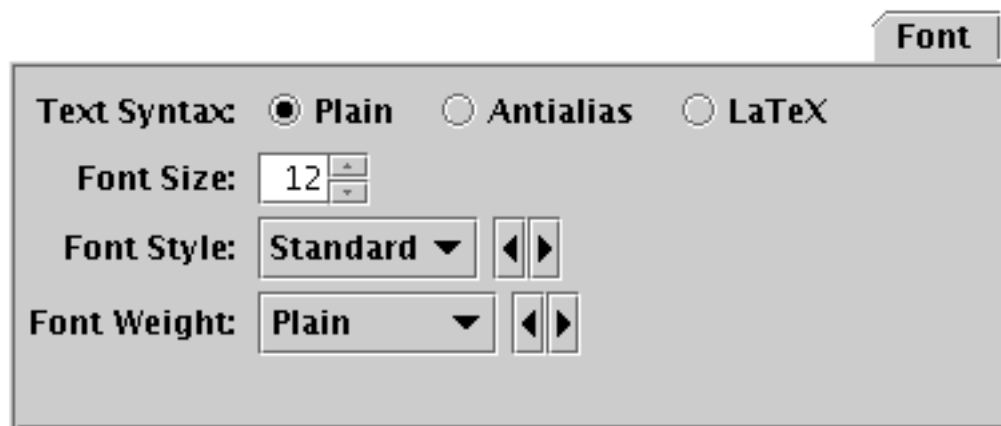
Use the slider to influence how many tick marks are drawn on each axis.



The image shows the 'Labels' tab of a control panel. It has a header bar with tabs: 'Coords', 'Navigation', 'Range', 'Grid', 'Labels' (selected), and 'Font'. Below the header, there are two rows of controls. The first row is for the 'X Label', with a text input field containing 'feh / dex' and a checked 'Auto' checkbox. The second row is for the 'Y Label', with a text input field containing 'V-I' and an unchecked 'Auto' checkbox.

**Labels tab of plane Axes control**

The **Labels** tab controls the text labels on the axes. If the **Auto** checkbox is set, the text will be taken from one of the data coordinates being plotted on that axis. To override those with your own axis labels, unset Auto and type text in to the **Label** fields.



The image shows the 'Font' tab of a control panel. It has a header bar with a single tab: 'Font'. Below the header, there are four rows of controls. The first row is 'Text Syntax' with three radio buttons: 'Plain' (selected), 'Antialias', and 'LaTeX'. The second row is 'Font Size' with a text input field containing '12' and up/down arrow buttons. The third row is 'Font Style' with a dropdown menu showing 'Standard' and left/right arrow buttons. The fourth row is 'Font Weight' with a dropdown menu showing 'Plain' and left/right arrow buttons.

**Font tab**

The **Font** tab configures the font used for axis annotation. It also affects some other things like the legend.

**Text Syntax**

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at face value, but **Antialias** smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

**Font Size**

Size of the font in points.

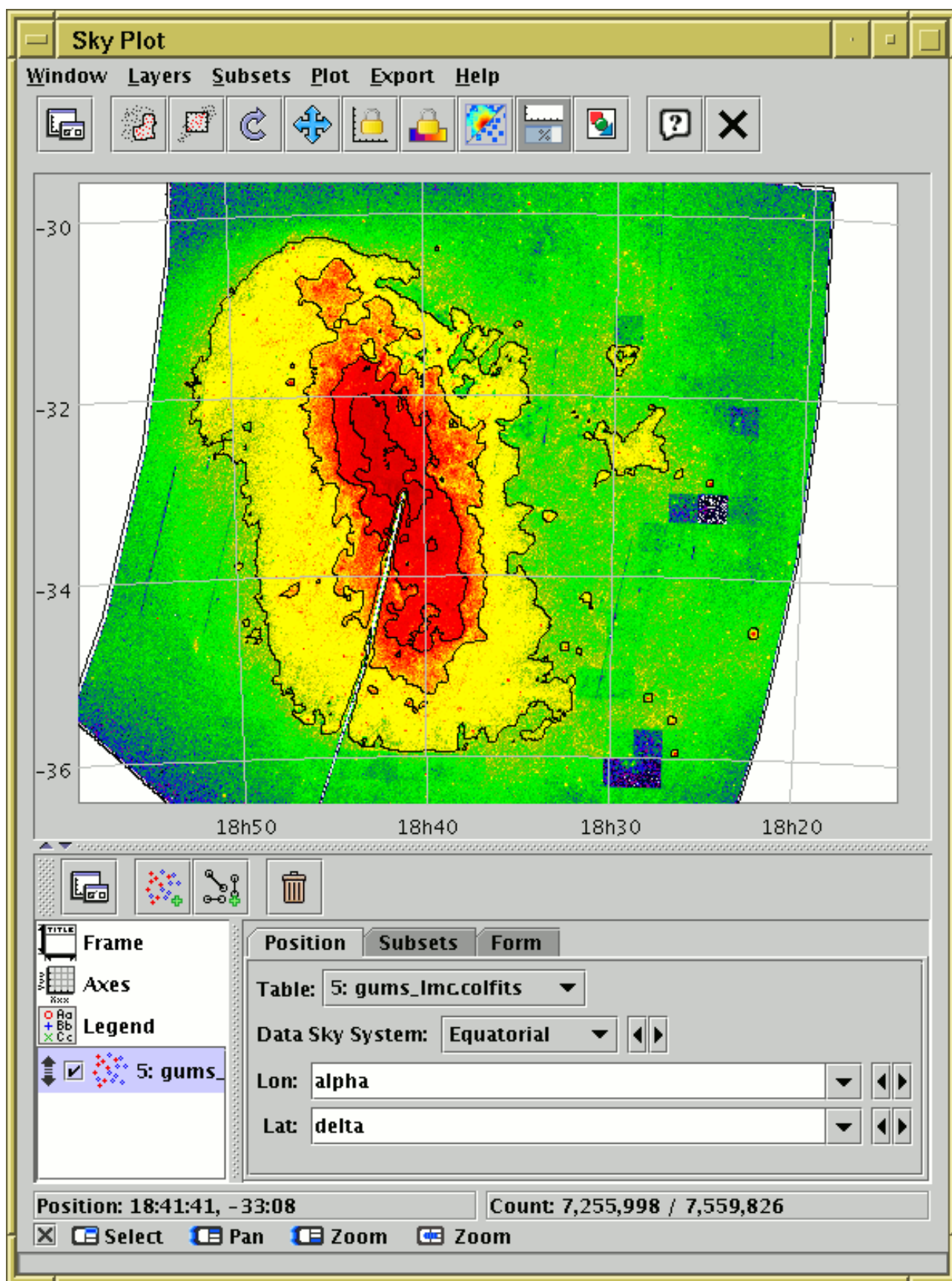
**Font Style**

Style of the font.


**Font Weight**

Whether the font is plain, bold or italic.

**A.4.9 Sky Plot Window**
















### Sky Plot Window

The **Sky Plot** () plots longitude/latitude positions onto the celestial sphere. It can plot to a number of projections (currently Sin, Aitoff and Plate Carrée).



The positional coordinates are **Longitude** and **Latitude**, specified in degrees. When supplying them, you can specify an associated **Data Sky System** (Equatorial, Galactic, Supergalactic or Ecliptic2000). Note that this is the sky system of the data coordinates, not (necessarily) of the plot you want to see. To specify the coordinates you want the data to be plotted on, use the **View Sky System** option in the **Projection** tab of the Axes control. However, if you just want to view the data using the same system as the coordinates you are supplying, you can ignore leave these values as their default (both Equatorial) and no conversion will be done.

The sky plot offers the following plot controls:

-  Position Layer Control, with these form options:
  -  Mark
  -  Size
  -  SizeXY
  -  SkyVector
  -  SkyEllipse
  -  Text Label
  -  Contour
  -  SkyDensity
-  Pair Layer Control, with the following plot controls:
  -  Marks
  -  Link
-  Healpix Layer Control

See the Window Overview (Appendix A.4.2) for features common to all plotting windows. The following subsections describe navigation and axis configuration.

#### A.4.9.1 Sky Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

The navigation options for this window are:

##### Left drag

*Pan.* In the Sin projection, this attempts to rotate the sphere, with north staying vertical on the screen, in such a way that the same part of the sky stays under the cursor position. At low magnifications this is sometimes not possible; if the mouse starts on the sphere and ends up off it the navigation will be jerky.

The Aitoff and Plate Carée projections act like a piece of paper that can be dragged around in

two dimensions in the usual way.

#### **Wheel**

*Zoom.* Spinning the mouse wheel forwards/backwards will zoom in/out around the mouse position. For this Sin projection this will also have the effect of rotating the sphere to keep North vertical on the screen.

#### **Right drag (CTRL-drag)**

*Stretch zoom.* Dragging with the right mouse button up-and-right/down-and-left has just the same effect as spinning the mouse button forwards/backwards.

#### **Center drag (SHIFT-drag)**


*Frame zoom.* Dragging with the middle button to the right (and either up or down) drags out a frame; when the button is released, the plot will be zoomed in to cover (roughly) the area enclosed by the frame. Dragging left (and either up or down) does something like the opposite, you can zoom out using a similar (though not quite the same) mechanism.

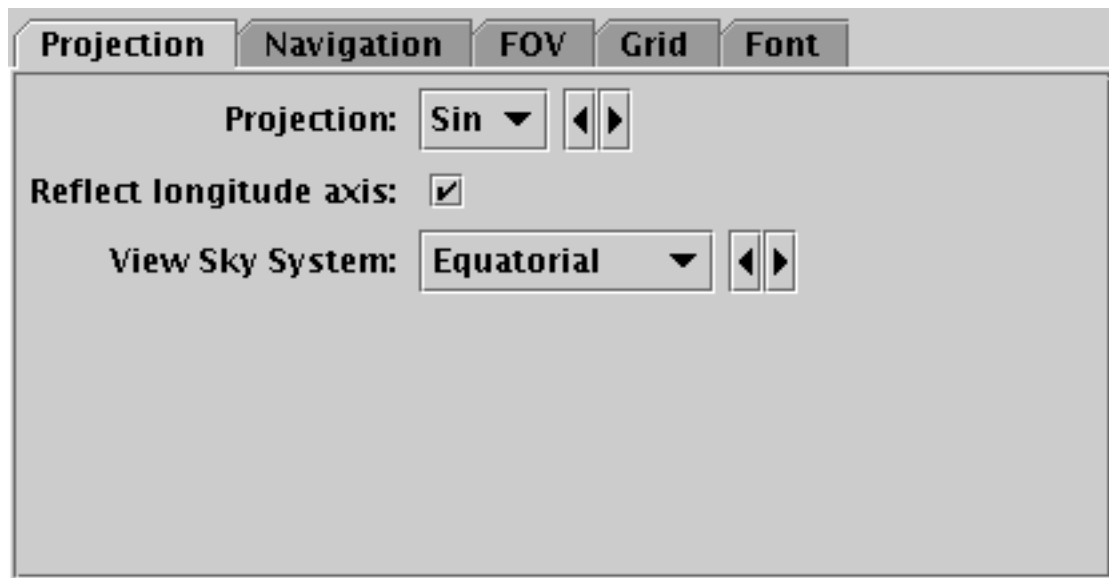
#### **Left click**

*Select.* If there is a plotted point near the cursor, it will plot a marker on it and activate (Section 8) it.

You can also navigate to a known sky position by coordinates or object name using the **FOV** tab of the Axes control.

### A.4.9.2 Sky Axes Control

The **Axes** control () for the sky plot window has the following tabs:



#### **Projection tab of the sky Axes control**

The **Projection** tab controls how the sky position coordinates are projected onto the screen.

##### **Projection**

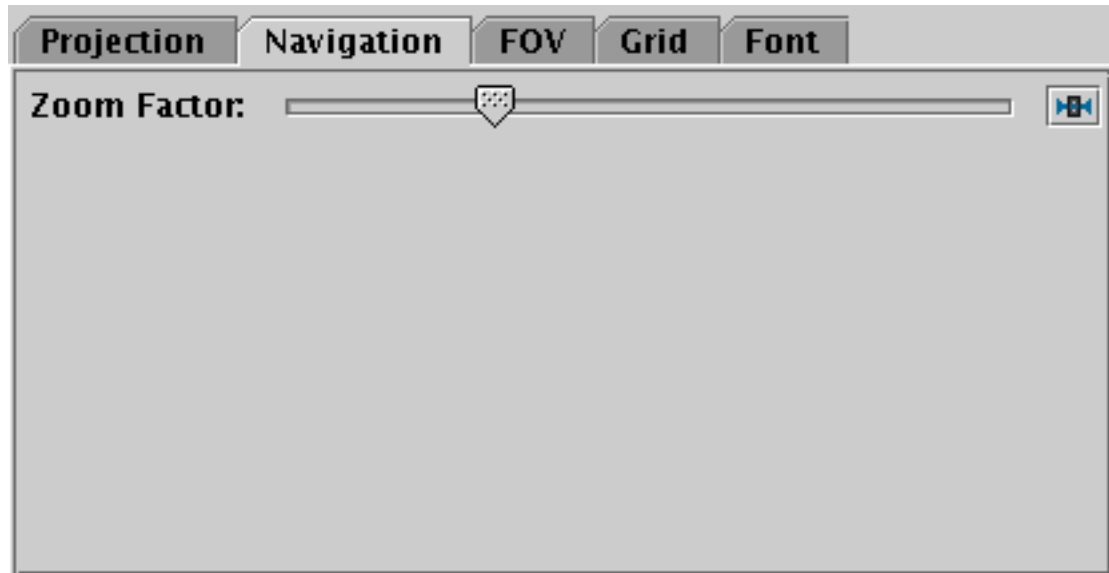
Selects the sky projection to use. The options are **Sin**, **Ait** for Aitoff and **Car** for Plate Carée. Sin is rotatable, the other two are essentially flat all-sky projections.

##### **Reflect longitude axis**

Determines whether longitude increases left to right or right to left.

##### **View Sky System**

Determines the sky coordinate system in which the data positions will be viewed. This interacts with the **Data Sky System** selected in the **Positions** tab of the table layer control for data coordinates; supplied data points are projected from the data system to the view system before being plotted. If you have (for instance) data in equatorial coordinates that you want to view in galactic coordinates, then select the Data Sky System as Equatorial and the View Sky System as Galactic. If the data and view systems are the same, it's OK to leave both as their defaults, even if they're not equatorial - the only effect that the chosen Data and View sky systems has is from the transformation between them.

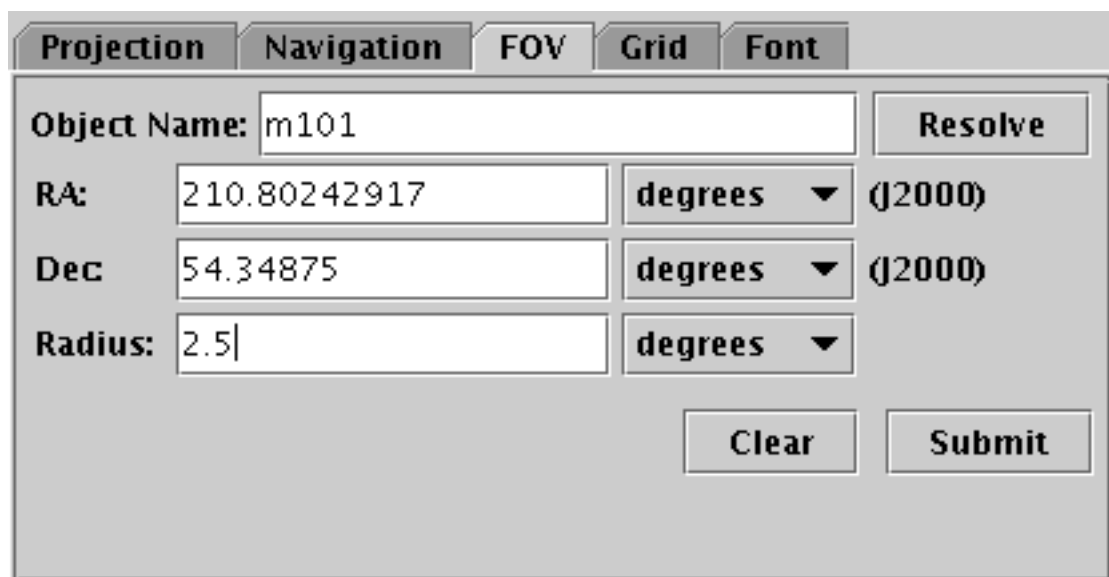


### Navigation tab of sky Axes control

The **Navigation** tab controls details of how the navigation works. It has the following option:

#### Zoom Factor

Controls the factor by which each zoom action zooms the plot. Moving this slider to the left/right makes the mouse more/less sensitive (one wheel click or dragging a fixed distance has more/less zoom effect).



### Field Of View tab of the sky Axes control

The **FOV** (Field Of View) tab allows you to enter a sky position or object name and a radius and

positions the view at that region. Filling in the fields and hitting **Submit** will reposition the sky view, but not vice versa; normal pan/zoom operations will not affect the content of this panel.

### Object Name

If you fill this in with the name of a celestial object and hit the **Resolve** button, a Simbad query will execute to fill in the **RA** and **Dec** fields with its position.

### RA/Dec

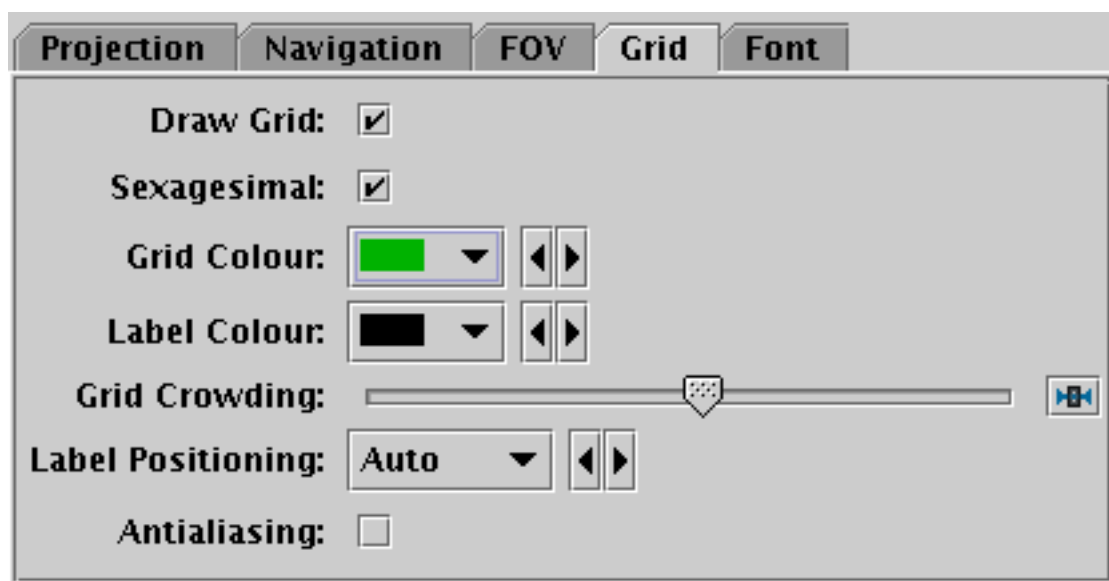
J2000 positions of the required field centre. These values can either be filled in by the object name resolution as described above, or by hand.

### Radius

Gives the radius of the desired field of view.

### Clear

Clears the fields in this tab.



### Grid tab of the sky Axes control

The **Grid** tab controls how the sky coordinate axes appear.

#### Draw Grid

If selected, grid lines will be drawn on the plot.

#### Sexagesimal

If selected sky coordinate annotations of the grid will be in sexagesimal format, otherwise in decimal degrees.

#### Grid Colour

Selects the colour with which grid lines will be drawn.

#### Label Colour

Selects the colour in which axis label text will be written.

#### Grid Crowding

Use the slider to control how closely packed grid lines are on the axes.

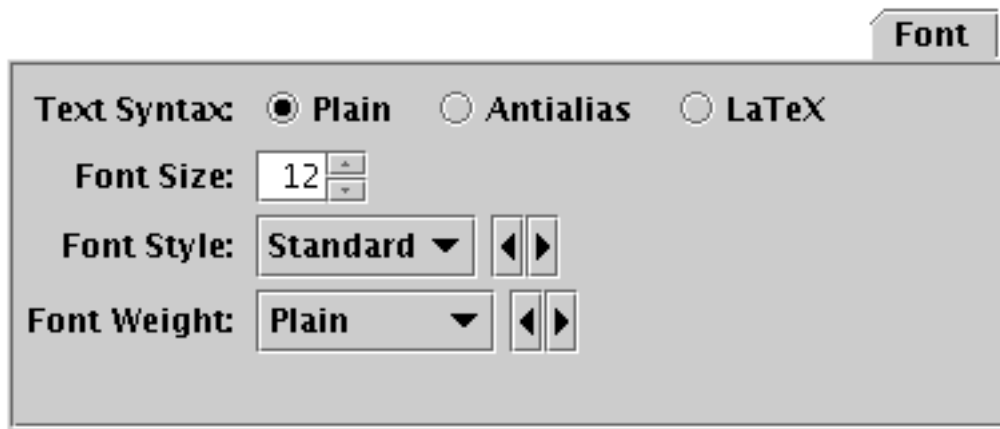
#### Label Positioning

Controls whether and where the numeric annotations of the lon/lat axes are put. The default option **Auto** usually does the sensible thing, but other options exist to force labelling internally or externally to the plot region, or to remove numeric labels altogether.

#### Antialiasing

Controls whether grid lines will be drawn antialiased (smoothed) or not. This option does not

affect exported plots.



### Font tab

The **Font** tab configures the font used for axis annotation. It also affects some other things like the legend.

#### Text Syntax

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at face value, but **Antialias** smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

#### Font Size

Size of the font in points.

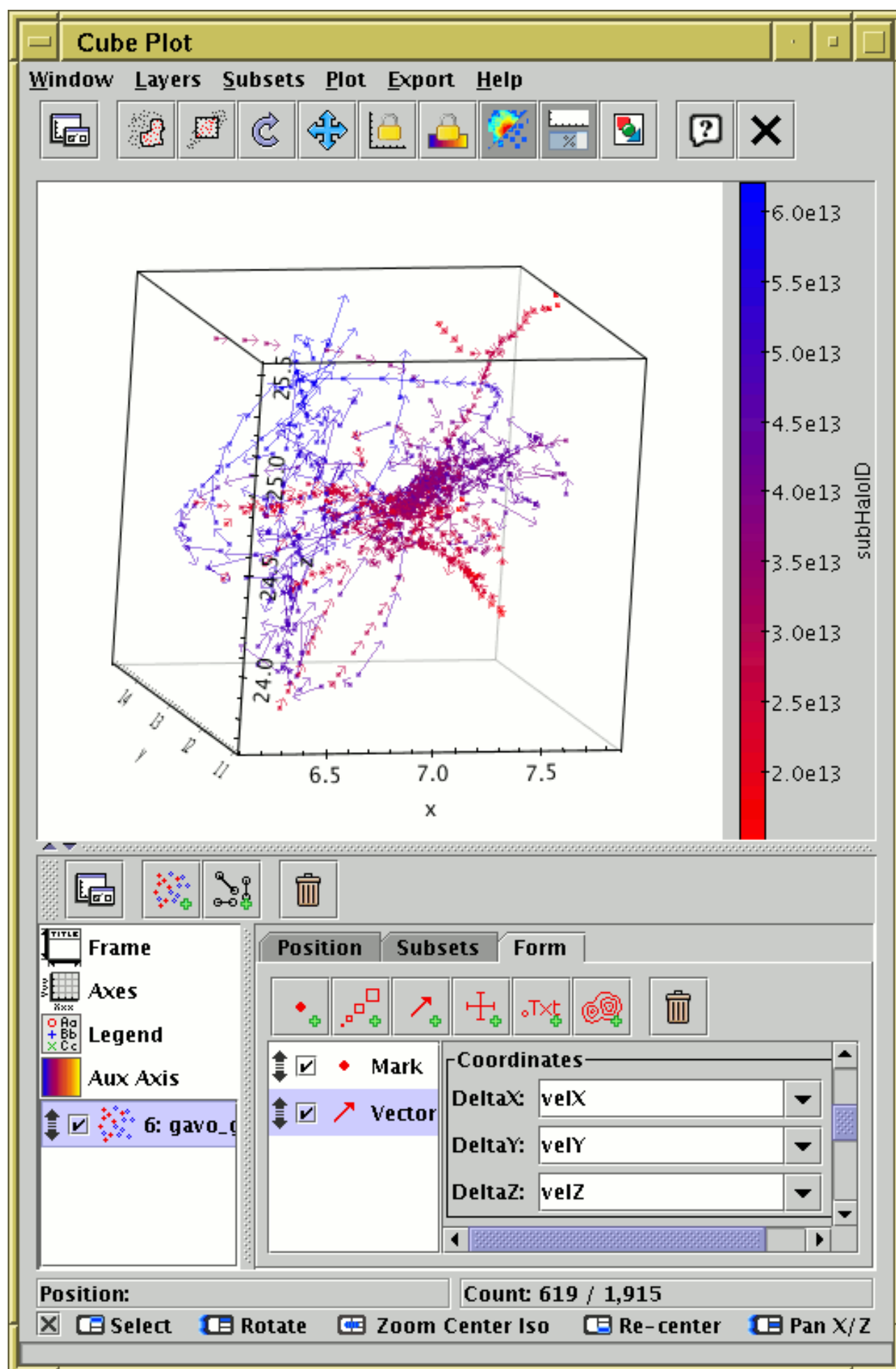
#### Font Style

Style of the font.


#### Font Weight

Whether the font is plain, bold or italic.

### A.4.10 Cube Plot Window














## Cube Plot Window

The **Cube Plot** () plots 3-dimensional Cartesian positions in a 3-d space.

The positional coordinates are **X**, **Y** and **Z**. To control the direction and linear/logarithmic scaling of the axes, see the **Coords** tab of the Axes control.

The cube plot offers the following plot controls:

-  Position Layer Control, with these form options:
  -  Mark
  -  Size
  -  SizeXY
  -  Vector
  -  Error Bars
  -  Text Label
  -  Contour
-  Pair Layer Control, with the following plot controls:
  -  Marks
  -  Link

Note that use of the Auto, Density and Weighted shading modes can be confusing in 3 dimensions with multiple datasets. This is because pixels based on density along a line of sight are not located at any point on that line, so shaded pixels can't appear at the "right" place in the 3-d space. The same applies to a lesser extent with contours. They work fine with a single dataset though.

See the Window Overview (Appendix A.4.2) for features common to all plotting windows. The following subsections describe navigation and axis configuration.

### A.4.10.1 Cube Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

Navigation in three dimensions with a two-dimensional screen and mouse is a bit of a challenge. However, the actions listed below should make it fairly straightforward to navigate around points in 3d space to zoom in on the regions that you are interested in.



#### Left drag

*Rotate.* Rotates the view cube about its center.



#### Wheel

*Isotropic zoom.* Spinning the mouse wheel forwards/backwards zooms the space in the view

cube in/out around the cube center. The mouse position does not affect it.

#### **Right drag (CTRL-drag)**

*2d stretch zoom.* Dragging with the right button does a stretch zoom in the two dimensions best aligned with the plane of the screen. There is no movement in the third (mostly perpendicular to the screen) direction. The zoom is around a line defined by the position of the mouse at the start of the zoom, pointing along the third dimension. You can stretch/squash in both directions by dragging the mouse up/down/left/right - it's easier to try it than to explain it.

#### **Center drag (SHIFT-drag)**

*2d pan.* Dragging with the center button pans the 3d space in the two dimensions best aligned with the plane of the screen. There is no movement in the third (mostly perpendicular to the screen) direction.

#### **Left click**

*Select.* If there are plotted points near the cursor it will identify one, plot a marker on it, and activate (Section 8) it.


In 3d, it's not obvious which is the nearest point to a 2d cursor, since a screen position represents a line not a point. To break the degeneracy, the point used is the one nearest the center of mass of plotted points along the line of sight represented by the cursor position. The upshot of this is that if you click on an isolated point, you'll pick that point, and if you click on a dense cluster, you'll get a point near the center (of mass) of that cluster.

#### **Right click (CTRL-click)**

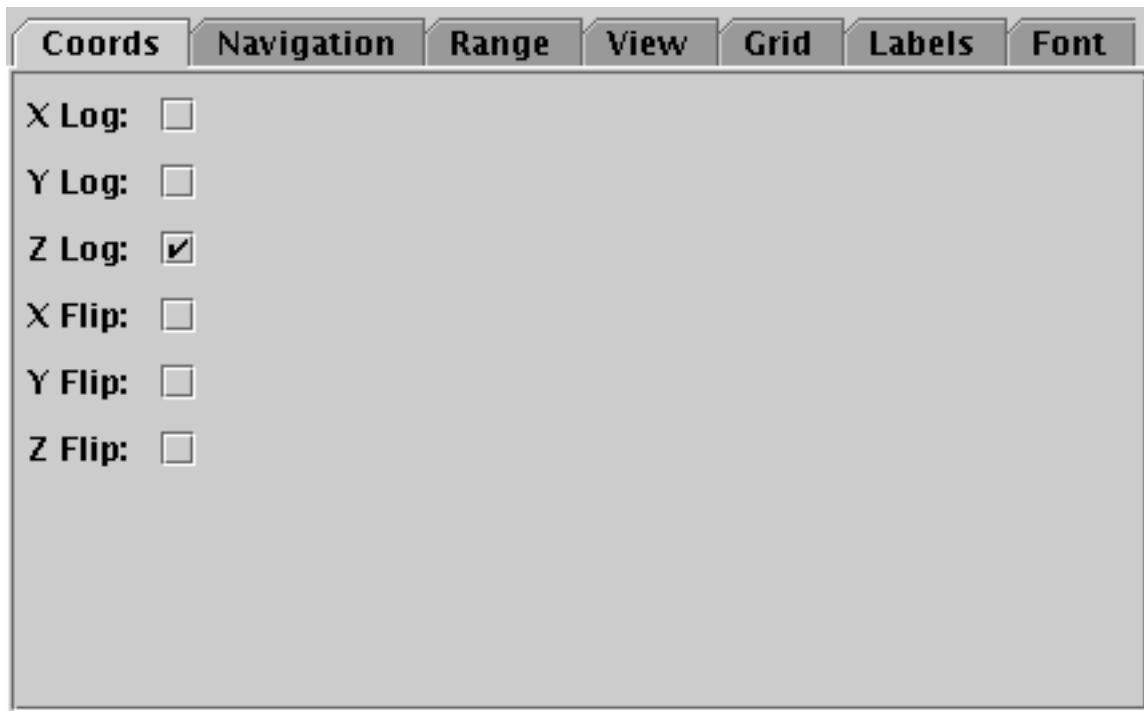
*Re-center.* Right-clicking identifies a position in a similar way to the left-click Select action, and then translates the plot so that point is at the center of the view cube. This means that clicking on a cluster will put that cluster in the center of the cube. If you click on an empty region, a position half way between front and back of the cube at that X/Y position will be used.

You can also manually fix the plot bounds using the **Range** tab of the Axes control.

### A.4.10.2 Cube Axes Control

The **Axes** control () for the cube plot window has the following tabs:





#### Coords tab of the cube Axes control

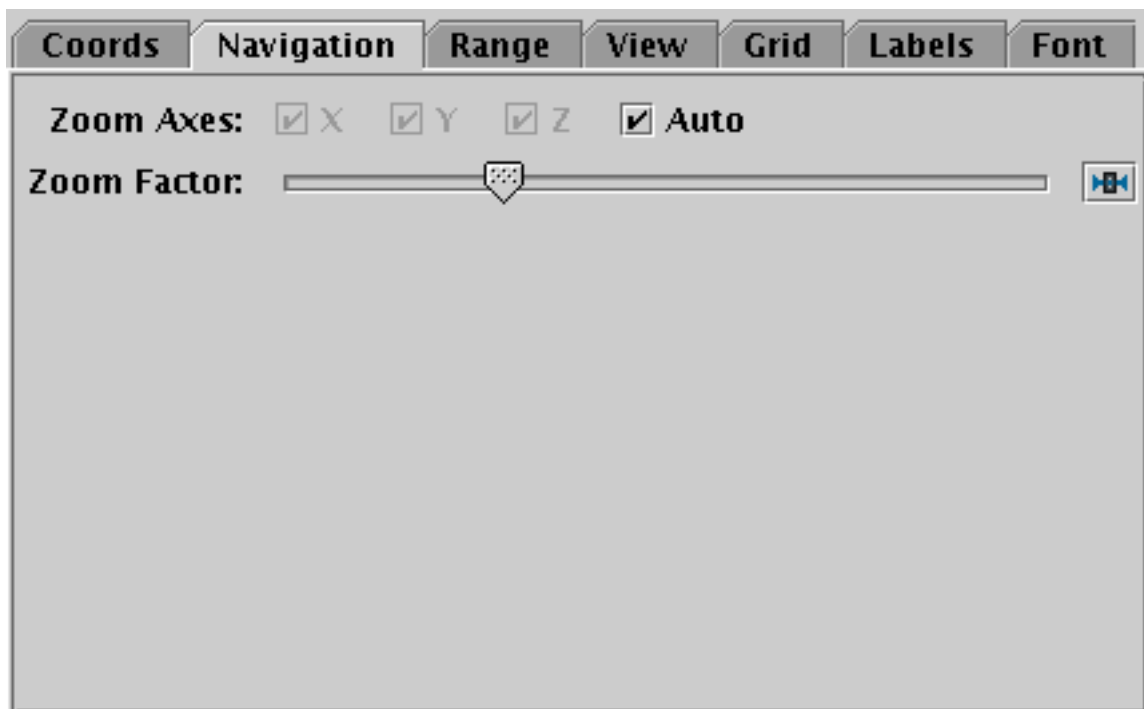
The **Coords** tab controls the axis coordinates. It has the following options:

##### X/Y/Z Log

If selected, X/Y/Z axis coordinates are logarithmic, otherwise they are linear.

##### X/Y/Z Flip

If selected, X/Y/Z axis coordinate axes run in the opposite direction to normal.



#### Navigation tab of the cube Axes control

The **Navigation** tab controls details of how the navigation works. It has the following options:

##### Zoom Axes

By default the **Auto** setting is in effect. This means that the mouse wheel zooms around the center of the cube, and right-button drag zooms in the two dimensions most closely aligned with the plane of the screen, with the reference position set by the initial position of the mouse. If **Auto** is unset, then all zoom operations are around the cube center, and affect the axes indicated by the **X/Y/Z** checkboxes.

### Zoom Factor

Controls the factor by which each zoom action zooms the plot. Moving this slider to the left/right makes the mouse more/less sensitive (one wheel click or dragging a fixed distance has more/less zoom effect).

The image shows a software interface for configuring a cube plot. It features a series of tabs: **Coords**, **Navigation**, **Range** (selected), **View**, **Grid**, **Labels**, and **Font**. Under the **Range** tab, there are input fields and sliders for the visible range of the plot along the X, Y, and Z axes. For the X-axis, the minimum is set to 0 and the maximum to 100. The Y and Z axes have empty minimum and maximum fields. Each axis also has a 'Subrange' slider to restrict the visible range within the specified min/max. At the bottom right, there are **Clear** and **Submit** buttons.

### Range tab of the cube Axes control

The **Range** tab provides manual configuration of the visible range of the plot. Making changes to this tab will reset the visible plot range, but not vice versa - zooming and panning in the usual way will not change the settings of this panel.

Filling in the **Minimum/Maximum** fields for one or more axes will constrain the corresponding range of the visible data. The limits corresponding to any of those fields that are left blank will initially be worked out from the data. The **Subrange** double-sliders restrict the ranges within the (explicit or automatic) min/max ranges. Note you can move both sliders at once by grabbing a position between the two.

The **Clear** button resets all the fields. The **Submit** button updates the plot with the current values in this tab, as does making any changes to it.

The image shows a software interface with a tabbed menu at the top. The tabs are 'Coords', 'Navigation', 'Range', 'View', 'Grid', 'Labels', and 'Font'. The 'View' tab is currently selected. Below the tabs, there are five control rows, each with a label on the left and a slider control on the right. The labels are: 'Rotation about Z axis:', 'Rotation from vertical:', 'Zoom factor:', 'X offset of centre:', and 'Y offset of centre:'. Each slider has a central knob with a small icon above it. To the right of each slider is a small icon that looks like a cube with axes. At the bottom right of the 'View' tab area are two buttons: 'Clear' and 'Submit'.

### View tab of the cube Axes control

The **View** tab can configure how the cube containing the data is viewed in the plot window, though it does not control the content of the cube.

#### Rotation about Z axis

Explicitly sets the rotation angle about the Z axis, regardless of its current value, which may have been set by mouse actions.

#### Rotation from vertical

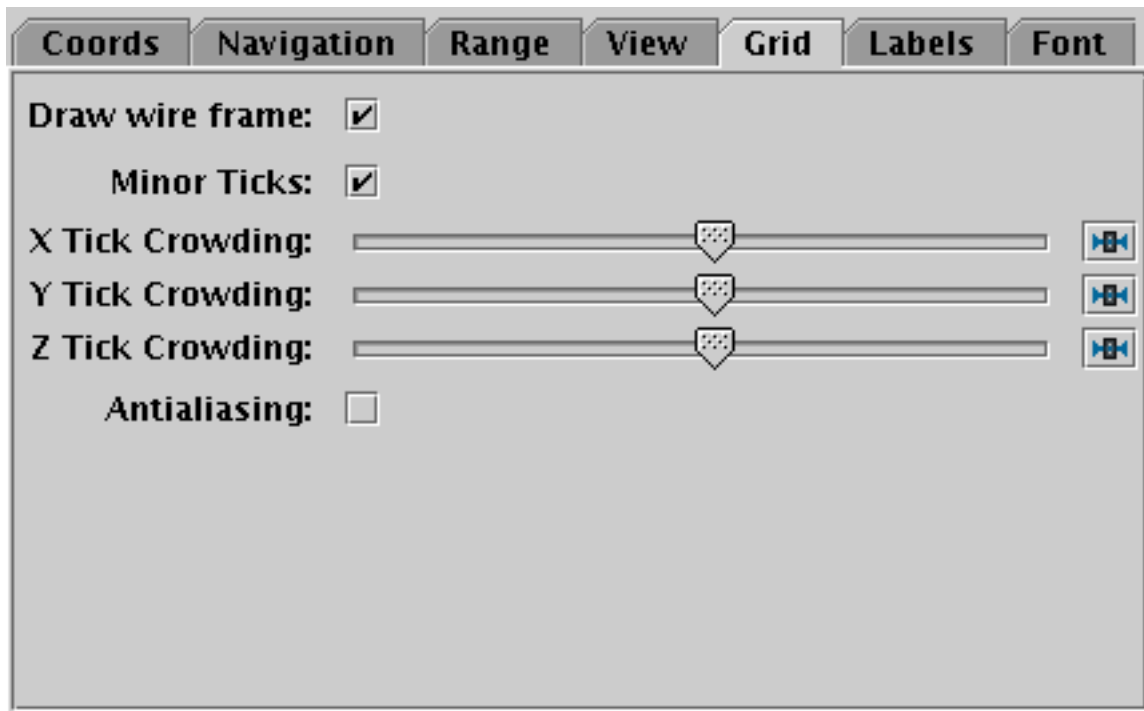
Explicitly sets the rotation angle from vertical, regardless of its current value, which may have been set by mouse actions.

#### Zoom factor

Sets the magnification of the cube wireframe itself, without affecting the data volume it contains. This cannot be done with the mouse.

#### X/Y/Z offset of centre

Controls where on the screen the cube wireframe is centred. This cannot be done with the mouse.



### Grid tab of the cube Axes control

The **Grid** tab configures the appearance of the cube wire frame enclosing the data volume.

#### **Draw wire frame**

Whether the enclosing cube is drawn at all.

#### **Minor Ticks**

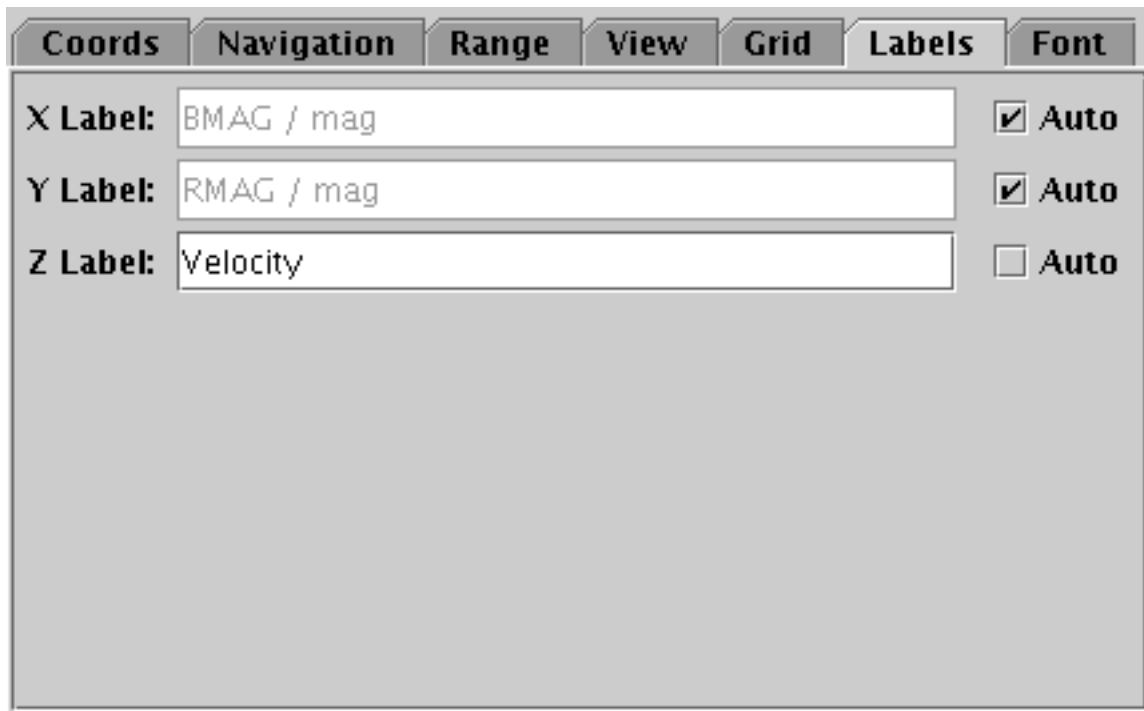
If set, minor (unlabelled) tick marks will be drawn between the major (labelled) ones.

#### **X/Y/Z Tick Crowding**

Use the slider to influence how many tick marks are drawn on each axis.

#### **Antialiasing**

Controls whether grid lines will be drawn antialiased (smoothed) or not. This option does not affect exported plots.

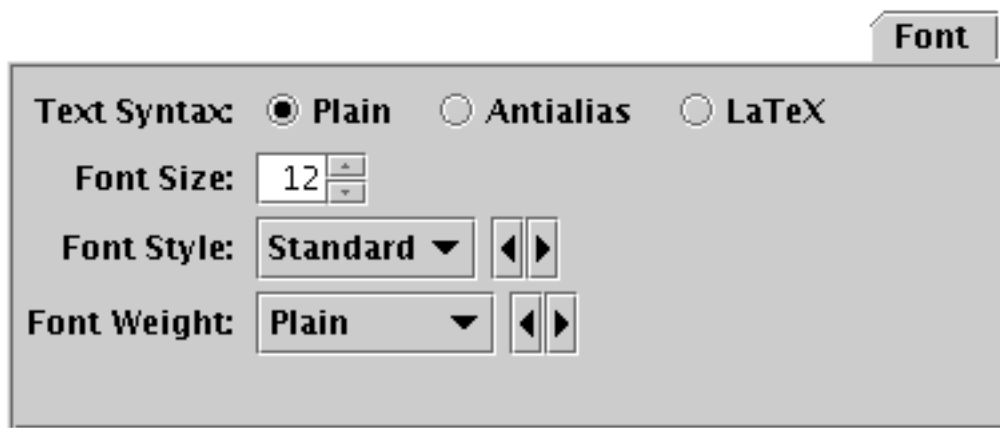


The image shows the 'Labels' tab of a software interface. It has a header bar with tabs: 'Coords', 'Navigation', 'Range', 'View', 'Grid', 'Labels' (selected), and 'Font'. Below the header, there are three rows of controls:

- X Label:** A text box containing 'BMAG / mag' and a checked checkbox labeled 'Auto'.
- Y Label:** A text box containing 'RMAG / mag' and a checked checkbox labeled 'Auto'.
- Z Label:** A text box containing 'Velocity' and an unchecked checkbox labeled 'Auto'.

### Labels tab of the cube Axes control

The **Labels** tab controls the text labels on the axes. If the **Auto** checkbox is set, the text will be taken from one of the data coordinates being plotted on that axis. To override those with your own axis labels, unset Auto and type text in to the **Label** fields.



The image shows the 'Font' tab of the same software interface. It has a header bar with the 'Font' tab selected. The controls include:

- Text Syntax:** Three radio buttons: 'Plain' (selected), 'Antialias', and 'LaTeX'.
- Font Size:** A text box with '12' and up/down arrow buttons.
- Font Style:** A dropdown menu showing 'Standard' and left/right arrow buttons.
- Font Weight:** A dropdown menu showing 'Plain' and left/right arrow buttons.

### Font tab

The **Font** tab configures the font used for axis annotation. It also affects some other things like the legend.

#### Text Syntax

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at face value, but **Antialias** smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

#### Font Size

Size of the font in points.

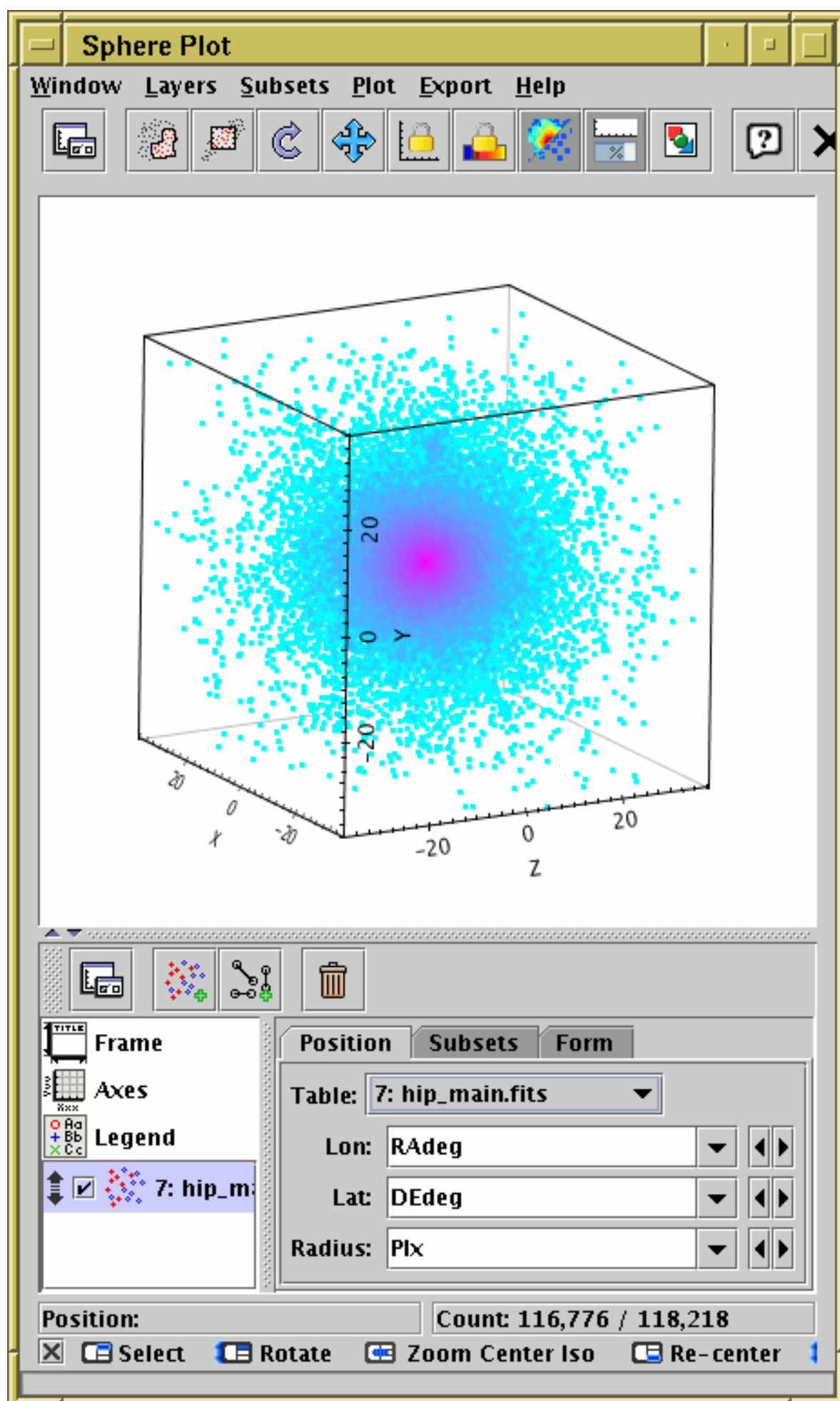
#### Font Style

Style of the font.


**Font Weight**

Whether the font is plain, bold or italic.

**A.4.11 Sphere Plot Window**












## Sphere Plot Window

The **Sphere Plot** () plots spherical polar coordinates in an isotropic 3-dimensional space.

Although the supplied coordinates are spherical polar, the visible space is not necessarily centred on the coordinate origin, and the visible axes are Cartesian. In many respects this works like the Cube Plot Window

The positional coordinates are **Longitude** and **Latitude** (in degrees) and **Radius**.

The sphere plot offers the following plot controls:

-  Position Layer Control, with these form options:
  -  Mark
  -  Size
  -  SizeXY
  -  Text Label
  -  Contour
-  Pair Layer Control, with the following plot controls:
  -  Marks
  -  Link

Note that use of the Auto, Density and Weighted shading modes can be confusing in 3 dimensions with multiple datasets. This is because pixels based on density along a line of sight are not located at any point on that line, so shaded pixels can't appear at the "right" place in the 3-d space. The same applies to a lesser extent with contours. They work fine with a single dataset though.

See the Window Overview (Appendix A.4.2) for features common to all plotting windows. The following subsections describe navigation and axis configuration.

### A.4.11.1 Sphere Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

Navigation in three dimensions with a two-dimensional screen and mouse is a bit of a challenge. However, the actions listed below should make it fairly straightforward to navigate around points in 3d space to zoom in on the regions that you are interested in. The actions are quite like for the Cube Window, but all zooming is isotropic.

#### Left drag

*Rotate.* Rotates the view cube about its center.

#### Wheel

*Zoom.* Spinning the mouse wheel forwards/backwards zooms the space in the view cube in/out around the cube center. The mouse position does not affect it.



### **Right drag (CTRL-drag)**

*Zoom.* Zooming up-and-right/down-and-left does just the same as spinning the mouse wheel forwards/backwards.

### **Center drag (SHIFT-drag)**

*2d pan.* Dragging with the center button pans the 3d space in the two dimensions best aligned with the plane of the screen. There is no movement in the third (mostly perpendicular to the screen) direction.

### **Left click**

*Select.* If there are plotted points near the cursor it will identify one, plot a marker on it, and activate (Section 8) it.


In 3d, it's not obvious which is the nearest point to a 2d cursor, since a screen position represents a line not a point. To break the degeneracy, the point used is the one nearest the center of mass of plotted points along the line of sight represented by the cursor position. The upshot of this is that if you click on an isolated point, you'll pick that point, and if you click on a dense cluster, you'll get a point near the center (of mass) of that cluster.

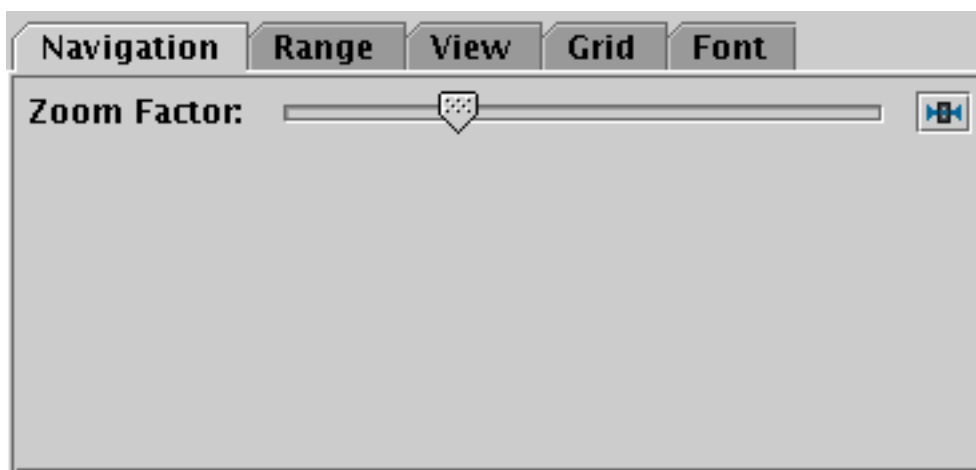
### **Right click (CTRL-click)**

*Re-center.* Right-clicking identifies a position in a similar way to the left-click Select action, and then translates the plot so that point is at the center of the view cube. This means that clicking on a cluster will put that cluster in the center of the cube. If you click on an empty region, a position half way between front and back of the cube at that X/Y position will be used.

You can also manually fix the plot bounds using the **Range** tab of the Axes control.

## A.4.11.2 Sphere Axes Control

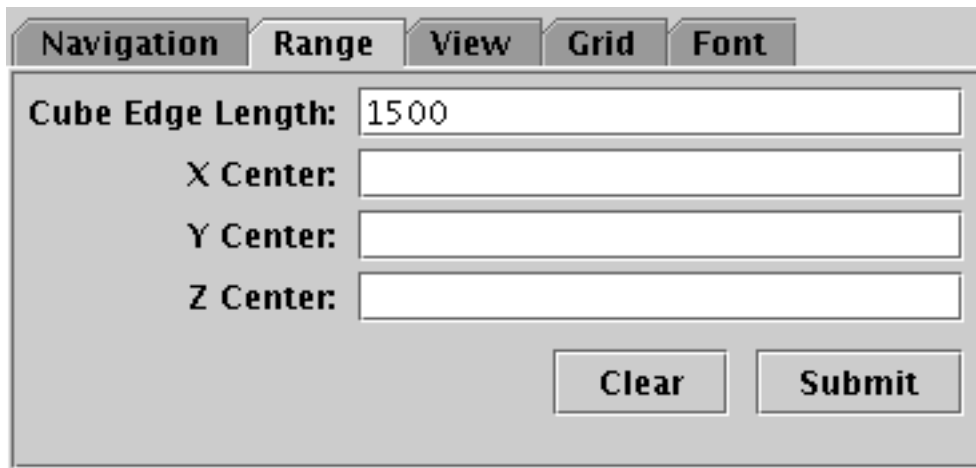
The **Axes** control () for the sphere plot window has the following tabs:



The **Navigation** tab controls details of how the navigation works. It has the following option:

### **Zoom Factor**

Controls the factor by which each zoom action zooms the plot. Moving this slider to the left/right makes the mouse more/less sensitive (one wheel click or dragging a fixed distance has more/less zoom effect).



The image shows the 'Range' tab of a control panel. It has five tabs at the top: 'Navigation', 'Range' (selected), 'View', 'Grid', and 'Font'. Below the tabs, there are four input fields: 'Cube Edge Length' with the value '1500', 'X Center', 'Y Center', and 'Z Center'. At the bottom right, there are two buttons: 'Clear' and 'Submit'.

### Range tab of the sphere Axes control

The **Range** tab provides manual configuration of the data range of the plot. Making changes to this tab will reset the visible plot range, but not vice versa - zooming and panning in the usual way will not change the settings of this panel. Any values not filled in will be determined from the data. The fields are:

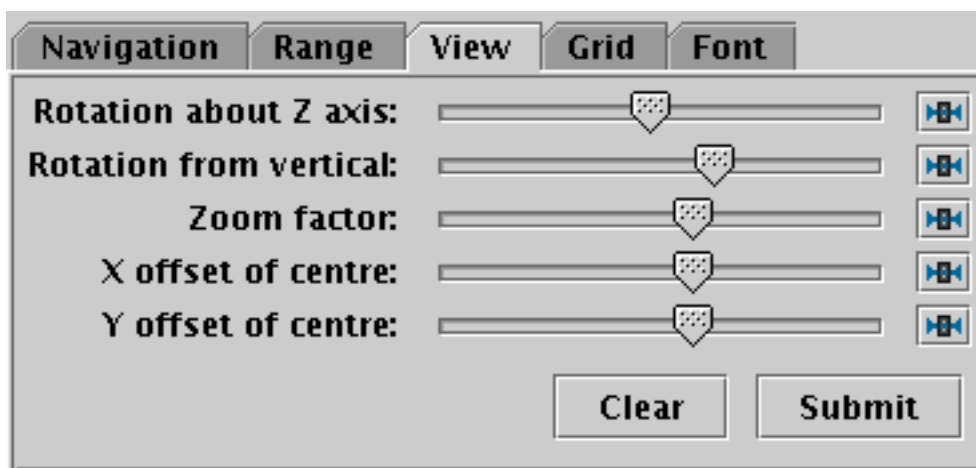
#### Cube Edge Length

Specifies the dimension along each side of the view cube in data units.

#### X/Y/Z Center

Gives the position of the center of the view cube in data coordinates.

The **Clear** button resets all the fields. The **Submit** button updates the plot with the current values in this tab, as does making any changes to it.



The image shows the 'View' tab of a control panel. It has five tabs at the top: 'Navigation', 'Range', 'View' (selected), 'Grid', and 'Font'. Below the tabs, there are five sliders: 'Rotation about Z axis', 'Rotation from vertical', 'Zoom factor', 'X offset of centre', and 'Y offset of centre'. Each slider has a dotted triangular handle and a blue double-headed arrow button on the right. At the bottom right, there are two buttons: 'Clear' and 'Submit'.

### View tab of the sphere Axes control

The **View** tab can configure how the cube containing the data is viewed in the plot window, though it does not control the content of the cube.

#### Rotation about Z axis

Explicitly sets the rotation angle about the Z axis, regardless of its current value, which may have been set by mouse actions.

#### Rotation from vertical

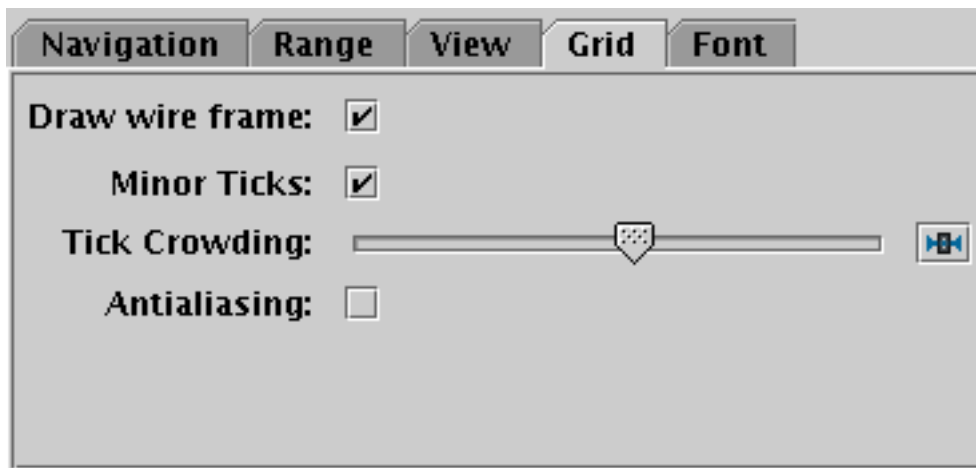
Explicitly sets the rotation angle from vertical, regardless of its current value, which may have been set by mouse actions.

#### Zoom factor

Sets the magnification of the cube wireframe itself, without affecting the data volume it contains. This cannot be done with the mouse.

### X/Y offset of centre

Controls where on the screen the cube wireframe is centred. This cannot be done with the mouse.



### Grid tab of the sphere Axes control

The **Grid** tab configures the appearance of the cube wire frame enclosing the data volume.

#### Draw wire frame

Whether the enclosing cube is drawn at all.

#### Minor Ticks

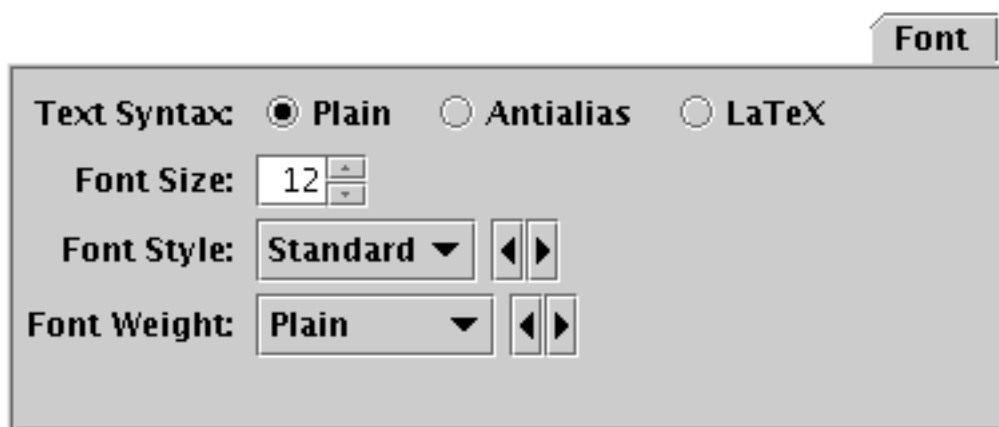
If set, minor (unlabelled) tick marks will be drawn between the major (labelled) ones.

#### Tick Crowding

Use the slider to influence how many tick marks are draw on the axes.

#### Antialiasing

Controls whether grid lines will be drawn antialiased (smoothed) or not. This option does not affect exported plots.



### Font tab

The **Font** tab configures the font used for axis annotation. It also affects some other things like the legend.

#### Text Syntax

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at

face value, but Antialias smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

### Font Size

Size of the font in points.

### Font Style

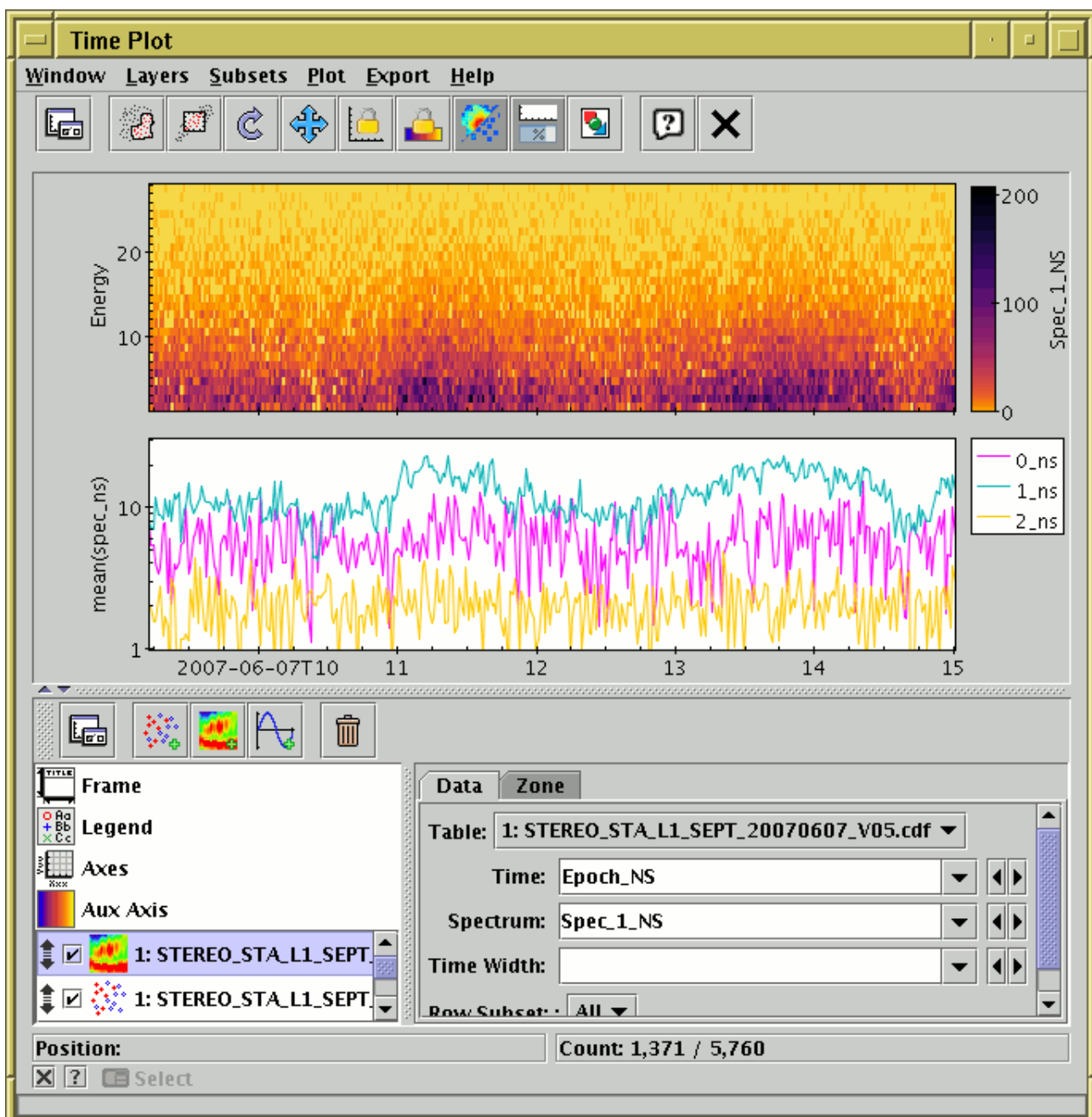
Style of the font.


### Font Weight

Whether the font is plain, bold or italic.

## A.4.12 Time Plot Window

*This plot window is experimental and lacks some useful features. The details of the user interface may be changed in future versions.*











The **Time Plot** () is intended for plotting time series data. You can find it in the Graphics menu of the main Control Window, but since it is currently experimental it does not appear in the main toolbar.

The horizontal axis represents time, and can be labelled accordingly, and the window can display functions of time, scatter plots and spectrograms.

Unlike the other plot windows (at time of writing) the Time plot can display different data plots in different plot Zones (Appendix A.4.12.1) stacked vertically on top of each other, so that different plots share a time axis but have their own Y axis.

The Time Plot offers the following plot controls:

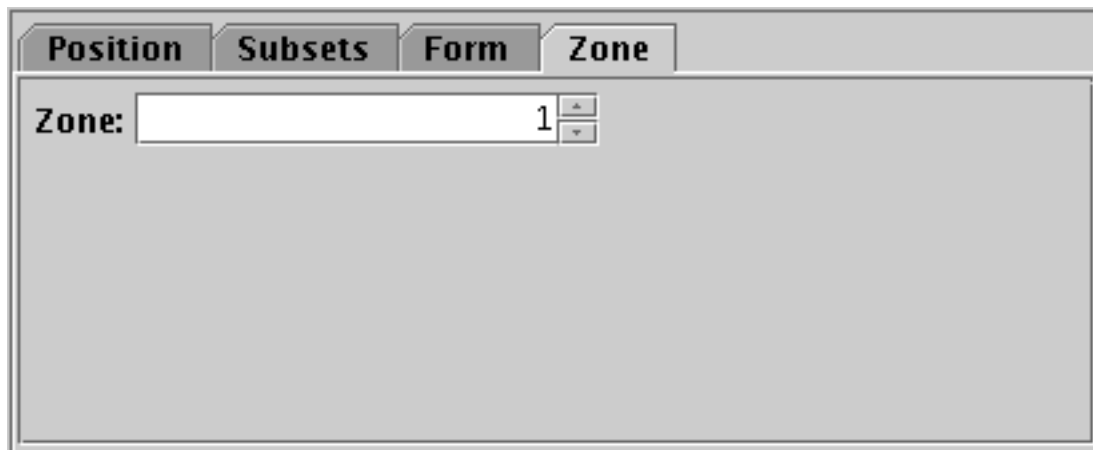
-  Position Layer Control, with these form options:
  -  Line
  -  Mark
  -  Error Bars (vertical bars only)
  -  Text Label
  -  Fill
-  Spectrogram Layer Control
-  Function Layer Control

See the Window Overview (Appendix A.4.2) for features common to all plotting windows. The following subsections describe zones, navigation and axis configuration.



#### A.4.12.1 Plot Zones

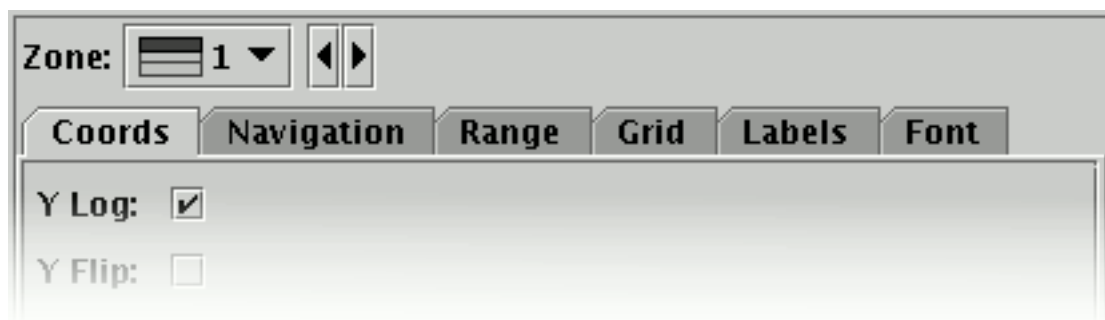
Unlike the other plot types, the Time plot can display data in multiple panels, known as **Zones**, stacked vertically. All plots therefore share the same time axis, but can have different Y axes.

By default, each new plot added is displayed in a new zone stacked beneath the existing ones. By using the **Zone** tab for each plot however, you can control which plots appear in which zones. Each zone has a numeric identifier, which by default increments by one for each new plot, but if you select the same identifier for several plots, they will all appear in the same zone.



### Zone selection tab

Some of the fixed controls (Axes  and Aux ) operate on a per-zone basis. If multiple zones are visible, then a zone selector is displayed above the tabs:



### Zone selector for Axes Control

In this case you should select the zone you want to control and adjust the configuration for that zone only. The selector shows an icon indicating the position of the zone in question. Each zone has to be configured separately. In a future release a global zone configuration option may be introduced as well.

This multi-zone feature is currently available only for the Time plot, but it may be added to other plot types at some point in the future.

### A.4.12.2 Time Navigation

For general comments on plot navigation, see Appendix A.4.2.1. Additional configuration options are available in the **Navigation** tab of the Axes control.

Horizontal navigation (panning and zooming left and right along the shared time axis) affects all the stacked zones, but vertical navigation (panning and zooming along the Y axis of each zone) is done independently for each zone.

The navigation actions for this window are:

#### Left drag

*Pan.* On the body of the plot this usually moves it around left/right. By default, vertical movement is inhibited since you don't normally want it for a time plot. To pan vertically, drag on the left of the Y axis for the zone you want to affect. It is also possible to configure panning on the body of the plot to be in both directions using the **Pan/Zoom Axes** in the **Navigation** axis configuration tab.

**☐ Right drag (CTRL-drag)**

*Stretch zoom.* On the body of the plot this usually stretches/squashes the plot centered on the horizontal position of the mouse. By default, vertical zooming is inhibited since you don't normally want it for a time plot. To zoom vertically, drag on the left of the Y axis for the zone you want to affect. It is also possible to configure zooming on the body of the plot to be in both directions using the **Pan/Zoom Axes** in the **Navigation** axis configuration tab.

**☐ Center drag (SHIFT-drag)**

*Frame zoom.* On the body of the plot, dragging right usually drags out a frame; when the button is released, the plot will be zoomed in to cover the area enclosed by the frame. Dragging left does something like the opposite, you can zoom out using a similar (though not quite the same) mechanism. By default, vertical zooming is inhibited since you don't normally want it for a time plot. To zoom in/out vertically, drag up/down on the left of the Y axis for the zone you want to affect. It is also possible to configure zooming on the body of the plot to be in both directions using the **Pan/Zoom Axes** in the **Navigation** axis configuration tab.

**☐ Wheel**

*Zoom.* Spinning the mouse wheel forwards/backwards will zoom horizontally in/out around the mouse position, just like dragging with the right button up-and-right or down-and-left.


**☐ Left click**

*Select.* If there is a plotted point near the cursor, it will plot a marker on it and activate (Section 8) it.

You can also manually fix the plot bounds using the **Range** tab of the Axes control.

**A.4.12.3 Time Axes Control**

**Note:** the Time Plot has multiple plot Zones (Appendix A.4.12.1), and the axes are configured individually for each zone.

The **Axes** () control for the time plot window has the following tabs:

**Coords tab of time Axes control**

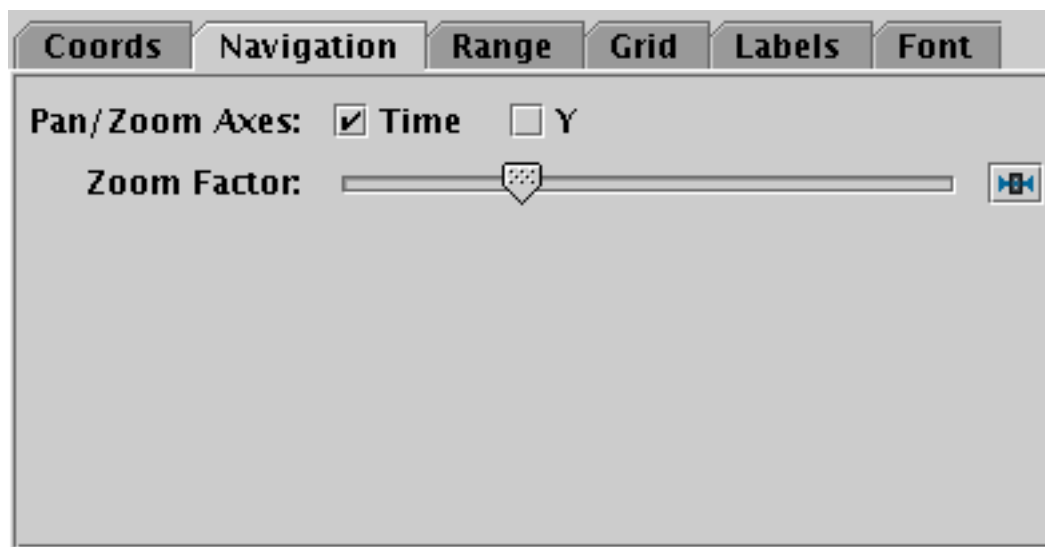
The **Coords** tab controls the vertical axis coordinates (you can't flip or rescale the time axis). It has the following options:

**Y Log**

If selected, vertical axis coordinates are logarithmic, otherwise they are linear.

### Y Flip

If selected, vertical axis coordinate axes increase down rather than up.



### Navigation tab of time Axes control

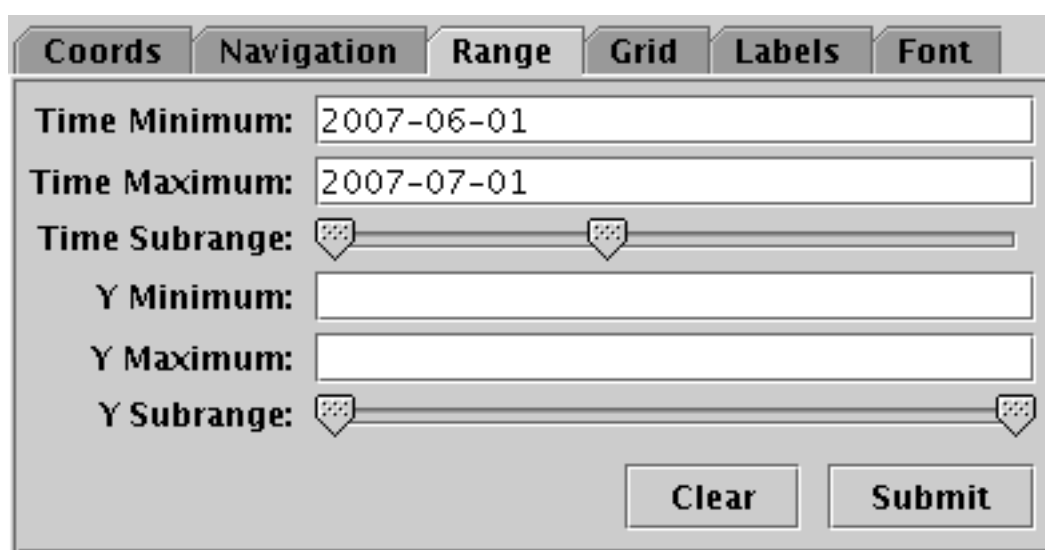
The **Navigation** tab controls details of how the navigation works. It has the following options:

#### Pan/Zoom Axes

By default, dragging with the mouse or using the mouse wheel on the body of the plot will pan or zoom in the horizontal (time) direction only, which is usually what you want for a time series. Using this control you can make it work in the vertical direction as well.

#### Zoom Factor

Controls the factor by which each zoom action zooms the plot. Moving this slider to the left/right makes the mouse more/less sensitive (one wheel click or dragging a fixed distance has more/less zoom effect).



### Range tab of time Axes control

The **Range** tab provides manual configuration of the visible range of the plot. Making changes to this tab will reset the visible plot range, but not vice versa - zooming and panning in the usual way will not change the settings of this panel.



Filling in the **Minimum/Maximum** fields for either or both axes will constrain the corresponding range of the visible data. The limits corresponding to any of those fields that are left blank will initially be worked out from the data. The **Subrange** double-sliders restrict the ranges within the (explicit or automatic) min/max ranges. Note you can move both sliders at once by grabbing a position between the two. For the time axis, the range may be entered as an ISO-8601 date/time value.

The **Clear** button resets all the fields. The **Submit** button updates the plot with the current values in this tab, as does making any changes to it.

The image shows a software interface with six tabs: 'Coords', 'Navigation', 'Range', 'Grid', 'Labels', and 'Font'. The 'Grid' tab is currently selected. Inside the 'Grid' tab, there are several controls:
 

- Time Format:** A dropdown menu showing 'ISO-8601' with left and right arrow buttons next to it.
- Draw Grid:** A checkbox that is checked.
- Minor Ticks:** A checkbox that is checked.
- Time Tick Crowding:** A horizontal slider with a central handle and a small icon on the right.
- Y Tick Crowding:** A horizontal slider with a central handle and a small icon on the right.

### Grid tab of time Axes control

The **Grid** tab configures the appearance of the axis grid. It has the following options:

#### Time Format

Selects the representation for time/date in which the horizontal axis is labelled. Options are ISO-8601, decimal year, Modified Julian Day, and Unix (seconds since midnight on 1 Jan 1970).

#### Draw Grid

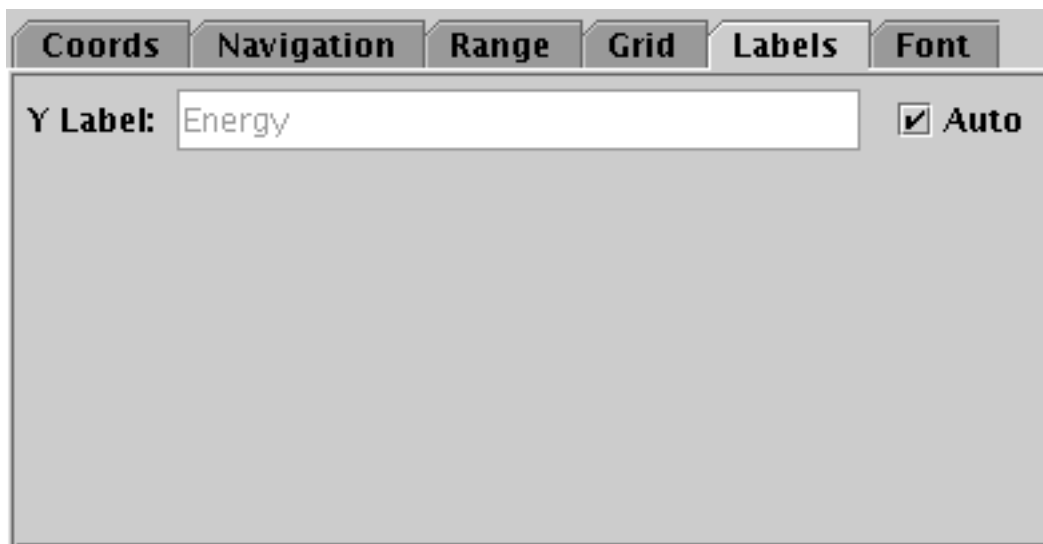
If true, grid lines will be drawn across the plot for every tick mark.

#### Minor Ticks

If set, minor (unlabelled) tick marks will be drawn between the major (labelled) ones.

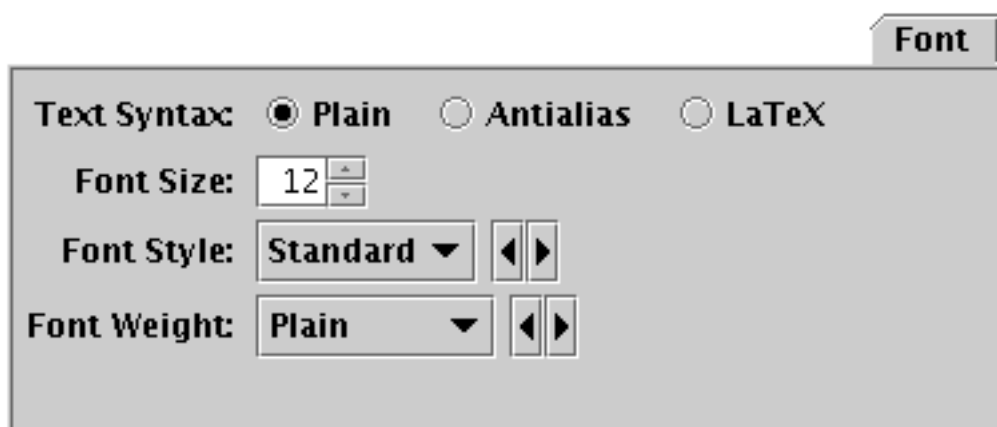
#### Time/Y Tick Crowding

Use the slider to influence how many tick marks are drawn on each axis.



### Labels tab of time Axes control

The **Labels** tab controls the text label on the Y axis (the horizontal axis is not labelled). If the **Auto** checkbox is set, the text will be taken from one of the data coordinates being plotted on that axis. To override that with your own axis label, unset **Auto** and type text in to the **Y Label** field.



### Font tab

The **Font** tab configures the font used for axis annotation. It also affects some other things like the legend.

#### Text Syntax

How to turn the text into characters on the screen. **Plain** and **Antialias** both take the text at face value, but **Antialias** smooths the characters. Antialiased text usually looks nicer, but can be perceptibly slower to plot. At time of writing, on MacOS antialiased text seems to be required to stop the writing coming out upside-down for non-horizontal text. **LaTeX** interprets the text as LaTeX source code and typesets it accordingly.

#### Font Size

Size of the font in points.

#### Font Style

Style of the font.

#### Font Weight

Whether the font is plain, bold or italic.

## A.5 Old-Style Plot Windows

*This section describes the old-style plotting windows used as standard in TOPCAT versions 2 and 3. Since version 4 (2013), the visualisation has been rewritten, and the new standard plotting windows are described in Appendix A.4. The windows described in this section are mildly deprecated, and will not be developed further. Most of their capabilities are handled better by the new-style plotting windows. However, these ones are still functional and if you want to use them you can find them in the **Graphics** menu of the Control Window, below the new-style plot options.*

Common features of the old-style plotting windows are described in the first subsection below; the specific windows themselves are described in the later subsections.

### A.5.1 Common Features

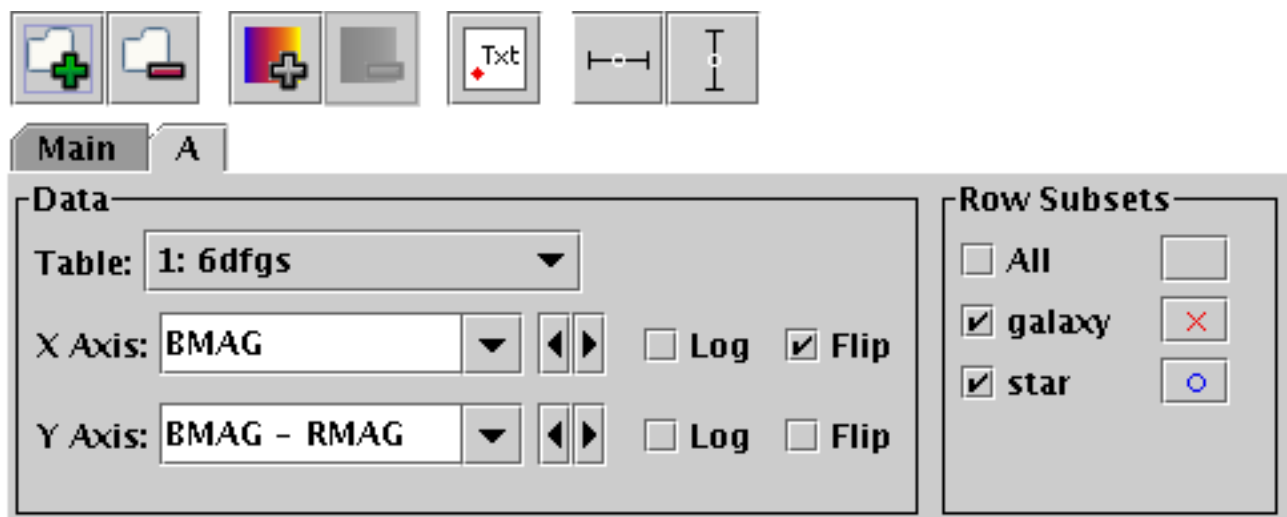
The various types of old-style plotting windows have different characteristics to fulfil their different functions, but they share a common way of doing things. Each window contains a number of controls including toolbar buttons, menu items, column selectors and others. In general any change that you make to any of the controls will cause the plot to be redrawn to take account of that change. If this requires a read or re-read of the data, a progress bar at the bottom of the window may show this progressing - except for very large tables it is usually pretty fast.

Each of the graphics windows is displayed by clicking its menu item in the **Graphics** menu of the Control Window. If one of the tables in the list is selected at the time (the Current Table) the new plot window will initially be displayed showing data from some of its columns (generally the first few numeric columns) by way of illustration. You will usually want to change the controls so it displays the quantities you are interested in.

The following subsections describe some of the features which work the same for most or all of the old-style graphics windows.

#### A.5.1.1 Dataset Selectors

All the old-style graphics windows provide one or more axes on which to plot - the histogram has 1, the 2d scatter and density plots have 2, the 3d scatter plot has 3 and the spherical plot has 2 or 3. In each case you select one or more dataset to plot on these axes, and select what plotting style to use for each set. A dataset is typically a number of columns from a table (the number matching the dimensionality of the plot) and a selection of row subsets associated with that table. You select this and the plotting style(s) using the panel at the bottom of each plot window. Here is dataset selector for the 2d scatter plot:



Default dataset selector from 2d scatter plot window

The different parts of this control work as follows:

### Data panel

The **Table** selector gives the identifier of the table (one of the ones loaded into TOPCAT) that the data comes from.

The **Axis** selectors (here **X Axis** and **Y Axis**) give the quantities to be plotted. If you click the little down arrow at the right of each selector you get a list of all the numeric columns in the chosen table, from which you can select one. If you click the little left and right arrows to the right of the selector it will cycle through all the columns in the table. However, if you prefer you can type in an expression to use here. This may be more convenient if there's a very long list of columns (another way to deal with this is to hide most of the columns using the Column Window). However, what you type in doesn't have to be a column name, it can be an algebraic expression based on columns in the table, or even a constant. In the example, the X axis is a straight column name, and the Y axis is an expression. The expression language syntax is explained in Section 7.

The **Log** checkbox for each axis is used to select whether the scale should be logarithmic or linear.

The **Flip** checkbox for each axis is used to select whether the axis values increase in the conventional direction (left to right, bottom to top) or its opposite.

Some of the buttons in the toolbar shown will modify what is visible in this panel, for instance inserting new selectors to allow selection of error values. All the selectors work in the same way however.

### Row Subsets panel

This defines which row subsets will be plotted in this window, and what plotting style should be used for them. In this case there are three defined subsets, All, galaxy and star. The checkboxes on the left indicate which ones will be displayed - here, only the latter two. Sets of points are generally plotted in the order they are selected for viewing; since points plotted afterwards can obscure ones plotted before ("underneath") this makes a difference. If you want to see a set of points without it being obscured by other ones in the plot, then deselect it and reselect it again (clicking twice in the corresponding checkbox), and this will ensure that its points are plotted on top.

The buttons to the right of each subset name show the symbol that is used in the plot to display the data from that subset, in this case a red cross and a blue circle. These are selected automatically when the subset is first selected for viewing (the initial default style set depends mainly on how many rows there are in the selected table - many rows gives small dots, few gives big ones). However, you have a lot of freedom to configure how they appear. If you click the button with the symbol on it a dialogue will pop up which allows you to select colour, shape, transparency and so on, as well as error bar style if appropriate and things like whether fitted lines will be plotted for that subset. The options available differ according to the kind of plot, and are described along with the different graphics windows in the following subsections. The style window stays visible until you dismiss it, but if you click on another of the buttons in the Row Subsets panel its contents will change to allow you to select values for the corresponding subset. Most graphics windows have a **Marker Style** menu. This allows you to change all the styles in the plot at once to be members of a uniform set, for instance different coloured pixels, or black open shapes. If you select one of these it will overwrite the existing style selections, but they can be edited individually afterwards.



### Dataset Tool Bar

The toolbar shown above the data panel in the figure contains buttons which affect the dataset selector itself. The first two buttons add and remove dataset **Tabs** (see below) and are present for all plots. The other items configure optional selectors appearing in the **Data Panel** - the ones shown here are concerned with Auxiliary Axes (Appendix A.5.1.5), Point Labels (Appendix A.5.1.4) and Error Bars (Appendix A.5.1.3), but not all the types of plot have

exactly the same ones.

### Tabs

The example shows two tabs: **Main** and **A**; the currently visible one is **A**. You can select a tab by clicking on its name. In each tab you select a table and a set of columns/expressions, and if they are all filled in it will contribute points (or bars, or whatever) to the plot. The **Main** dataset determines the initial values for the axis labels, but the data comes equally from all of them. The numerical values of the coordinates are treated the same for all the datasets, but their meanings might be different, for instance one dataset may be plotting *V* magnitude against ellipticity and another plotting *B* magnitude against ellipticity.

The **Add Dataset** () and **Remove Dataset** () buttons in the toolbar add a new tab or remove the selected one respectively. Initially only the **Main** tab is present, and this one cannot be removed.

Sometimes (high-dimensional plots, auxiliary axes, error bars) a lot of information needs to be entered into the data panel, and the bottom part of the window can get quite large. Normally, the plot in the upper part of the window shrinks to accommodate it. You can of course resize the window to gain more space, but if your screen is small you may still end up with an uncomfortably small plot. If this happens, you can use the following button from the main toolbar:



### Split Window

When this toggle button is on, the dataset selector can be resized by dragging the bar between it and the plot itself up or down. If there is insufficient space for all the components in the selector, a scrollbar will appear. When it is off (the default), the full height of the selector will be visible, and the plot will shrink to accommodate it.

## A.5.1.2 Axis Configuration and Zooming

In general terms the axes on which the graphics are plotted are defined by the datasets you have selected. The axis labels are set from the column names or expressions chosen for the **Main** dataset, and the ranges are determined so that all the data from the chosen datasets can be seen. However, these things can be adjusted manually.

The following features are available directly from the window for configuring axis range:

### X-Y Zoom

In some of the windows (2d scatter plot, histogram and density map), you can change both axis ranges by zooming in or out with the mouse on the plot surface itself. To zoom in, place the mouse at the top left of the region you want to examine, press the button, drag it to the bottom right corner, and release the button. To zoom out, drag up and left instead. A box is drawn as you drag so you can see what you're doing.

### Centre Zoom

The 3d and spherical plots allow you to zoom in on the central part of the window. The 'active region' for dragging is to the left or right of the plot (the region on the right is rather thin, and does not include the width used by the legend). When the pointer is in these regions, the mouse cursor symbol should change to indicate that zooming can be done. Drag down to zoom in and up to zoom out.

An easier alternative for zooming in the 3D windows is to use the mouse wheel, if you have one: wheel forward to zoom in and backward to zoom out.

### Axis Zoom

In some of the windows (2d scatter plot, histogram, density map, stacked lines) you can



modify the range on each axis independently by dragging the mouse over where the axis is drawn. The 'active region' for dragging is just below the X axis and just to the left of the Y axis, in the region where the numeric and text labels are written. When the pointer is in one of these regions, the mouse cursor symbol should change to indicate that zooming can be done. As for the X-Y Zoom, drag left-to-right or up-to-down to zoom in and right-to-left or down-to-up to zoom out.


### Auxiliary (Colour) Axis Zoom

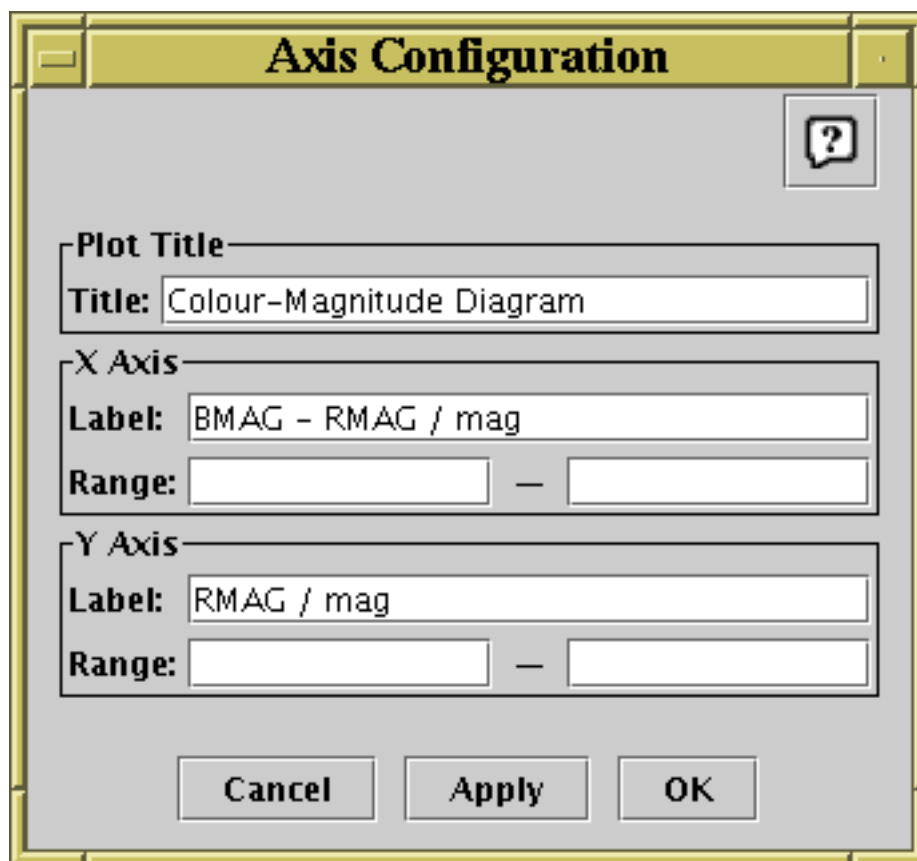
When Auxiliary Axes are in use, you can zoom in and out of them by dragging up and down on the colour bar to the right of the plot, in the same way as for a normal Axis Zoom above.



### Rescale

If you find you're zoomed to a region you don't want to be in, you can use the Rescale toolbar button to return to the default scale (full coverage). Note this affects any auxiliary axes as well as the spatial ones. Some windows may have per-axis rescale buttons too (, ).

For more control over axis range and labelling, use the **Configure Axes and Title** () toolbar button, which will pop up a dialogue like the following:



The dialog box is titled "Axis Configuration" and has a yellow title bar. It contains the following fields and buttons:

- Plot Title:** A text box containing "Colour-Magnitude Diagram".
- X Axis:**
  - Label:** A text box containing "BMAG - RMAG / mag".
  - Range:** Two empty text boxes separated by a minus sign.
- Y Axis:**
  - Label:** A text box containing "RMAG / mag".
  - Range:** Two empty text boxes separated by a minus sign.
- Buttons:** "Cancel", "Apply", and "OK" at the bottom.

### Axis Configuration Dialogue for 2-d axes

You can fill in these values for each axis as follows:

#### Label

For each axis the label box contains the text used to annotate the axis in the plot. By default this is the same as the text in the **Main** dataset column selector (usually a column name), followed by the units if known. However, you can change it by typing whatever text you like.

#### Range

The range boxes allow you to specify the lower and upper limits of each axis. By default these are blank, meaning that the plot will size its axes so that all the data can be seen. However, if you fill in one or both of the boxes with a suitable numeric value, the lower/upper bound will be fixed at that. Note that the lower bound (left box) must be numerically less than the upper bound (right box).

Both values are reset if the plot's axis is changed (a new column or expression is selected for the Main dataset), or if the range is reset in some other way (e.g. by zooming).

The plot title may also be set in the **Plot Title** panel of this window:

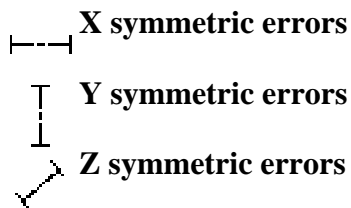
#### **Title**

Any text entered here will be displayed at the top of the plot to provide a title.

### **A.5.1.3 Error Bars**

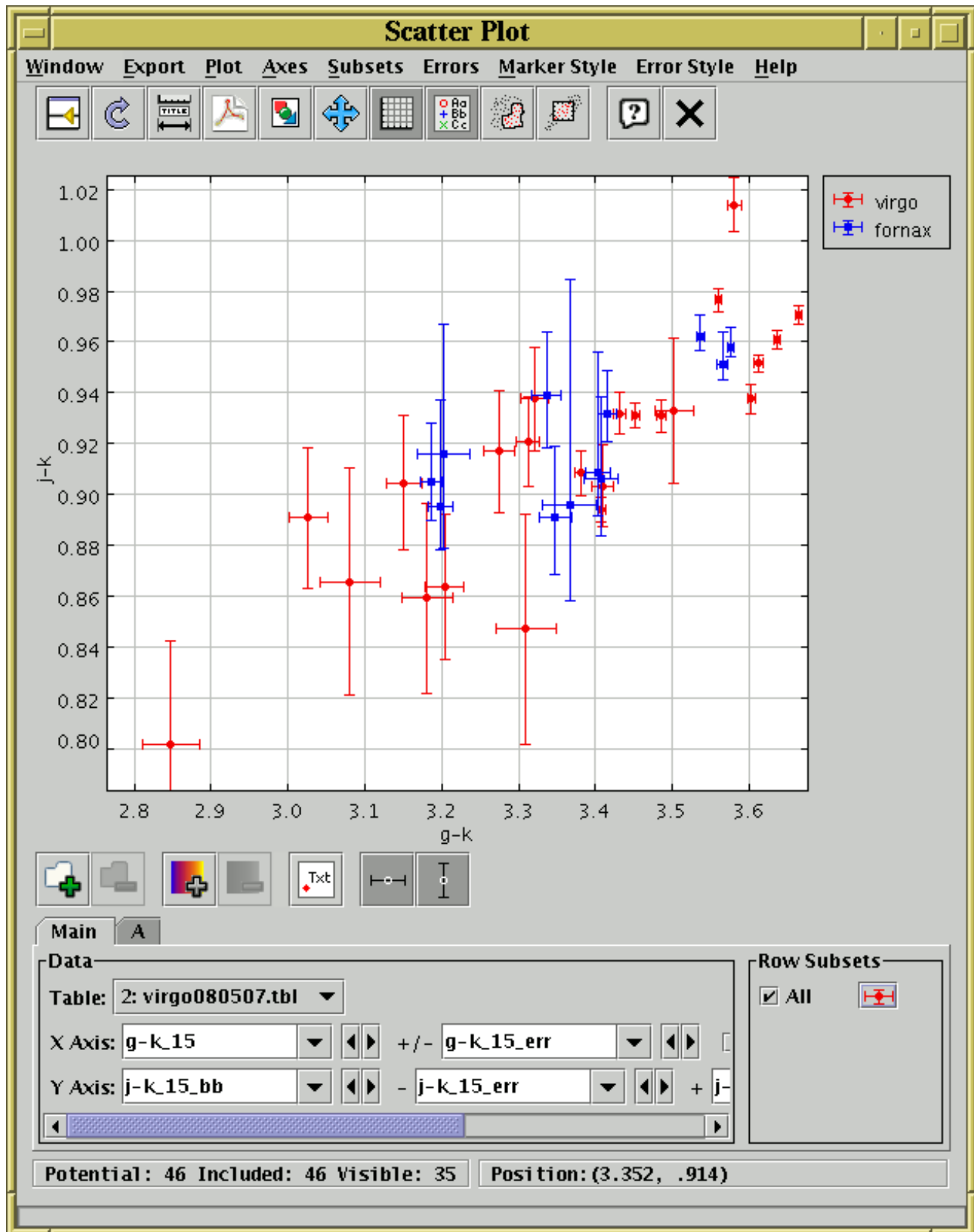
TOPCAT provides quite flexible graphical representation of symmetric or asymmetric errors in 1, 2 and 3 dimensions. The plots with error bar support are the 2D, 3D and spherical scatter plots and the stacked lines plot.

By default, error bar drawing is switched off. The simplest way to activate it is to use the relevant error bar button(s) in the data selector tool bar (the one below the plot). For the Cartesian (2D, 3D, lines) plots, some or all of the following buttons are present:



Any combination of them can be active at once. Clicking one of these buttons toggles symmetric error bar drawing for that axis on or off. When it is on, an additional column selector will appear to the right of the main column selector for the axis in question. If you fill this in with a column name or expression which gives the error for that axis, then the error bars will be plotted. TOPCAT may make a guess based on column names and UCDs about which columns provide error values for which other columns, so the error selector may get filled in automatically. However, in most cases you will need to provide the error values by selecting a column yourself, and occasionally you may need to correct TOPCAT's guesses.

Here is a 2D plot in which symmetric X and asymmetric Y errors are being used:

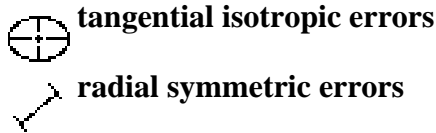


### Plot window with symmetric X and asymmetric Y errors

You can see that with the error column selector, the panel has become too wide for the window so a scrollbar has appeared at the bottom - you can scroll this left and right or enlarge the window to see the parts that you need to.

For the spherical plot the following error toggle buttons are present:





These work in a similar way to the Cartesian errors above, except that the tangential one adds a single column selector, with an associated unit selector, near the latitude and longitude selectors to determine the isotropic angular size of error small circles.


If you want to use **asymmetric** or one-sided errors, use the options in the **Errors** menu instead of the toolbar buttons. For instance the options for X axis error bars in the 2D scatter plot are:

**None**

 **Symmetric**

 **Lower Only**

 **Upper Only**

 **Lower & Upper**

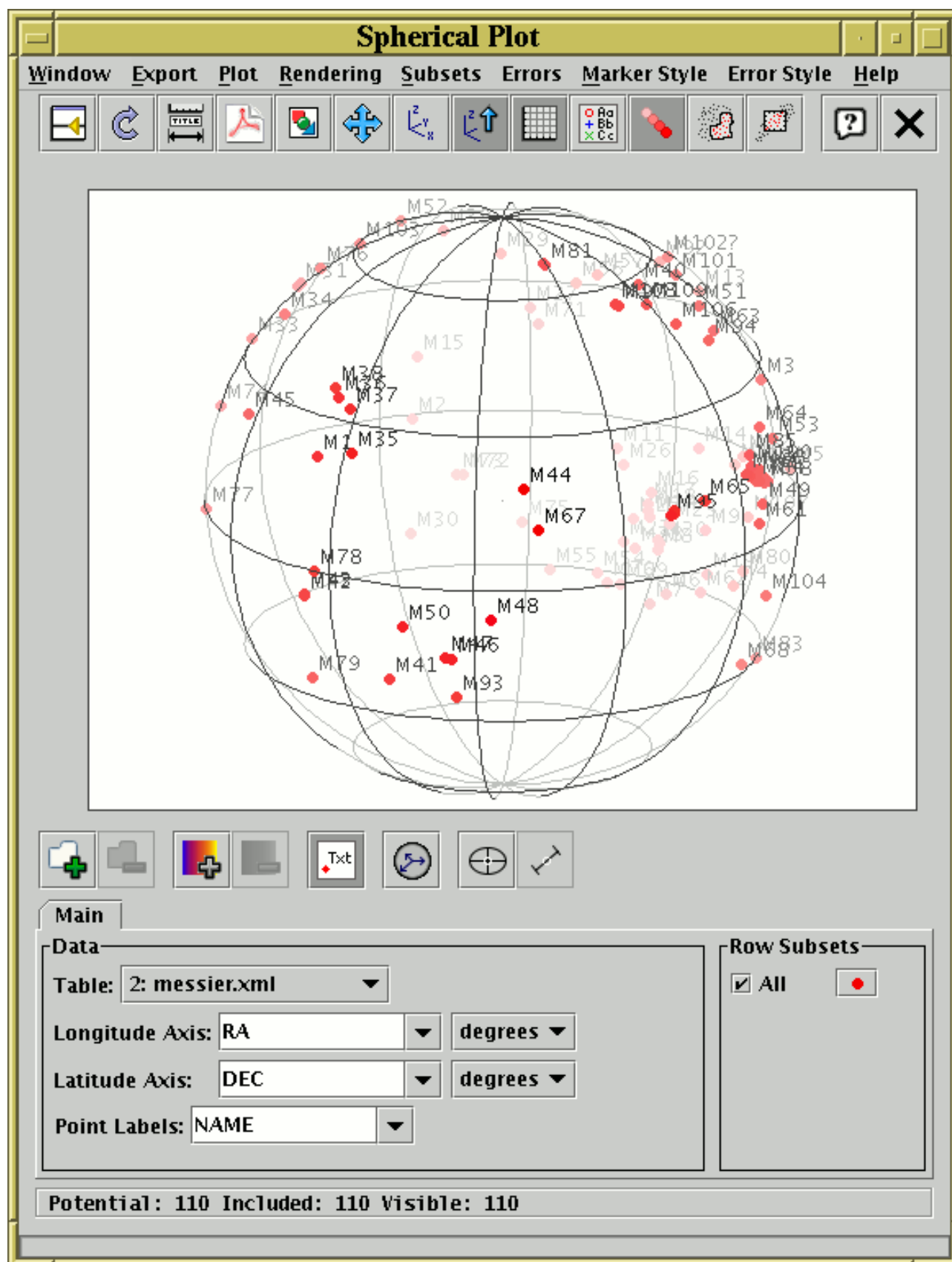
These give you different column selector boxes, but work in much the same way as the symmetric ones.

There are many options for the plotting style of one, two and three dimensional error bars, including capped and uncapped bars, crosshairs, ellipses and rectangles. This plotting style is controlled from the plot window's **Style Editor** window (see e.g. Appendix A.5.3.1), which can be viewed by clicking on the marker icon in the Row Subsets panel at the bottom right of the window. The available error bar styles will depend on which axes currently have errors; if none do, then the error bar selector will be disabled. You can also use the **Error Style** menu to change the error style for all the visible datasets at once.

#### A.5.1.4 Point Labels

On the 2-d and 3-d scatter plots you can write text labels adjacent to plotted points. To do this click the **Draw Labels** () button in the dataset toolbar (below the plotting area in the plot window).

This will reveal a new **Point Labels** selector below the existing spatial ones. Using this you can select any of the table columns (not just the numeric ones as for the other selectors), or give a string or numeric expression involving them. When this selector is filled in, every point in the dataset which has a non-blank value for this quantity will have it written next to the point on the display.



### Point Labelling for Messier objects in the spherical plot


In this example the NAME column has been selected, so that each point plotted (in this case all the Messier objects) is labelled with its name. As you can see, where many labels are plotted near to each other they can get in each others' way. In some cases TOPCAT will omit plotting labels in crowded regions, in others not - but in any case if you have labels too tightly grouped they are

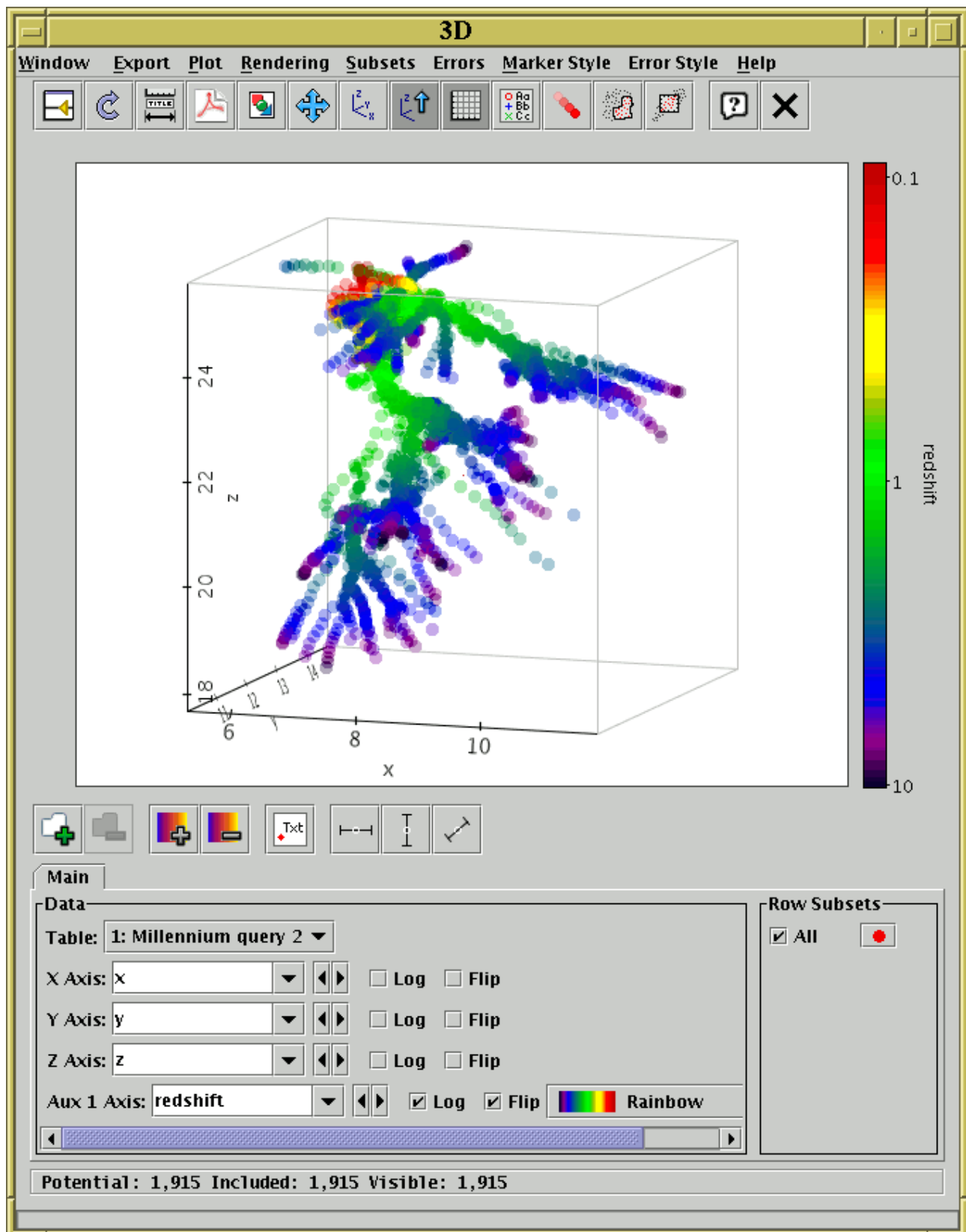
unlikely to be legible.

### A.5.1.5 Auxiliary Axes

TOPCAT can plot data in one, two or three spatial dimensions, but sometimes the data which you need to visualise is of higher dimensionality. For this purpose, some of the plotting windows (2D and 3D scatter plots) allow you to control the colouring of plotted points according to values from one or more additional columns (or calculated expressions), which gives you more visual information about the data you are examining.

To use this facility, click the **Add auxiliary axis** () button in the dataset toolbar (below the plot

area in a plot window). A new axis selector will appear below the existing spatial ones, labelled **Aux 1 Axis**. It has log and flip checkboxes like the spatial axes, and to the right (you may need to widen the window or use the scrollbar at the bottom to see it) is a selector depicting a number of colourmaps to choose from - the default one resembling a rainbow is usually quite suitable, but you can pick others. If you enter a column name or expression into the selector, each plotted point will be coloured according to the value of that quantity in the corresponding row of data. If that quantity is null for a row, the corresponding point will not be plotted. A scale on the right of the plot indicates how the colour map corresponds to numeric values. To remove the auxiliary axis and go back to normally-coloured points, simply click the **Remove auxiliary axis** () button.



**3D plot of simulation data showing X, Y, Z spatial position with the auxiliary axis indicating timestep.**

There are two types of colour maps you can choose from: **colour fixing** and **colour modifying**. The fixing ones are easiest to understand: the original colour of the point (as drawn in the legend) is ignored, and it is coloured according to the relevant value on the selected auxiliary axis. The colour modifying maps take the original colour and affect it somehow, for instance by changing its transparency or its blue component. These are marked with an asterisk ("\*") in the colour map

selector. They can be used to convey more information but are often harder to interpret visually - for one thing the shading of the colour bar in the legend will not correspond exactly to the colours of the plotted points.




By using modifying colour maps it is possible to perform plots with more than one auxiliary axis - typically the first one will be a fixing map and subsequent ones will be modifying. So the first auxiliary axis could have the (fixing) Rainbow map, and the second could have the (modifying) Transparency map. The colour alterations are applied in order. It is possible, but pointless, to have multiple fixing maps applied to the same points - the last-numbered one will determine the colour and earlier ones will get ignored. Multiple aux axes can be obtained by clicking the **Add auxiliary axis** button more than once. When combining several maps some thought has to be given to which ones to use - some good combinations are the three RGB ones or the three YUV ones.

A fairly wide range of colour maps of both kinds is provided by default. If these do not suit your needs, it is possible to provide your own custom colour fixing maps using the **lut.files** system property - see Section 10.2.3.

It is easy to generate attractive screenshots using auxiliary axes. Making visual sense of the results is a different matter. One visualisation expert tried to dissuade their introduction in TOPCAT on the grounds that the graphics they produce are too hard for humans to interpret - I hope that these plots can assist with some analysis, but it is a somewhat experimental feature which may or may not end up being widely useful. The maximum number of auxiliary axes which can be used together is currently three. This could be increased on request, but if you feel you can generate an intelligible plot using more than this then you're considerably smarter than me.


#### A.5.1.6 Defining Subsets by Region

When quantities are plotted in one of the graphics windows it becomes easy to see groupings of the data which might not otherwise be apparent; a cluster of (X,Y) points representing a group of rows may correspond to a physically meaningful grouping of objects which you would like to treat separately elsewhere in the program, for instance by calculating statistics on just these rows, writing them out to a new table, or plotting them in a different colour on graphs with different coordinates. This is easily accomplished by creating a new Row Subset containing the grouped points, and the graphics windows provide ways of doing this.

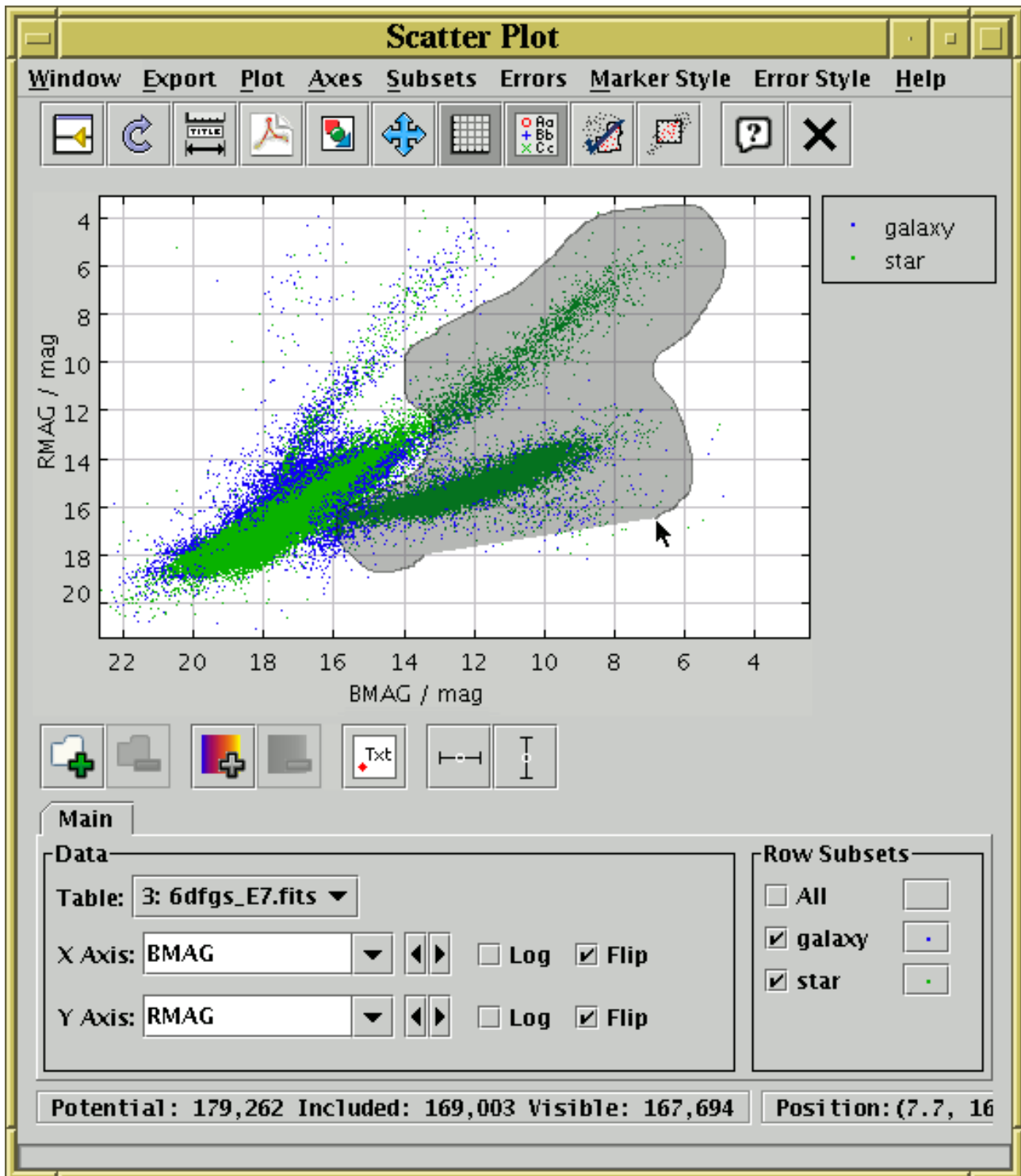
In some of the plots (Histogram 2d Scatter plot Density map and Spherical plot) you can set the axis ranges (either manually or by zooming with the mouse - see Appendix A.5.1.2) so that only the points you want to identify are visible, and then click the **New Subset From Visible** toolbar button (the icon is , , or  depending on the plot type). This defines a subset consisting of all

the points that are visible on the current plot. This is only useful if the group you are interested in corresponds to a rectangular region in the plotting space.


A more flexible way is to draw a region or regions on the plot which identify the points you are interested in. To do this, hit the **Draw Subset Region** () toolbar button. Having done this, you

can drag the mouse around on the plot (keep the left mouse button down while you move) to encircle the points that you're interested in. As you do so, a translucent grey blob will be left behind - anything inside the blob will end up in the subset. You can draw one or many blobs, which may be overlapping or not. If you make a mistake while drawing a sequence of blobs, you can click the right mouse button, and the most recently added blob will disappear. When you're in this region-drawing mode, you can't zoom or resize the window or change the characteristics of the plot, and the **Draw Subset Region** button appears with a tick over it () to remind you you're in it.

Here's what the plot looks like while you're drawing:



### Region-Drawing Mode

When you're happy with the region you've defined, click the  toolbar button again.

In either case, when you have indicated that you want to define a new row subset, a dialogue box will pop up to ask you its name. As described in Section 3.1.1, it's a good idea to use a name which is just composed of letters, numbers and underscores. You can optionally select a subset name which has been used before from the list, which will overwrite the former contents of that subset. When you enter a name and hit the **OK** button, the new subset will be created and the points in it will be shown straight away on the plot using a new symbol. As usual, you can toggle whether the points in this subset are displayed using the **Row Subsets** box at the bottom of the Plot Window.

### A.5.1.7 Exporting Graphics

All the old-style graphics windows have the following export options in the toolbar:



**Export as PDF**



**Export as GIF**

and additionally, the **Export** menu contains:



**Export as Encapsulated PostScript**



**Export as Gzipped Encapsulated PostScript**



**Export as JPEG**



**Export as PNG**

These can be used to export the currently visible plot to an external graphics format for later use.

Exporting to the pixel-based formats (GIF, JPEG, PNG) is fairly straightforward: each pixel on the screen appears as one pixel in the output file. PNG is the most capable, but it is not supported by all image viewers. GIF works well in most cases, but if there are more than 255 colours some of the colour resolution will be lost. JPEG can preserve a wide range of colours, but does not support transparency and is lossy, so close inspection of image features will reveal blurring.

When exporting to Portable Document Format (PDF) or Encapsulated PostScript (EPS), which are vector graphics formats, there are a few things to consider:

#### **Positional Quantisation**

The output file will render text, lines and markers properly, with smooth edges rather than steps where pixel boundaries would be on the screen. However, the *positional* resolution is the same as it would be on the screen, so if you have a 400-pixel high plot for instance, there are only 400 possible Y coordinates at which a marker can be plotted. In general this is not obvious by looking at the output plot, but you may find it helpful to increase the size of the plot on the screen by resizing the window before performing an export to EPS. This reduces the effect of the positional quantisation, but note it will also have the effect of making the text labels proportionally smaller to the graphics.

#### **Transparency**

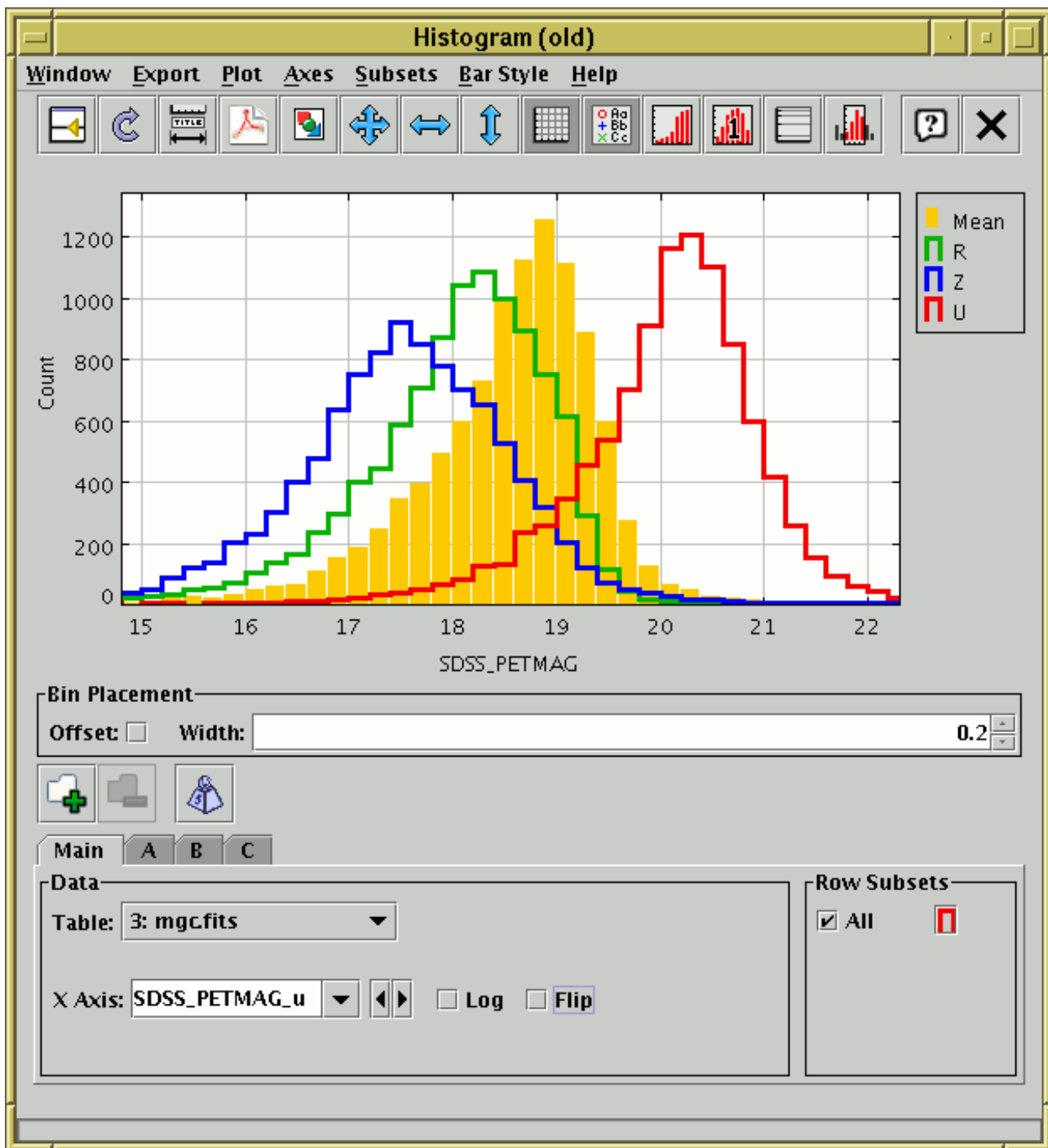
For technical reasons transparent markers cannot easily be rendered when a plot is exported to PostScript. In some cases the plot is done using a bitmap in the PostScript output to permit transparency and in some cases the points are just plotted opaque. Try it and see; if the points come out opaque instead of transparent you may need to export to GIF instead. Better workarounds may be provided in future releases.

#### **File Size**


In some cases (2D and 3D scatter plots with many thousands of points) output EPS files can get extremely large; the size scales with the number of points drawn, currently with a factor of a few hundred bytes per point. In some cases you can work round this by plotting some points as transparent so that the plot is rendered as a bitmap (see the discussion of transparency above) which scales as the number of pixels rather than the number of points. The Gzipped EPS format helps somewhat (though can be slow); PDF output is better still. Even PDF files may be unmanageably large for very many points however.

### A.5.2 Histogram (old-style)

This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4.



### Histogram Window

The histogram window lets you plot histograms of one or more columns or derived quantities. You can display it using the **Histogram** (  ) item in the Control Window's **Graphics** menu.

You select the quantity or quantities to plot using the dataset selector (Appendix A.5.1.1) at the bottom of the window. You can configure the axes, including **zooming** in and out, with the mouse (drag on the plot or the axes) or manually as described in Appendix A.5.1.2.



The **Bin Placement** box below the main plot controls where the bars are drawn. Select the horizontal range of each bar using the **Width** entry box - either type in the value you want or use the tiny up/down arrows at the right to increase/decrease the bin size. The **Offset** checkbox on the left determines where the zero point on the horizontal axis falls in relation to the bins; if the box is checked then zero (or one for logarithmic X axis) is in the centre of a bin, and if it's unchecked then zero (or one) is on a bin boundary.

The following buttons are available on the toolbar:



#### **Split Window**

Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



#### **Replot**

Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



#### **Configure Axes and Title**

Pops up a dialogue to allow manual configuration of axis ranges, axis labels and plot title - see Appendix A.5.1.2.



#### **Export as PDF**

Pops up a dialogue which will write the current plot as a PDF file.



#### **Export as GIF**

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



#### **Rescale**

Rescales the axes so that the width and height are sufficient to accommodate all the non-zero bars in the histogram for all the currently selected datasets. By default the plot will be scaled like this, but it may have changed because of changes in the subset selection or from zooming in or out.



#### **Rescale X**

Rescales the horizontal axis to accommodate all the currently plotted bars. The vertical axis scaling is not changed.



#### **Rescale Y**

Rescales the vertical axis to accommodate all the currently plotted bars. The horizontal axis scaling is not changed.



#### **Grid**

Toggles whether a grid is drawn over the plotting surface or not.



### Show Legend

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



### Cumulative Plot

Toggles whether the histogram should be normal or cumulative. Normally the height of each bar is determined by counting the number of points which fall into the range on the X axis that it covers. For a cumulative plot, the height is determined by counting all the points between negative infinity and the upper bound of the range on the X axis that it covers.



### Normalisation

Toggles whether histograms are normalised. When selected, each dataset will be normalised so that the sum of the counts of all its bars over the whole range of data is equal to one.



### Log Y Axis

Toggles whether the Y axis is scaled logarithmically or not.



### Subset From Visible

Defines a new **Row Subset** consisting of only the data in the bars which are visible in the current plot. See Appendix A.5.1.6 for more explanation.

The **Dataset Toolbar** contains the following options:



### Add/Remove dataset

Adds/removes tabs in which the data for extra datasets can be entered. See Appendix A.5.1.1.



### Weight Counts

If this toggle button is on, an additional **Weight Axis** selector appears below the **X Axis** selector. If this is filled in with a column name or expression, then instead of simply accumulating the number of points per bin, the Y axis will show the sum over the weighting quantity for points in each bin. Not having a weight axis is equivalent to filling in its value with the quantity 1. Note that with weighting, the figure drawn is no longer strictly speaking a histogram.

When weighted, bars can be of negative height. An anomaly of the plot as currently implemented is that the Y axis never descends below zero, so any such bars are currently invisible. This may be amended in a future release (contact the author to vote for such an amendment).

The **Export** menu contains additional image export options and the following options specific to the histogram:



### Save as Table

The bin counts/sums corresponding to the currently plotted histogram will be written to disk in tabular form. The first two columns give the lower and upper bounds of each bin, and the subsequent columns give the occupancies of each bin for each plotted data set. If only one dataset is plotted, there will only be three columns.



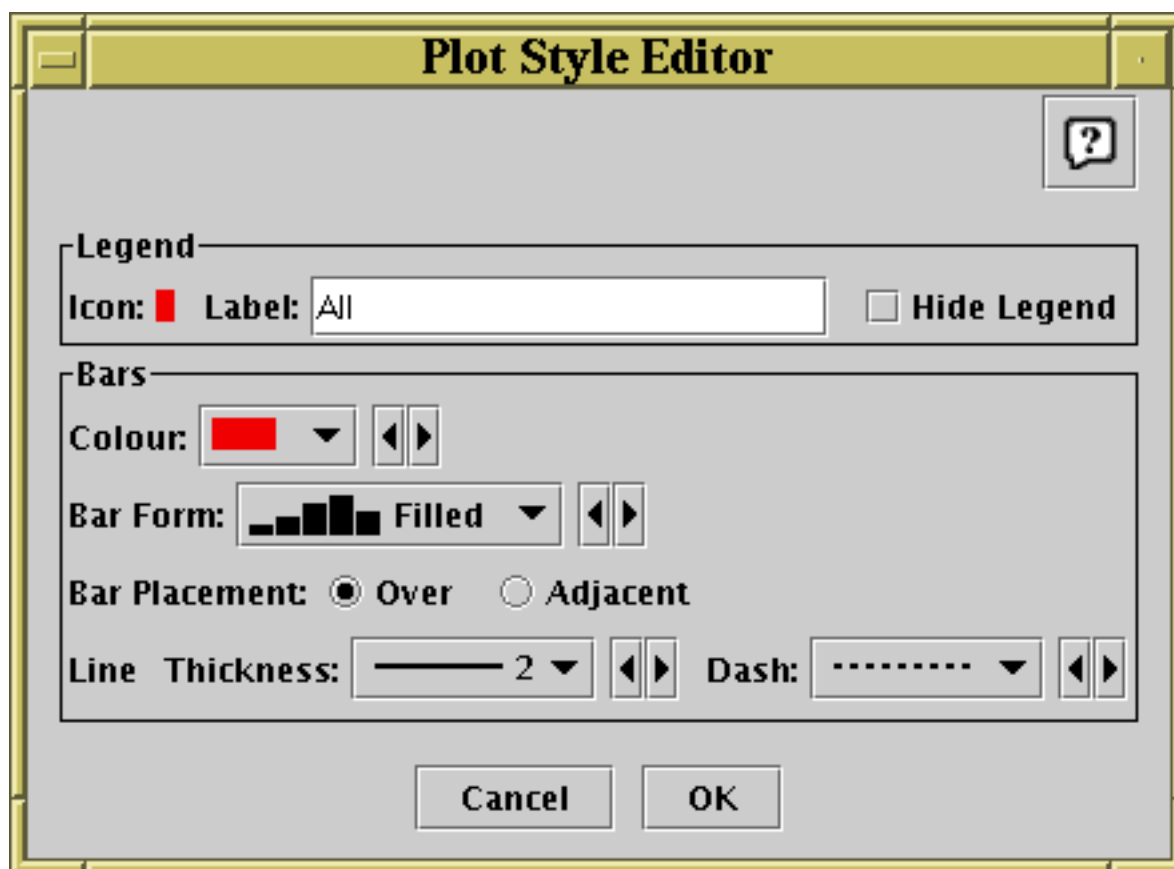
### Import as Table

Assembles a table as per the **Save** option above, but rather than writing it to disk imports it directly into TOPCAT, where it can be manipulated in all the usual ways.


You have considerable freedom to configure how the bars are drawn; controlling this is described in the following subsection.

### A.5.2.1 Histogram Style Editor

The bins in a histogram can be represented in many different ways. A representation of how a bar will be displayed is shown on a button to the right of the name of each visible subset, at the bottom right of the histogram window. If you click this button the following dialogue will pop up which enables you to change the appearance.



Style editor dialogue for histogram bars

The **Legend** box defines how the selected set will be identified in the legend which appears alongside the plot (though the legend will only be visible if **Show Legend**  is on):

#### Icon

Displays the icon which will be shown to identify the points in the selected set. Its appearance depends on the selections you make in the rest of this dialogue window.

#### Label

Gives the name written in the legend to label the subset. By default this is derived from the Row Subset's name and, if it's not part of the main dataset, the name of the dataset's tab. You can type in a new value to change what is written in the legend.

#### Hide Legend

If this checkbox is selected, then no entry for the selected set will appear in the legend.

The **Bins** panel describes the form of the bars to be plotted for each data set.

**Colour**

Selects the colour for the bar or line which will represent this set.

**Bar Form**

Selects the style of bar which will be plotted. Available styles are filled rectangles, open rectangles, stepped lines and spikes.

**Bar Placement**

Selects where each bar will be placed on the X axis within the ordinate region which it represents. There are currently two options: **Over** means that it covers the whole of its range, and **Adjacent** means that it covers only a proportion of its range, so that multiple datasets plotted on the same graph don't obscure each other (if 2 sets are plotted, the first one will take the left half and the second will take the right half of each bin region, and so on). In the case that there is only a single data set plotted, it doesn't matter which of these is chosen.

**Line Thickness**

Determines the thickness in pixels of any lines which are drawn. Only applies to those **Bar Forms** which use lines rather than solid blocks.

**Dash**

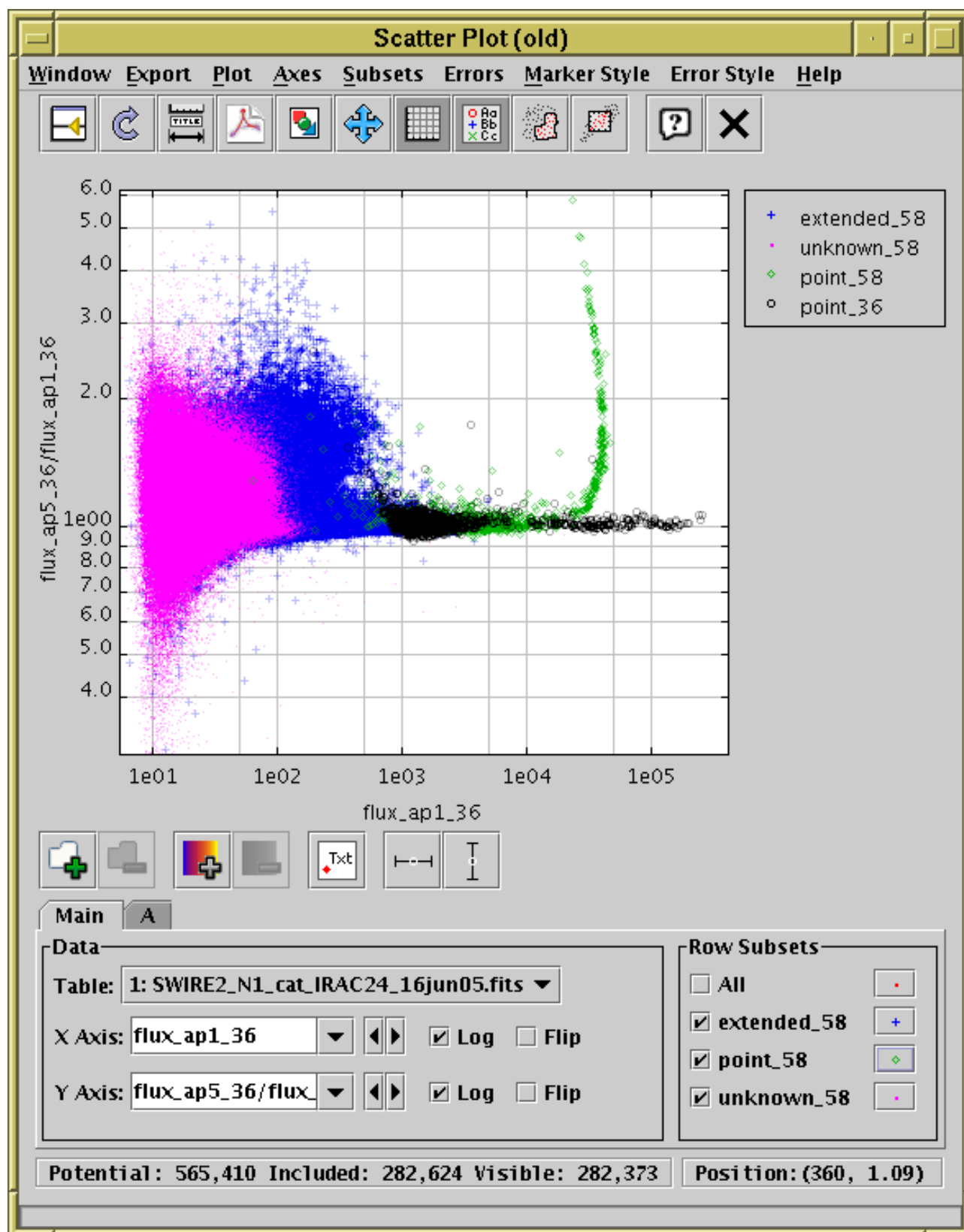
Determines the dash style (solid, dotted, dashed or dot-dashed) for any lines which are drawn. Only applies to those **Bar Forms** which use lines rather than solid blocks.

Any changes you make in this window are reflected in the plot straight away. If you click the **OK** button at the bottom, the window will disappear and the changes remain. If you click **Cancel** the window will disappear and any changes you made will be discarded.


You can also change all the plotting styles at once by using the **Bar Style** menu in the histogram window. Here you can select a standard group of styles (e.g. all filled adjacent bars with different colours) for the plotted sets.

### A.5.3 2D Plot (old-style)

*This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4.*



### Plot Window

The plot window allows you to do 2-dimensional scatter plots of one or more pair of table columns (or derived quantities). You can display it using the **Plot** (  ) item in the Control Window's

**Graphics** menu.

On the plotting surface a marker is plotted for each row in the selected dataset(s) at a position

determined by the values in the table columns selected to provide the X and Y values. A marker will only be plotted if both the X and Y values are not blank. Select the quantities to plot and the plotting symbols with the dataset selector (Appendix A.5.1.1) at the bottom. You can configure the axes, including **zooming** in and out, with the mouse (drag on the plot or the axes) or manually as described in Appendix A.5.1.2.

Clicking on any of the plotted points will **activate** it - see Section 8.

The following buttons are available on the toolbar:



#### Split Window

Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



#### Replot

Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



#### Configure Axes and Title

Pops up a dialogue to allow manual configuration of axis ranges, axis labels and plot title - see Appendix A.5.1.2.



#### Export as PDF

Pops up a dialogue which will write the current plot as a PDF file. In general this is a faithful and high quality rendering of what is displayed in the plot window. However, if plotting is being done using the transparent markers, the markers will be rendered as if they were opaque. Plots with very many points can result in rather large output PDFs.



#### Export as GIF

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



#### Rescale

Rescales the axes of the current plot so that it contains all the data points in the currently selected subsets. By default the plot will be scaled like this, but it may have changed because of changes in the subset selection or from zooming in or out.



#### Grid

Toggles whether a grid is drawn over the plotting surface or not.



#### Show Legend

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



#### Draw Subset Region

Allows you to draw a region on the screen defining a new Row Subset. When you have

finished drawing it, click this button again to indicate you're done. See Appendix A.5.1.6 for more details.



### Subset From Visible

Defines a new **Row Subset** consisting of only the points which are currently visible on the plotting surface. See Appendix A.5.1.6 for more explanation.

The **Dataset Toolbar** contains the following options:



### Add/Remove dataset

Adds/removes tabs in which the data for extra datasets can be entered. See Appendix A.5.1.1.



### Add/Remove auxiliary axis

Adds/removes a selector for entering an auxiliary axis to modify point colours etc. See Appendix A.5.1.5.



### Toggle point labelling

Allows text labels to be drawn near plotted points. See Appendix A.5.1.4.



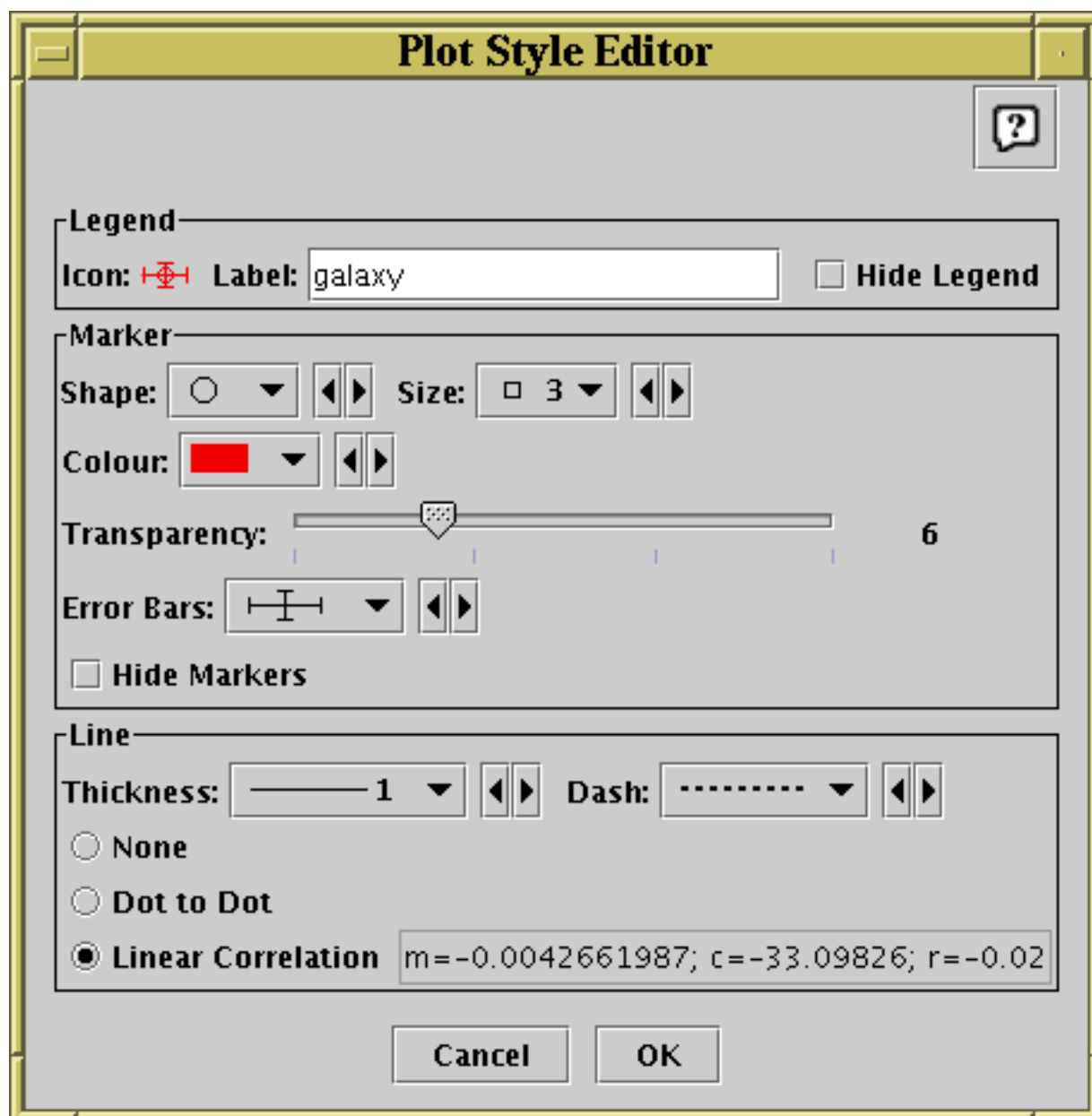
### Toggle X/Y error bars

Switches between drawing symmetric error bars and no error bars in the X and Y directions respectively. Other options are available in the **Errors** menu. See Appendix A.5.1.3.


You have considerable freedom to configure how the points are plotted including the shape, colour and transparency of symbols, the type of lines which join them if any, and the representation of error bars if active. These options are described in the following subsection.

## A.5.3.1 Plot Style Editor

When plotting points in a scatter plot there are many different ways that each point can be displayed. By default, TOPCAT chooses a set of markers on the basis of how many points there are in the table and uses a different one for each plotted set. The marker for each set is displayed in a button to the right of its name in the dataset selector panel at the bottom of the plot window. If you click this button the following dialogue will pop up which enables you to change the appearance.



### Style editor dialogue for 2d scatter plot

The **Legend** box defines how the selected set will be identified in the legend which appears alongside the plot (though the legend will only be visible if **Show Legend**  is on):

#### Icon

Displays the icon which will be shown to identify the points in the selected set. Its appearance depends on the selections you make in the rest of this dialogue window.

#### Label

Gives the name written in the legend to label the subset. By default this is derived from the Row Subset's name and, if it's not part of the main dataset, the name of the dataset's tab. You can type in a new value to change what is written in the legend.

#### Hide Legend

If this checkbox is selected, then no entry for the selected set will appear in the legend.

The **Marker** box defines how the markers plotted for each data point will appear:

#### Shape



Choose from a variety of shapes such as open or filled circles, squares, crosses etc.

### Size

Choose the size of the marker; the value given is approximate radius in pixels. If a size of zero is chosen, then the shape doesn't matter, the points will be plotted as single pixels.

### Colour

Choose the colour in which the markers, and any line if one is drawn, will be plotted.

### Transparency

Choose transparency of the plotted symbols. The scale on the slider is logarithmic, with 1 at the left hand end. The actual value chosen is an integer written at the right of the slider. This number gives the number of markers for this set which need to be plotted in the same position to result in fully opaque pixels - any fewer and the background, or other markers plotted underneath, will show through to some extent. Setting this to some value greater than 1 is very useful if you have a very large number of points being plotted (especially if it's comparable with the number of the pixels on the screen), since it enables you to distinguish regions where there are lots of points on top of each other from those where there are only a few.

### Error Bars

If error bars are active for this plot, allows you to select the way they will appear. The options which can be selected here will depend on whether X and/or Y errors are in use.

### Hide Markers

This check box is only enabled if a line and/or error bars are being plotted; it allows the markers to be invisible, so that only the line/error bars are seen.

The **Line** box determines if any lines are drawn associated with the current set and if so what their appearance will be.

### Thickness

Selects the line thickness in pixels.

### Dash

Selects a dash pattern (solid, dotted, dashed or dot-dashed) for the line.

### Type

The other radio buttons determine what kind of line, if any, will be plotted for these points. There are three options:

#### None

No line is drawn - this is the default.

#### Dot to Dot

A straight line segment is drawn between each of the points. If the points effectively form an ordered set of samples of a function, this will result in a more-or-less smooth drawing of that function on the plot. Note that the lines are drawn in the order that the points appear in the basic table, and if this doesn't match the 'ordinate' order the result will be a mess. Really, the drawing order ought to be the table's current sort order - that it is not is a misfeature which may be corrected at some point. Note also that if you try this with a huge table you're likely to get a result which (a) is messy and (b) takes a very long time to draw.

#### Linear Correlation

If you select this option then a linear regression line will be plotted. The correlation coefficients will also be displayed to the right of the radio button (you may need to resize the window to see them all). The values cited are  $m$  (gradient),  $c$  (intercept) and  $r$  (product moment correlation coefficient). You can cut and paste from this text.

**Note** that for both the plotted line and the quoted coefficients the data is taken only from the points which are currently visible - that means that if you've zoomed the axes to

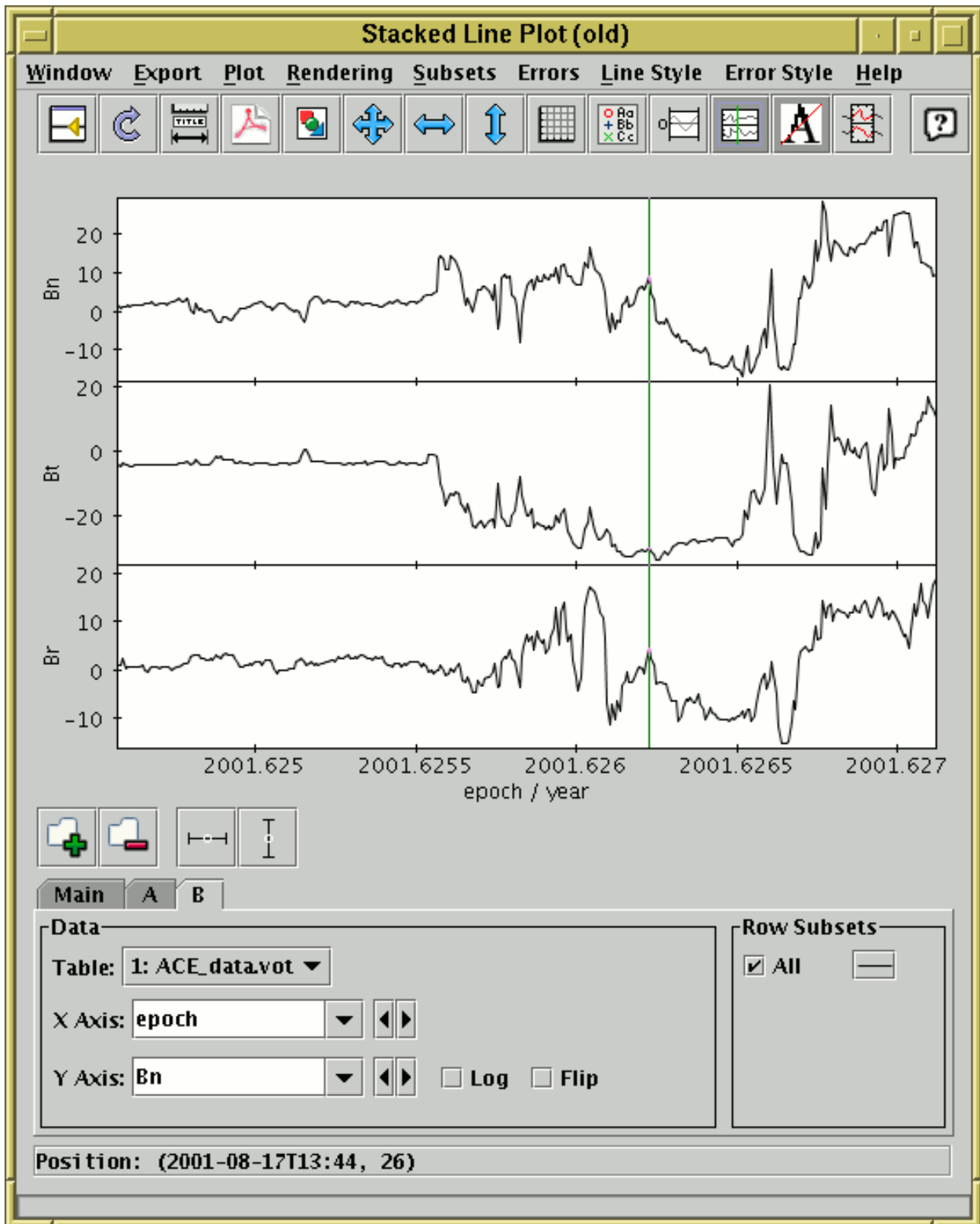
exclude some of the data points, they will not be contributing to the calculated statistics.

Any changes you make in this window are reflected in the plot straight away. If you click the **OK** button at the bottom, the window will disappear and the changes remain. If you click **Cancel** the window will disappear and any changes you made will be discarded.


You can also change all the plotting styles at once by using the **Marker Style** menu in the plot window. Here you can select a standard group of styles (e.g. all open 2-pixel markers with different colours and shapes) for the plotted sets. Similarly, error styles can be changed all at once using the **Error Style** menu.

#### **A.5.4 Stacked Line Plot (old-style)**

*This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4, though note that at time of writing (v4.1) no equivalent new-style plotting window is available for plotting stacked lines.*




### Stacked Lines Window


The stacked line plot window allows you to plot one or more ordinate (Y) quantities against a monotonic abscissa (X) quantity. For clarity, the different plots are displayed on vertically displaced graphs which share the same X axis. You can display this window using the **Lines**  item in


the Control Window's **Graphics** menu.

The display initially holds a single X-Y graph, usually with lines connecting adjacent points. The

points will be reordered before drawing if necessary so that the line is displayed as a function of X, rather than of an invisible third independent variable (in the Scatter Plot this isn't done which can lead to lines being scribbled all over the plot). If one of the columns in the table appears to represent a time value, this will be selected as the default X axis. Otherwise, the 'magic' **index** variable will be used, which represents the row number. Of course, these can be changed from their default values using the selectors in the usual way.

To add a new graph with a different Y axis, use the **Add Dataset** () button in the **Dataset**

**Toolbar** at the bottom of the window. This has a slightly different effect from what it does in the other plot windows, in that it inserts a new plotting region with its own Y axis at the top of the plot on which the specified data is drawn, rather than only causing a new set of points to be plotted on the existing plot region. Thus all the datasets appear in their own graphs with their own Y axes (though if you have multiple row subsets plotted for the same dataset they will appear on the same part of the plot as usual). To remove one of the graphs, select its tab and use the **Remove Dataset** () button as usual.

**Zooming** can only be done on one axis at a time rather than dragging out an X-Y region on the plot surface, since there isn't a single Y axis to zoom on. To zoom the X axis in/out, position the mouse just below the X axis at the bottom of the plot and drag right/left. To zoom one of the Y axes in or out, position the mouse just to the left of the Y axis you're interested in and drag down/up. To set the ranges manually, use the **Configure Axes and Title** () button as usual, but note that there

is one label/range setting box for each of the Y axes. These things work largely as described in Appendix A.5.1.2, as long as you bear in mind that the range of each of the Y axes is treated independently of the others.

Clicking on any of the points will **activate** it - see Section 8.

The following buttons are available on the toolbar:



#### **Split Window**

Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



#### **Configure Axes and Title**

Pops up a dialogue to allow manual configuration of the range and label for the X axis and each of the Y axes, as well as a plot title - see Appendix A.5.1.2.



#### **Export as PDF**

Pops up a dialogue which will write the current plot as a PDF file.



#### **Export as GIF**

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the

same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



### Rescale

Rescales all the plots so that all points in the plotted datasets can be seen. The X axis and all the Y axes are rescaled to fit the data.



### Rescale X Axis

Rescales the X axis only. The X axis is rescaled to cover the lowest and highest values on any of the plotted datasets, but the Y ranges are left as they are.



### Rescale Y Axes

Rescales the Y axes only. Each of the plotted Y axes are independently rescaled so that they cover the lowest and highest values within the currently visible X range.



### Full Grid

Toggles whether X and Y grid lines are drawn over the plots or not. If this is selected, the single  $y=0$  grid line (see next item) will automatically be deselected.



### Show Legend

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



### $y=0$ Grid Lines

Toggles whether a single horizontal line at  $y=0$  is drawn. If this is selected, the full grid (see previous item) will automatically be deselected.



### Show Vertical Crosshair

Toggles whether a vertical line follows the mouse when it is positioned over the plot. This can be useful to compare features in different graphs at the same X coordinate position.



### Antialias

Toggles whether lines are drawn with antialiasing. Antialiasing means smoothing lines so that they appear less pixelised, and generally improves the aesthetic appearance of the plot, but in some circumstances it might look better not antialiased. The state of this button does not affect images exported to postscript.



### Subset From X Range

Defines a new **Row Subset** in each of the plotted tables consisting of only the points in the currently visible range on the horizontal axis. Points will be included even if they are outside the current ranges of the Y axes.

The **Dataset Toolbar** contains the following options:



### Add/Remove dataset

Adds/removes a dataset for plotting, which both adds/removes a tab from the dataset selector and adds/removes a plot in the currently visible stack in the plotting area. See above for more explanation.



**Toggle X/Y error bars**

Switches between drawing symmetric error bars and no error bars in the X and Y directions respectively. Other options are available in the **Errors** menu. See Appendix A.5.1.3.


You can determine how the data are plotted using lines and/or markers as described in the following subsection.

#### A.5.4.1 Lines Style Editor

The default plotting style for the stacked lines plot is a simple black line for each graph. Since the plots typically do not overlap each other, this is in many cases suitable as it stands. However, you can configure the plotting style so that the points are plotted with markers as well as or instead of lines, and change the colours, marker shapes, line styles etc. The style for each row subset is displayed in a button to the right of its name in the bottom right of the plotting window. If you click this button the following dialogue will pop up which entables you to configure the plotting style.



#### Stacked Line Plot Style Editor

The **Legend** box defines how the selected set will be identified in the legend which appears alongside the plot (though the legend will only be visible if **Show Legend** (  ) is on):

##### Icon

Displays the icon which will be shown to identify the points in the selected set. Its appearance depends on the selections you make in the rest of this dialogue window.

##### Label

Gives the name written in the legend to label the subset. By default this is derived from the Row Subset's name and, if it's not part of the main dataset, the name of the dataset's tab. You can type in a new value to change what is written in the legend.

#### **Hide Legend**

If this checkbox is selected, then no entry for the selected set will appear in the legend.

The **Display** box defines how the markers plotted for each data point will appear:

#### **Colour**

Choose the colour in which the lines and/or markers will be plotted.

#### **Line/Markers**

Select from the radio buttons whether you want just lines between the data points, or markers at each point, or both.

#### **Error Bars**

If error bars are active for this plot, allows you to select the way they will appear. The options which can be selected here will depend on whether X and/or Y errors are in use.

The **Line** box defines how the lines joining the points will look. These controls will only be active if the Display selection is **Line** or **Both**.

#### **Thickness**

Determines line thickness.

#### **Dash**

Determines line dash pattern.

The **Markers** box defines how markers at the data points will look. These controls will only be active if the Display selection is **Markers** or **Both**.

#### **Size**

Determines the size of the markers in pixels. If a size of zero is chosen then the shape doesn't matter, the points will be plotted as single pixels.

#### **Shape**

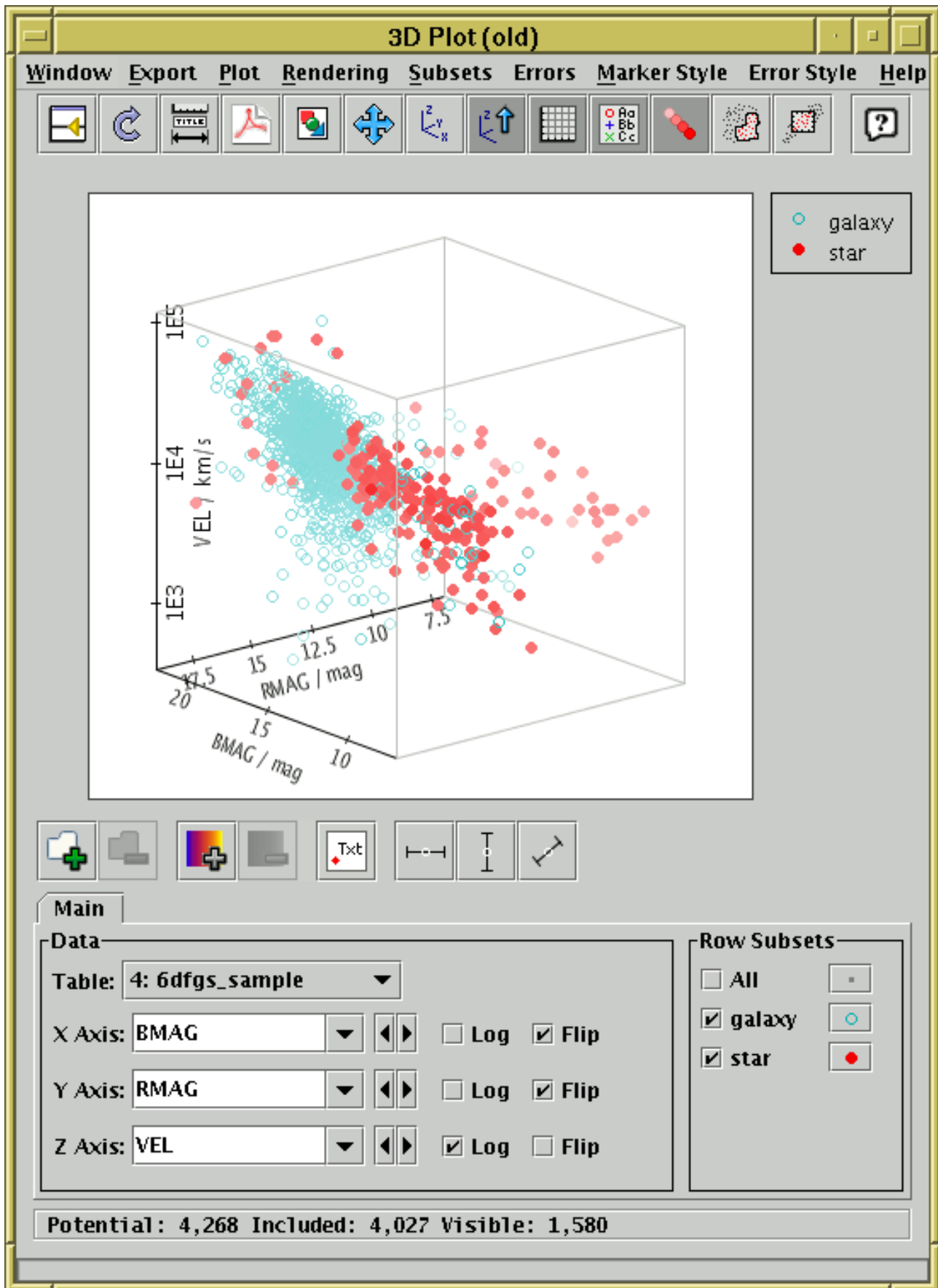
Determines the shape of the markers from a selection such as open or filled circles, squares, crosses etc.

Any changes you make in this window are reflected in the plot straight away. If you click the **OK** button at the bottom, the window will disappear and the changes remain. If you click **Cancel** the window will disappear and any changes you made will be discarded.


You can also change all the plotting styles at once by using the **Line Style** menu in the stacked lines plot window. Here you can select a standard group of styles (e.g. dashed lines, coloured lines) for the plotted sets. Similarly, error styles can be changed all at once using the **Error Style** menu.

### **A.5.5 3D Plot (old-style)**

*This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4.*



### 3D scatter plot window


The 3D plot window draws 3-dimensional scatter plots of one or more triples of table columns (or derived quantities) on Cartesian axes. You can display it using the **3D**  item in the Control



Window's **Graphics** menu.

On the display a marker is plotted for each row in the selected dataset(s) at a position determined by the values in the table columns selected to provide the X, Y and Z values. A marker will only be plotted if none of the X, Y and Z values are blank. Select the quantities to plot and the plotting symbols with the dataset selector (Appendix A.5.1.1) at the bottom.

The 3D space can be rotated by dragging the mouse around on the surface - it will rotate around the point in the centre of the plotted cube. The axis labels try to display themselves the right way up and in a way which is readable from the viewing point if possible, which means they move around while the rotation is happening. By default the points are rendered as though the 3D space is filled with a 'fog', so that more distant points appear more washed out - this provides a visual cue which can help to distinguish the depth of plotted points. However, you can turn this off if you want. If there are many points, then you may find that they're not all plotted while a drag-to-rotate gesture is in progress. This is done to cut down on rendering time so that GUI response stays fast. When the drag is finished (i.e. when you release the mouse button) all the points will come back again.

**Zooming** is also possible. You can zoom in around the centre of the plot so that the viewing window only covers the middle. The easiest way to do this is to use the mouse wheel if you have one - wheel forward to zoom in and backward to zoom out. Alternatively you can do it by dragging on the region to the left of the plot, like the **Axis Zoom** in some of the 2-d plots. Drag the mouse down to zoom in or up to zoom out on this part of the window. Currently it is only possible to zoom in/out around the centre of the plot. When zoomed you can use the **Subset From Visible** ()

toolbar button to define a new **Row Subset** consisting only of the points which are currently visible. See Appendix A.5.1.6 for more explanation.

Clicking on any of the plotted points will **activate** it - see Section 8.

The following buttons are available on the toolbar:



#### Split Window

Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



#### Replot

Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



#### Configure Axes and Title

Pops up a dialogue to allow manual configuration of axis ranges, axis labels and plot title - see Appendix A.5.1.2.



#### Export as PDF

Pops up a dialogue which will write the current plot as a PDF file. In general this is a faithful and high quality rendering of what is displayed in the plot window. However, if plotting is being done using the transparent markers, the markers will be rendered as if they were opaque. Plots with very many points can result in rather large output PDFs.



### Export as GIF

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



### Rescale

Rescales the axes of the current plot so that it contains all the data points in the currently selected subsets. By default the plot will be scaled like this, but it may have changed because of changes in the subset selection.



### Reorient

Reorients the axes of the current plot to their default position. This can be useful if you've lost track of where you've rotated the plot to with the mouse. This also resets the zoom level to normal if you've changed it.



### Stay Upright

Toggle button which when selected ensures that the Z axis is oriented vertically on the screen at all times. By default this is off which means you can drag the axes round to any orientation, but it can be convenient to switch it on to keep the plot pointing in a sensible direction.



### Grid

Toggles whether the cubic frame bounding the plot is drawn or not.



### Show Legend

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



### Fog

Toggles whether rendering is done as if the space is filled with fog. If this option is selected, distant points will appear more washed out than near ones.



### Draw Subset Region

Allows you to draw a region on the screen defining a new Row Subset. When you have finished drawing it, click this button again to indicate you're done. The subset will include points at all depths in the viewing direction which fall in the region you have drawn. See Appendix A.5.1.6 for more details.



### Subset From Visible

Defines a new **Row Subset** consisting of only the points which are currently visible on the plotting surface. See Appendix A.5.1.6 for more explanation.

The following additional item is available as a menu item only:



### Antialias

Toggles whether the axes and their annotations are drawn antialiased. Antialiased lines are smoother and generally look more pleasing, especially for text at a sharp angle, but it can slow the rendering down a bit.

The **Dataset Toolbar** contains the following options:



**Add/Remove dataset**

Adds/removes tabs in which the data for extra datasets can be entered. See Appendix A.5.1.1.



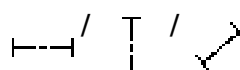
**Add/Remove auxiliary axis**

Adds/removes a selector for entering an auxiliary axis to modify point colours etc. See Appendix A.5.1.5.



**Toggle point labelling**

Allows text labels to be drawn near plotted points. See Appendix A.5.1.4.



**Toggle X/Y/Z error bars**

Switches between drawing symmetric error bars and no error bars in the X, Y and Z directions respectively. Other options are available in the **Errors** menu. See Appendix A.5.1.3.


You have considerable freedom to configure how the points are plotted including the shape, colour and transparency of symbols and the representation of error bars if used. These options are described in the following subsection.

#### A.5.5.1 3D Plot Style Editor

When plotting points in a 3D plot there are many different ways that each point can be displayed. By default, TOPCAT chooses a set of markers on the basis of how many points there are in the table and uses a different one for each plotted set. The marker for each set is displayed in a button to the right of its name in the dataset selector panel at the bottom of the plot window. If you click this button the following dialogue will pop up which enables you to change the appearance.



Style editor dialogue for 3d plots

The **Legend** box defines how the selected set will be identified in the legend which appears alongside the plot (though the legend will only be visible if **Show Legend**  is on):

#### Icon

Displays the icon which will be shown to identify the points in the selected set. Its appearance depends on the selections you make in the rest of this dialogue window.

#### Label

Gives the name written in the legend to label the subset. By default this is derived from the Row Subset's name and, if it's not part of the main dataset, the name of the dataset's tab. You can type in a new value to change what is written in the legend.

#### Hide Legend

If this checkbox is selected, then no entry for the selected set will appear in the legend.

The **Marker** box defines how the markers plotted for each data point will appear:

#### Shape

Choose from a variety of shapes such as open or filled circles, squares, crosses etc.

#### Size

Choose the size of the marker; the value given is approximate radius in pixels. If a size of zero is chosen, then the shape doesn't matter, the points will be plotted as single pixels.

#### Colour

Choose the colour in which the markers will be plotted.

#### Transparency

Choose transparency of the plotted symbols. The scale on the slider is logarithmic, with 1 at the left hand end. The actual value chosen is an integer written at the right of the slider. This number gives the number of markers for this set which need to be plotted in the same position to result in fully opaque pixels - any fewer and the background, or other markers plotted underneath, will show through to some extent. Setting this to some value greater than 1 is very useful if you have a very large number of points being plotted (especially if it's comparable with the number of the pixels on the screen), since it enables to you distinguish regions where there are lots of points on top of each other from those where there are only a few. If a finite transparency is set, you may find it useful to turn off fogging (see above).

### **Error Bars**

If error bars are active for this plot, allows you to select the way they will appear. The options which can be selected here will depend on whether X, Y and/or Z errors are in use.

### **Hide Markers**

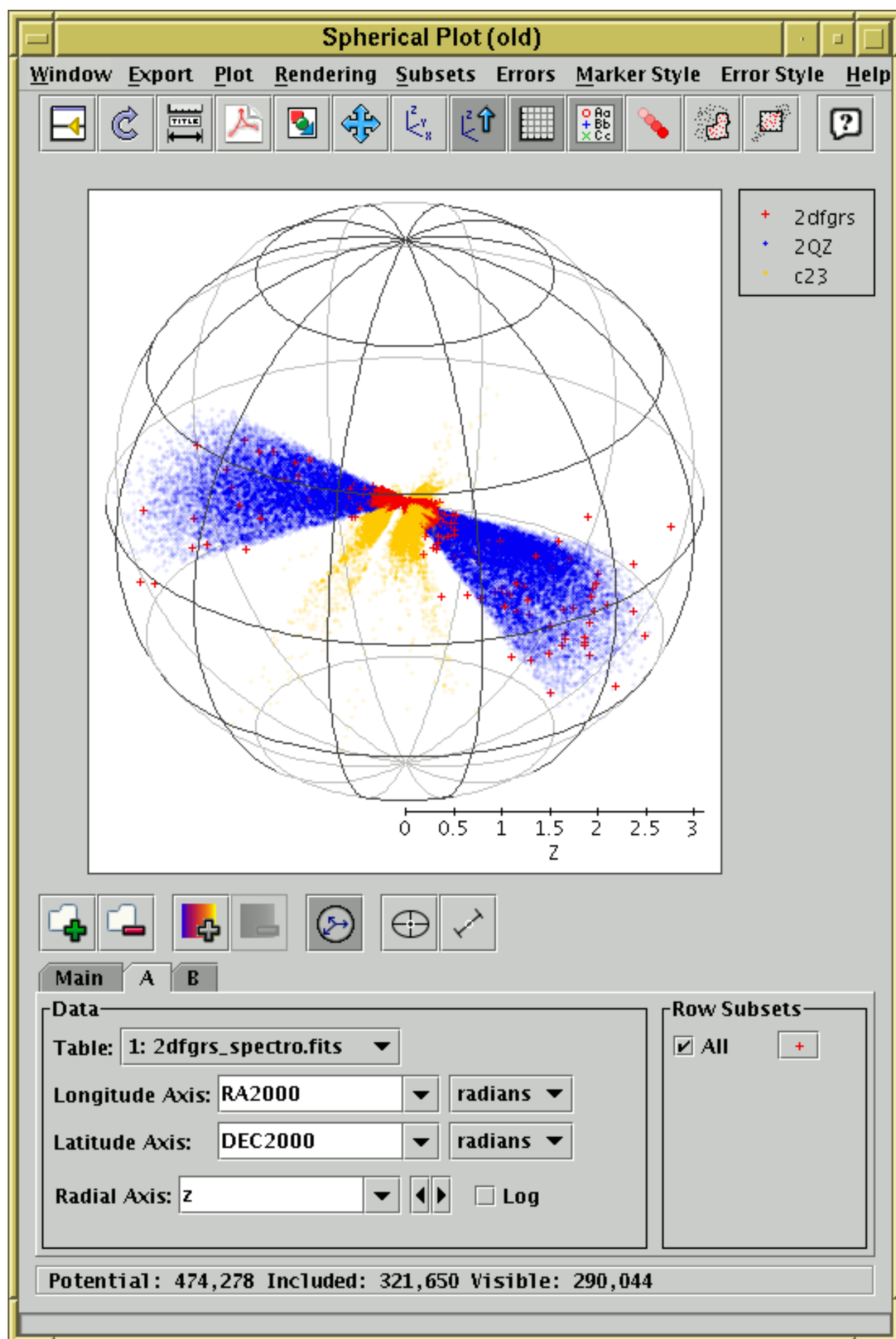
This check box is only enabled if error bars are being plotted; it allows the markers to be invisible, so that only the error bars are seen.

Any changes you make in this window are reflected in the plot straight away. If you click the **OK** button at the bottom, the window will disappear and the changes remain. If you click **Cancel** the window will disappear and any changes you made will be discarded.


You can also change all the plotting styles at once by using the **Marker Style** menu in the plot window. Here you can select a standard group of styles (e.g. all open 2-pixel markers with different colours and shapes) for the plotted sets. Similarly, error styles can be changed all at once using the **Error Style** menu.

## **A.5.6 Spherical Plot (old-style)**


*This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4.*




## Spherical plot window

The spherical plot window draws 3-dimensional scatter plots of datasets from one or more tables on spherical polar axes, so it's suitable for displaying the position of coordinates on the sky or some other spherical coordinate system, such as the surface of a planet or the sun. You can display it using the **Sphere** () item in the Control Window's **Graphics** menu.

In most respects this window works like the 3D Plot window, but it uses spherical polar axes rather than Cartesian ones. You have to fill in the dataset selector (Appendix A.5.1.1) at the bottom with longitude- and latitude-type coordinates from the table. Selectors are included to indicate the units of those coordinates. If TOPCAT can locate columns in the table which appear to represent Right Ascension and Declination, these will be filled in automatically. If only these two are filled in, then the points will be plotted on the surface of the unit sphere - this is suitable if you just want to inspect the positions of a set of objects in the sky.

If the **Radial Coordinates** () button is activated, you can optionally fill in a value in the **Radial Axis** selector as well. In this case points will be plotted in the interior of the sphere, at a distance from the centre given by the value of the radial coordinate.

The 3D space can be rotated by dragging the mouse around on the surface - it will rotate around the centre of the sphere. By default the points are rendered as though the 3D space is filled with a 'fog', so that more distant points appear more washed out - this provides a visual cue which can help to distinguish the depth of plotted points. However, you can turn this off if you want. If there are many points, then you may find that they're not all plotted while a drag-to-rotate gesture is in progress. This is done to cut down on rendering time so that GUI response stays fast. When the drag is finished (i.e. when you release the mouse button) all the points will come back again.

**Zooming** is also possible. You can zoom in around the centre of the plot so that the viewing window only covers the middle. The easiest way to do this is to use the mouse wheel if you have one - wheel forward to zoom in and backward to zoom out. Alternatively you can do it by dragging on the region to the left of the plot, like the **Axis Zoom** in some of the 2-d plots. Drag the mouse down to zoom in or up to zoom out on this part of the window. Currently it is only possible to zoom in/out around the centre of the plot. When zoomed you can use the **Subset From Visible** ()

toolbar button to define a new **Row Subset** consisting only of the points which are currently visible. See Appendix A.5.1.6 for more explanation.

Clicking on any of the plotted points will **activate** it - see Section 8.

The following buttons are available on the toolbar:



Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



### Configure Axes and Title

Pops up a dialogue to allow manual configuration of axis ranges, axis labels and plot title - see Appendix A.5.1.2. The only configurable axis range is the upper limit of the radial axis.



### Export as PDF

Pops up a dialogue which will write the current plot as a PDF file. In general this is a faithful and high quality rendering of what is displayed in the plot window. However, if plotting is being done using the transparent markers, the markers will be rendered as if they were opaque. Plots with very many points can result in rather large output PDFs.



### Export as GIF

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



### Rescale

Rescales the axes of the current plot so that it contains all the data points in the currently selected subsets. By default the plot will be scaled like this, but it may have changed because of changes in the subset selection.



### Reorient

Reorients the axes of the current plot to their default position. This can be useful if you've lost track of where you've rotated the plot to with the mouse. This also resets the zoom level to normal if you've changed it.



### Stay Upright

Toggle button which when selected ensures that the north pole (latitude = +90 degrees) is oriented vertically on the screen at all times. By default this is on.



### Grid

Toggles whether the spherical wire frame bounding the plot is drawn or not.



### Show Legend

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



### Fog

Toggles whether rendering is done as if the space is filled with fog. If this option is selected, distant points will appear more washed out than near ones.



### Draw Subset Region

Allows you to draw a region on the screen defining a new Row Subset. When you have finished drawing it, click this button again to indicate you're done. The subset will include points at all depths in the viewing direction which fall in the region you have drawn. See Appendix A.5.1.6 for more details.



### Subset From Visible

Defines a new **Row Subset** consisting of only the points which are currently visible on the plotting surface. See Appendix A.5.1.6 for more explanation.



The following additional item is available as a menu item only:



### Antialias

Toggles whether the axes and their annotations are drawn antialiased. Antialiased lines are smoother and generally look more pleasing, especially for text at a sharp angle, but it can slow the rendering down a bit.

The **Dataset Toolbar** contains the following options:



### Add/Remove dataset

Adds/removes tabs in which the data for extra datasets can be entered. See Appendix A.5.1.1.



### Add/Remove auxiliary axis

Adds/removes a selector for entering an auxiliary axis to modify point colours etc. See Appendix A.5.1.5.



### Toggle point labelling

Allows text labels to be drawn near plotted points. See Appendix A.5.1.4.



### Toggle Radial Coordinates

When activated, a column selector labelled **Radial Axis** will be visible below the Longitude and Latitude Axis selectors. This enables you to enter a value for the radial coordinate of each point. If this is disabled, or if it has a blank value, then all the points will have an assumed radial coordinate of unity and will be plotted on the surface of the sphere.



### Toggle tangential error bars

When activated, an additional column selector appears in the dataset panel to the right of the Longitude and Latitude selectors, along with its own unit selector. You can fill this in with an isotropic error value representing the radius of a small circle associated with the selected points, in units of arcsec, arcmin, degrees or radians. This will cause 2-d error bars to be plotted. You can configure their appearance (e.g. crosshairs, ellipses, rectangles, ...) using the style editor in the usual way. See Appendix A.5.1.3 for more information.



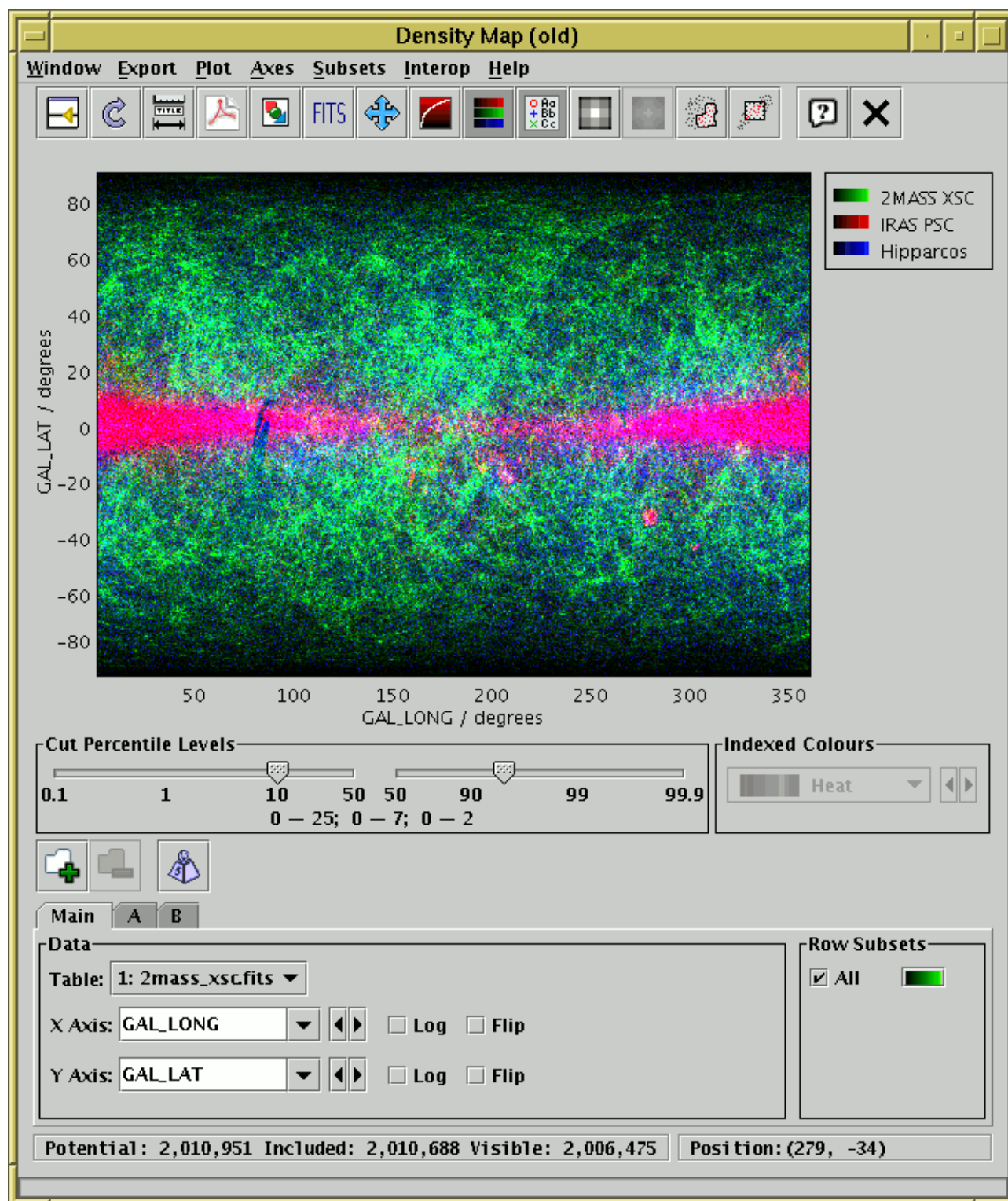
### Toggle radial error bars

Switches radial error bars on and off. See Appendix A.5.1.3. This button will only be enabled if the Radial Coordinates are in use.

You have considerable freedom to configure how points are plotted including the shape, colour and transparency of symbols and the representation of errors if used. This works exactly as for the Cartesian 3D plot as described in Appendix A.5.5.1.

## A.5.7 Density Map (old-style)

*This section describes an old-style plotting window. The standard plotting windows are described in Appendix A.4.*



### Density map window in RGB mode

The density map window plots a 2-dimensional density map of one or more pairs of table columns (or derived quantities); the colour of each pixel displayed is determined by the number of points in the data set which fall within its bounds. Another way to think of this is as a histogram on a 2-dimensional grid, rather than a 1-dimensional one as in the Histogram Window. You can optionally weight these binned counts with another value from the table.

Density maps are suitable when you have a very large number of points to plot, since in this case it's important to be able to see not just whether there is a point at a given pixel, but how many points fall on that pixel. To a large extent, the transparency features of the other 2d and 3d plotting

windows address this issue, but the density map gives you a bit more control. It can also export the result as a FITS image, which can then be processed or viewed using image-specific software such as GAIA or Aladin.


This window can be operated in two modes:

### Indexed Mode

Each pixel represents a single scalar value, corresponding to a count or sum as indicated by the selected dataset(s). If multiple datasets are being plotted at once, the values from each will be summed to give the result at each point. The mapping from numeric value to pixel colour at each point on the plot is determined by the colour map selected in the **Indexed Colours** selector below the plot. In this case the style editor colour selectors have no effect and are disabled. A fairly wide range of colour maps is provided by default. If these do not suit your needs, it is possible to provide your own custom colour maps using the **lut.files** system property - see Section 10.2.3.

### RGB Mode

Each pixel has up to three independent contributions, its intensity in Red, Green and Blue channels. These can come from different datasets, as configured in the style editor. If more than one dataset is assigned the same colour, the contributions are summed for that channel. In this case the Indexed Colours colour map selector has no effect and is disabled.

Switch between the modes using the **RGB** () button.

You can configure the axes, including **zooming** in and out, with the mouse (drag on the plot or the axes) or manually as described in Appendix A.5.1.2.

Two controls specific to this window are shown below the plot itself:

### Cut Percentile Levels

This controls how the number of counts in each pixel maps to a brightness. There are two sliders, one for the lower bound and one for the upper bound. They are labelled (logarithmically) with percentile values. If the upper one is set to 90, it means that any pixel above the 90th percentile of the pixels in the image in terms of count level will be shown with maximum brightness, and similarly for the lower one. These values apply independently to each colour channel if more than one is in use. Immediately below the sliders, the pixel values which correspond to minimum and maximum brightness are displayed. In indexed mode there is one range, and in RGB mode there may be up to three. If the image is not fairly completely covered, this control doesn't give you as much freedom as you might like - the user interface may be improved in future releases.

### Indexed Colours

When in indexed (non-RGB) mode only, this allows you to select a colour map which determines how pixel values (counts or sums per bin) are turned into colours on the screen. The lowest value corresponds to the colour at the left side of the icon and the highest value to the right side. In RGB mode this is disabled.

The following buttons are available on the toolbar:



Allows the dataset selector to be resized by dragging a separator between it and the plot area. Good for small screens.



Redraws the current plot. It is usually not necessary to use this button, since if you change any of the plot characteristics with the controls in this window the plot will be redrawn automatically. However if you have changed the data, e.g. by editing cells in the Data Window, the plot is not automatically redrawn (since this is potentially an expensive operation and you may not require it). Clicking this button redraws the plot taking account of any changes to the table data.



### **Configure Axes and Title**

Pops up a dialogue to allow manual configuration of axis ranges, axis labels and plot title - see Appendix A.5.1.2.



### **Export as PDF**

Pops up a dialogue which will write the current plot as a PDF file.



### **Export as GIF**

Pops up a dialogue which will output the current plot to a GIF file. The output file is just the same as the plotted image that you see. Resize the plotting window before the export to control the size of the output GIF.



### **Export as FITS**

Pops up a dialogue which will output the plotted map as a FITS array. If only one channel is visible (either one colour channel or indexed mode) then the output FITS file will be a 2d array with dimensions the same as the displayed image. If there are multiple RGB channels then the output array will be 3d with the third dimension having an extent of 2 or 3, depending on the number of colour channels visible. In either case the FITS file will have a single (primary) HDU. Basic coordinate system information, as well as DATAMIN and DATAMAX cards, will be written to the header. The type of the output array will be double precision for weighted values, or some integer type of sufficient length for unweighted ones.



### **Rescale**

Rescales the axes of the current plot so that it contains all the data points in the currently selected subsets. By default the plot will be scaled like this, but it may have changed because of changes in the subset selection or from zooming in or out.



### **Log Intensity**

Toggles between linear and logarithmic mapping for colour intensity as a function of number of counts.



### **Colour**

Toggles between indexed and RGB modes (see the explanation above).



### **Show Legend**

Toggles whether a legend showing how each data set is represented is visible to the right of the plot. Initially the legend is shown only if more than one data set is being shown at once.



### **Bigger Pixels**

Increments the size of screen pixel corresponding to one density map bin.



### **Smaller Pixels**

Decrements the size of screen pixel corresponding to one density map bin.



### Draw Subset Region

Allows you to draw a region on the screen defining a new Row Subset. When you have finished drawing it, click this button again to indicate you're done. See Appendix A.5.1.6 for more details.



### Subset From Visible

Defines a new **Row Subset** consisting of only the points which are currently visible on the plotting surface. See Appendix A.5.1.6 for more explanation.

The **Dataset Toolbar** contains the following options:



### Add/Remove dataset

Adds/removes tabs in which the data for extra datasets can be entered. See Appendix A.5.1.1.



### Weight Counts

If this toggle button is on, an additional **Weight Axis** selector appears below the **X Axis** selector. If this is filled in with a column name or expression, then instead of simply accumulating the number of points per bin, each pixel will represent the sum over the weighting quantity for points in each bin. Not having a weight axis is equivalent to filling in its value with the quantity 1. Note that with weighting, the figure drawn is no longer strictly speaking a histogram or density map.

The **Export** menu provides a number of ways to export the displayed image for external viewing or analysis. As well as options to export as GIF, JPEG, EPS and FITS, there is also the option to transmit the FITS image to one or all applications listening using the SAMP or PLASTIC tool interoperability protocol which will receive images. In this way you can transmit the image directly to SAMP/PLASTIC-aware image manipulation tools such as GAIA or Aladin. See Section 9 for more information about tool interoperability.


How to set the colour channel corresponding to each dataset is explained in the following subsection.

#### A.5.7.1 Density Style Editor

For a density map in RGB mode, each dataset is assigned a colour channel to which it contributes. A representation of this is displayed in a button to the right of its name in the dataset selector panel at the bottom of the density map window. If you click this button the following dialogue will pop up which enables you to change the colour channel.



### Style editor dialogue for density map

The **Legend** box defines how the selected set will be identified in the legend which appears alongside the plot (though the legend will only be visible if **Show Legend**  is on):

#### Icon

Displays the icon which will be shown to identify the points in the selected set. Its appearance depends on the selections you make in the rest of this dialogue window.

#### Label

Gives the name written in the legend to label the subset. By default this is derived from the Row Subset's name and, if it's not part of the main dataset, the name of the dataset's tab. You can type in a new value to change what is written in the legend.

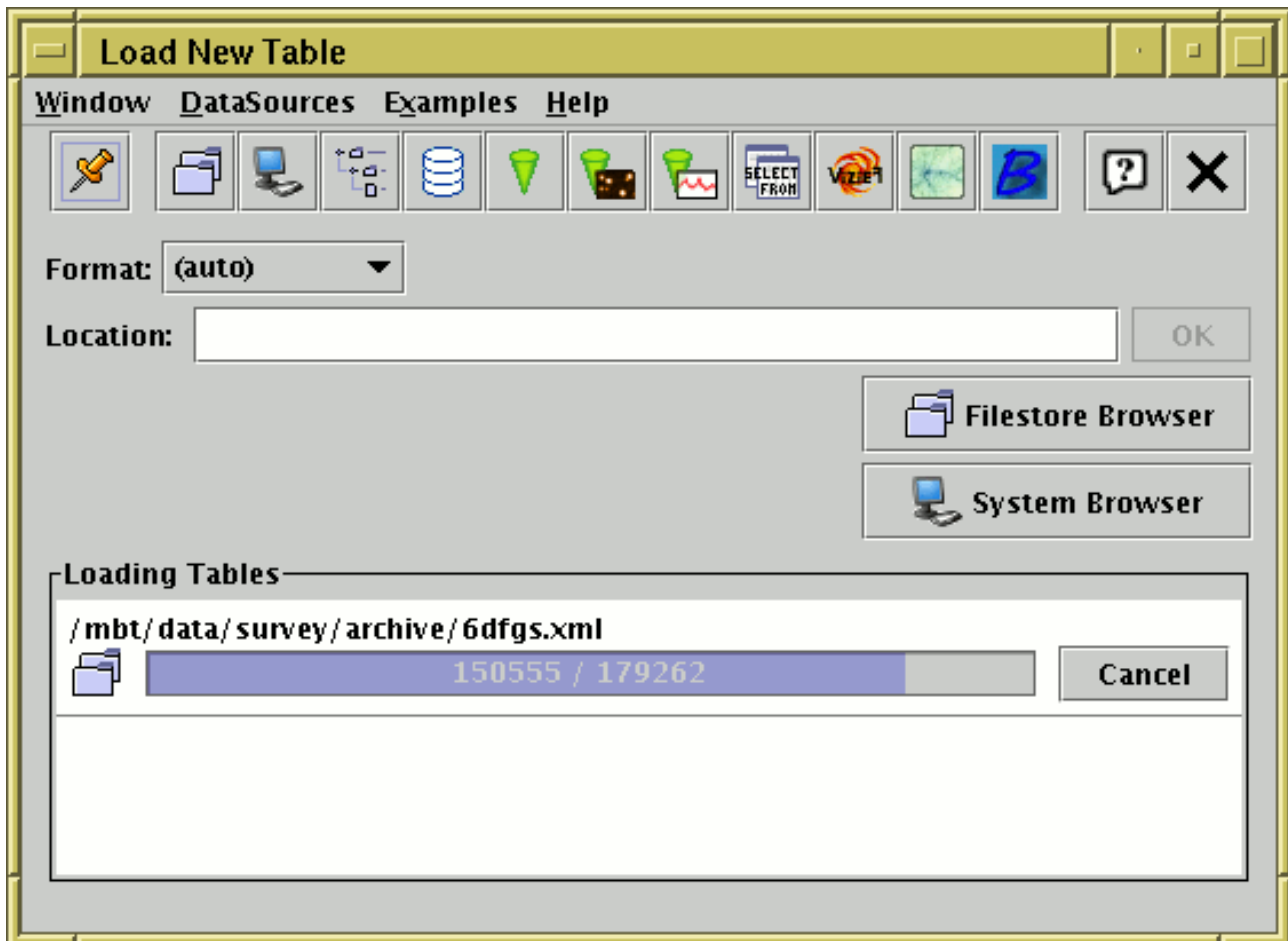
#### Hide Legend

If this checkbox is selected, then no entry for the selected set will appear in the legend.


The **Channel** selector allows you to select either the Red, Green or Blue channel for this dataset to contribute to. Note that this is only enabled in RGB mode; in indexed mode it has no effect and is disabled.

## A.6 Load Window





## Load Window

The Load Window is used for loading tables from an external location (e.g. disk or URL) into TOPCAT. It is obtained using the **Load Table** button () in the Control Window toolbar or **File** menu.












This dialogue allows you to specify a new table to open in a number of different ways, described below. If you perform a successful load using any of these options, a new entry or entries will be added into the Table List in the Control Window, which you can then use in the usual ways. If you choose a location which can't be turned into a table (for instance because the file doesn't exist), a window will pop up telling you what went wrong. The panel at the bottom of the window displays progress for tables currently loading; it is used to show progress for tables loaded from other sources too, for instance received via SAMP or specified on the command line.

In the simplest case, you can type a name into the **Location** field and hit return or the **OK** button. This location can be a filename or a URL, possibly followed by a '#' character and a 'fragment identifier' to indicate where in the file or URL the table is located; the details of what such fragment identifiers mean can be found in the relevant subsection within Section 4.1.1. Allowed URL types are described in Section 4.2. You should select the relevant table format from the **Format** selector box - you can leave it on **(auto)** for loading FITS tables or VOTables, but for other formats such as ASCII or CSV you must select the right one explicitly (again, see Section 4.1.1 for details).


There are many other ways of loading tables however, described in the following subsections. The **Filestore Browser** and **System Browser** buttons are always visible below the location field. Depending on startup options, there may be other buttons here. There are also a number of toolbar buttons which display various load dialogues; the **DataSources** contains all of these along with some lesser-used ones. The following subsections describe most of the available options, though

others may be available in your installation.

The options available on the toolbar by default are these:

-  Filestore Browser
-  System Browser
-  Hierarchy Browser
-  SQL Query
-  Cone Search
-  Simple Image Access (SIA) Query
-  Simple Spectral Access (SSA) Query
-  Table Access Protocol (TAP) Query
-  VizieR Catalogue Service Query
-  GAVO Millennium Database Query
-  BaSTI Theory Database Query

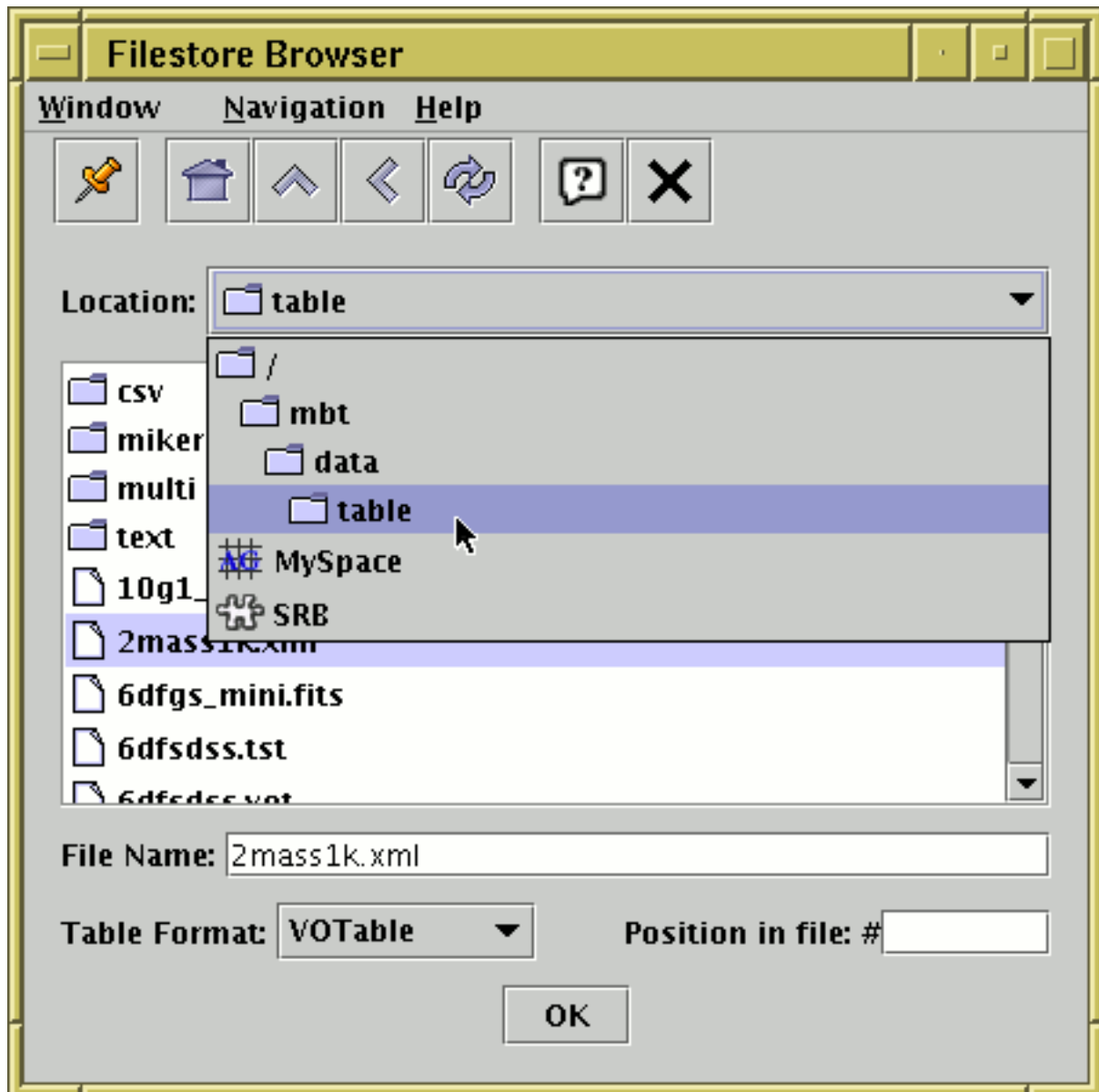
All of these load dialogues have an **OK** button. Once you click it, whatever load you have specified will start. If the load takes more than a few hundredths of a second, a progress bar will appear in the **Loading Tables** panel of the load window, as in the screenshot above. This will report on how many rows have been loaded, and if known, how many there are in total. If you want to interrupt the load for any reason while it is in progress, click the **Cancel** button, and the load will cease. If the load completes without cancellation, a new table will appear in the Table List of the main Control Window, ready for use.

By default, when a table load has completed, both the Load Window itself and whichever specific load dialogue window you used will close. If you don't want this to happen use the **Stay Open** (  ) item in the **Window** menu or toolbar of the Load Window and/or specific load dialogue. If


this is selected, the window will not automatically close. This can be very convenient if you want to load many tables from a similar place, for instance several files from the same directory or several cone searches to different services.

### A.6.1 Filestore Browser





### Filestore Browser window

By clicking the **Filestore Browser** button or toolbar button (  ) in the Load Window, you can obtain a file browser which will display the files in a given directory. The way this window works is almost certainly familiar to you from other applications.

Unlike a standard file browser however, it can also browse files in remote filestores: currently supported are MySpace and SRB. MySpace is a distributed storage system developed for use with the Virtual Observatory by the AstroGrid project, and SRB (Storage Resource Broker) is a similar general purpose system developed at SDSC. To make use of these facilities, select the relevant entry from the selector box at the top of the window as illustrated above; this will show you a **Log In** button which prompts you for username, password etc, and you will then be able to browse the remote filestore as if it were local. The same button can be used to log out when you are finished, but the session will be logged out automatically when TOPCAT ends in any case. Access to remote filesystems is dependent on certain optional components of TOPCAT, and it may not be available if you have the topcat-lite configuration.

The browser initially displays the current directory, but this can be changed by typing a new directory into the **File Name** field, or moving up the directory hierarchy using the selector box at the top, or navigating the file system by clicking the up-directory button or double-clicking on


displayed directories. The initial default directory can be changed by setting the `user.dir` system property (Section 10.2.3).

All files are shown, and there is no indication of which ones represent tables and which do not. To open one of the displayed files as a table, double-click on it or select it by clicking once and click the **Open Table** button. The **Table Format** selector must be set correctly: the "(auto)" setting will automatically detect the format of VOTable or FITS tables, otherwise you will need to select the option describing the format of the file you are attempting to load (see Section 4.1.1). If you pick a file which cannot be converted into a table an error window will pop up.

In most cases, selecting the file name and possibly the format is all you need to do. However, the **Position in file** field allows you to add information about where in the file the table you want is situated. The meaning of this varies according to the file format: for FITS files, it is the index or EXTNAME of the HDU containing the table you're after (see Section 4.1.1.1 for details), and for VOTables it is the index of the TABLE element (the first TABLE encountered is numbered 0). If you leave this blank, you will get all the tables in the file; this is usually just one table, since most file formats cannot accommodate more than one table per file, and even for those which can (FITS and VOTable) most files contain only a single file in any case.

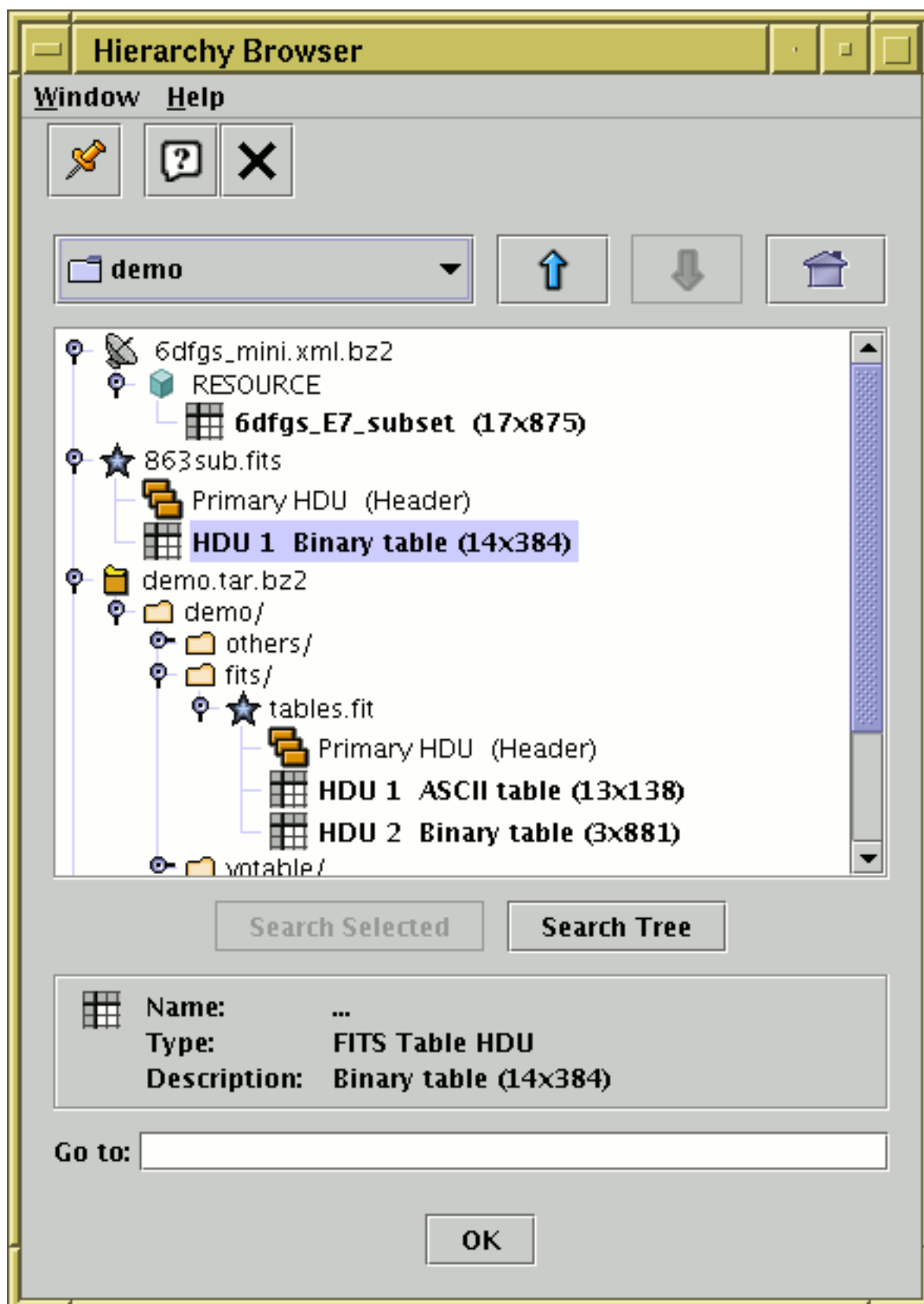
For a more table-aware view of the file system, use the Hierarchy Browser (Appendix A.6.3) instead.

### A.6.2 System Browser


By clicking the **System Browser** button or toolbar button () in the Load Window, you can use your Operating System's native file browser to choose a file to load from. What this looks like is entirely dependent on your OS; it may or may not contain system-specific features like shortcuts to commonly-used directories.

Since TOPCAT has no control over the way this dialogue looks, it cannot contain the **Table Format** selector, and certain other things such as load cancel may not work as well as for the other dialogue types. To select the table format, you will need to use the selector in the main **Load Window**.

### A.6.3 Hierarchy Browser




### File load Hierarchy Browser window



By selecting the **Hierarchy Browser** button () from the Load Window's toolbar or

**DataSources** menu, you can obtain a browser which presents a table-aware hierarchical view of the file system. (Note that a freestanding version of this panel with additional functionality is available in the separate Treeview application).

This browser resembles the Filestore Browser in some ways, but with important differences:

- It shows the file system in a 'tree-like' fashion, so that multiple levels of the hierarchy are displayed at once
- It understands which items in the hierarchy represent tables that can be automatically detected and which represent other kinds of object (for instance directories, zip files, or plain text files)
- It can look inside hierarchical files, so for instance it can investigate a Tar or Zip archive which may contain table entries, or display multiple tabular HDUs in a FITS file, or multiple TABLE elements at different levels in a VOTable document

The main part of the window shows a "tree" representation of the hierarchy, initially rooted at the current directory (the initial directory can be changed by setting the `user.dir` system property (Section 10.2.3)). Each line displayed represents a "node" which may be a file or some other type of item (for instance an HDU in a FITS file or an entry in a tar archive). The line contains a little icon which indicates what kind of node it is and a short text string which gives its name and maybe some description. Nodes which represent tables are indicated by the  icon. For nodes which have

some internal structure there is also a "handle" which indicates whether they are collapsed () or expanded (). You can examine remote filesystems (MySpace, SRB) as well as local ones in the same way as with the Filestore Browser.

If you select a node by clicking on it, it will be highlighted and some additional description will appear in the panel below the hierarchy display. The text is in **bold** if the node in question can be opened as a table, and non-bold if it is some non-table item.

**Note:** an important restriction of this browser is that it will only pick up tables which can be identified automatically - this includes FITS and VOTable files, but does not include text-based formats such as ASCII and Comma-Separated Values. If you want to load one of the latter types of table, you will need to use one of the other load methods and specify table format explicitly.

You can see how this browser works on an example directory of tables as described in Appendix A.6.12.

Note that this window requires certain optional components of the TOPCAT installation, and will not be available if you have the topcat-lite configuration.

### A.6.3.1 Navigation

Navigation is a bit different from navigation in the **File Browser** window. To expand a node and see its contents, click on its handle (clicking on the handle when it is expanded will collapse it again). When you have identified the table you want to open, highlight it by clicking on it, and then click the **Open Table** button at the bottom.

To move to a different directory, i.e. to change the root of the tree which is displayed, use one of the buttons above the tree display:

#### Selector box

Allows you to move straight to any directory higher up than the current one.



Up

Moves to the parent of the current directory.



Down

Moves to the currently selected (highlighted) node.

**Home**

Moves to the user's home directory.

Alternatively, you can type in a new directory in the **Go to** field at the bottom of the window.

(In fact the above navigation options are not restricted to changing the root to a new directory, they can move to any node in the tree, for instance a level in a Tar archive.)




### A.6.3.2 Table Searches

There are two more buttons in the browser, **Search Selected** and **Search Tree**. These do a recursive search for tables in all the nodes starting at the currently selected one or the current root respectively. What this means is that the program will investigate the whole hierarchy looking for any items which can be used as tables. If it finds any it will open up the tree so that they are visible (note that this doesn't mean that the only nodes revealed will be tables, ancestors and siblings will be revealed too). This can be useful if you believe there are a few tables buried somewhere in a deep directory structure or Tar archive, but you're not sure where. Note that this may be time-consuming - a busy cursor is displayed while the search is going on. Changing the root of the tree will interrupt the search.

### A.6.4 SQL Query

**SQL Query**

Window Help

**Protocol:** jdbc:  ▼

**Host:**   ▼

**Database name:**

**User name:**

**Password:**

**SQL Query: #**

```
SELECT
  pos_eq_ra, pos_eq_dec, phot_flux_peak
FROM first1
WHERE phot_flux_peak < 21.5
LIMIT 10000
```

**OK**

### SQL Query Dialogue

If you want to read a table from an SQL database (Section 4.1.1.10), you can use a specialised dialogue to specify the SQL query by selecting **SQL Query** button from the Load Window's toolbar

() or **DataSources** menu.

This provides you with a list of fields to fill in which make up the query, as follows:

**Protocol**

The name of the appropriate JDBC sub-protocol. This is defined by the JDBC driver that you are using, and is for instance "mysql" for MySQL's Connector/J driver or "postgresql" for PostgreSQL's JDBC driver.

**Host**

The hostname of the machine on which the database resides (may be "localhost" if the database is local).

**Database name**

The name of the database.

**User name**

The username under which you wish to access the database. This is not strictly necessary if there is no access control for the database in question.

**Password**

The password for the given username. Again, whether this is necessary depends on the access policy of the database.

**SQL Query**

The text of the query which will define the resulting table. If you want to look at a table named XXX as it exists in the database, you can write something like "SELECT \* from XXX". In principle any SQL query on the database can be used here, but the details of what SQL syntax is permitted may be restricted by the JDBC driver you are using.

There are a number of criteria which must be satisfied for SQL access to work within TOPCAT (installation of appropriate drivers and so on) - see Section 10.3. If you don't take these steps, this dialogue may be inaccessible.

The way that this dialogue contacts the database makes some assumptions about how the JDBC driver works which are not always true, so for some databases and drivers it may not be possible to get it to work. It may be improved in the future; if you have particular problems, please contact the author or the mailing list.

**A.6.5 Cone Search**

See Appendix A.9.2 for details.

**A.6.6 SIA Query**

See Appendix A.9.3 for details.

**A.6.7 SSA Query**

See Appendix A.9.4 for details.


**A.6.8 TAP Query**

See Appendix A.9.8 for details.

### **A.6.9 VizieR Catalogue Service Query**

**VizieR Catalogue Service**

Window Help



**VizieR Server**

Server:

**Row Selection**

☒ Cone Selection

Object Name:

RA:  degrees (J2000)

Dec:  degrees (J2000)

Radius:  degrees

☐ All Rows

Maximum Row Count:

**Column Selection**

Output Columns:

**Catalogue Selection**

By Category By Keyword Surveys Missions


Wavelength	Mission	Astronomy
Radio	AKARI	Spectrophotometry
IR	ANS	Spectroscopy
optical	ASCA	Stars
UV	BeppoSAX	Stars:Emission

☐ Sub-Table Details ☐ Include Obsolete Tables

Name	Popul...	Density	De:
III/198	1692	1	Palomar/MSU nearby star spectr
J/ApJ/669/821	1304	1	CO Tully-Fisher relation for host
III/251	1029	1	ELODIE library V3.1 (Prugniel+ 2
II/149	921	1	Diffuse Interstellar Band Measure
J/ApJ/171/146	844	1	Population synthesis in the blue.



## VizieR load dialogue

The VizieR query dialogue can be opened using the **VizieR Catalogue Service** button () in the

Load Window's toolbar or the Control Window's **VO** menu. It allows you to make queries directly to the VizieR service operated by CDS. VizieR is a comprehensive library of very many published astronomical catalogues. These items can equally be accessed from the web or other interfaces, but this load dialogue makes it convenient to load data directly from VizieR into TOPCAT.

Note that VizieR's idea of a catalogue is more complex than a single table; this means that in some cases querying one of VizieR's catalogues may result in more than one table being loaded into TOPCAT (the **Sub-Table Details** checkbox described below can help to control this).

The dialogue consists of four parts: the **VizieR Server**, **Row Selection**, **Column Selection** and **Catalogue Selection** panels, arranged top to bottom in the window. These are described below.

The **VizieR Server** panel allows you to specify which VizieR server you want to use for data download. By default the server at CDS is used, but there are mirrors elsewhere, whose URLs can be chosen from the selector. If you see a popup window complaining that the server cannot be contacted, you can choose a different one; you might also want to select one that is close to you for performance reasons.

The **Row Selection** panel specifies which rows you want to retrieve from the catalogue that you will select. You can choose one of the two radio buttons according to the kind of query that you want to make:

### Cone Selection

In this case you must give a central sky position and the search radius to define a cone-shaped region of interest. Rows within that range will be returned. For the central position you can either fill in the **RA** and **Dec** fields directly, or you can fill in the **Object Name** field and hit the **Resolve** button; in the latter case, a SIMBAD query will be made to determine the coordinates corresponding to the named object.

### All Rows

Alternatively, you can choose to download the whole catalogue without spatial restrictions.

In either case, the **Maximum Row Count** selector indicates the largest number of rows which will be returned. If your query requests more rows than the limit given, extra rows will simply be omitted from the returned result (the limit seems to be approximate). It is possible to choose any value for this field, including very large ones or the special value "unlimited"; however consider before doing this whether you want to download a potentially very large data set. The server may in any case time out in the case of a very long connection, so it is probably not possible, even if it were desirable, to download for instance the entire 2MASS point source catalogue.

The **Column Selection** panel gives you some control over which columns will be included in the loaded table. This will include some or all of the columns the table has in the VizieR archive, and perhaps some standard ones added automatically by the service. The options are currently:

### standard

Contains a selection of those columns considered most interesting by the service.

### default

Contains the 'standard' columns plus numeric "\_RAJ2000" and "\_DEJ2000" positional columns inserted by the service; if the query is a Cone Selection rather than All Rows, it also contains a column "\_r" inserted by the service giving the distance between the selected position and the row's position.

### all

Contains all the columns from the archived catalogue.

VizieR experts may fill in custom column requirements here by typing them into the selector box rather than choosing one of the predefined options, for instance `-out.add=_GLON,_GLAT` would add galactic coordinates to the standard set; see <http://vizier.u-strasbg.fr/doc/asu-summary.htx> for more details on VizieR hacking. (In fact, this trick can be used to add VizieR parameters unrelated to column selection as well).

The **Catalogue Selection** panel offers several different ways to identify which of the catalogues in the VizieR archive you want to query. In all cases you will be presented with a JTable of VizieR catalogues, and you must select one by clicking on the relevant row. You can sort within the displayed table by clicking on the column header. **Note:** for some of these options it is necessary to fill in the Row Selection panel before you can operate them (the controls will be disabled if no row selection has been made). That is because the catalogues listed will depend on the region you are going to query; VizieR is smart enough to know which catalogues have some coverage in the region in question. The options for catalogue selection are as follows:

#### By Category

You may select one or more terms from one or more of the presented lists of predefined keywords in the categories **Wavelength**, **Mission** and **Astronomy** to restrict the catalogues that you are interested in. How you select multiple entries from the same list is platform-dependent, but CTRL-click may work. When you have made your selections, hit the **Search Catalogues** button, and those catalogues in the categories you have identified, and with coverage in the region defined by the Row Selection panel, will be listed below the category selection panel. Select one of these by clicking on it. The **Sub-Table Details** checkbox controls whether the list displays only top-level VizieR catalogues (each of which may contain multiple tables) or entries for each table within each catalogue as well. The **Include Obsolete Tables** checkbox controls whether just the most current, or all versions of each catalogue are shown.

#### By Keyword

A **Keywords** text field is shown; you may enter a space-separated list of words which will be matched against catalogue names and descriptions. When you have entered the search terms, hit the **Search Catalogues** button, and those catalogues which match your terms, and with coverage in the region defined by the Row Selection panel, will be listed below the Keywords field. Select one of these by clicking on it. The **Sub-Table Details** checkbox controls whether the list displays only top-level VizieR catalogues (each of which may contain multiple tables) or entries for each table within each catalogue as well. The **Include Obsolete Tables** checkbox controls whether just the most current, or all versions of each catalogue are shown.

#### Surveys

A (fairly short) list of large surveys held by VizieR is presented in a table. An indication of the size of each, in terms of number of thousands of rows, is given. Select one of these by clicking on it.

#### Missions

A (fairly short) list of data holdings at VizieR originating from large missions is presented in a table. An indication of the size of each, in terms of number of thousands of rows, is given. Select one of these by clicking on it.

Depending on the type of catalogue search you make, various attributes of the catalogues in question will be listed in the table for selection:

#### Name

Unique VizieR identifier for the catalogue

#### Description

Short description of contents

**KRows**

Approximate number of thousands of rows

**Rows**

Approximate number of rows

**Tables**

Number of sub-tables contained within the Vizier catalogue

**Popularity**

A measure of the number of queries on the catalogue served by Vizier

**Density**

A measure of the number of sources per unit solid angle on the sky

**Wavelengths**

Keywords describing wavelength regimes covered

**Astronomy**

Keywords describing subject domain

When you have made your selection of rows, columns and catalogue you can hit the **OK** button and TOPCAT will attempt to contact the Vizier service to load the resulting table or tables. You can cancel a request in progress with the **Cancel** button.

CDS make the following request:

*If the access to catalogues with Vizier was helpful for your research work, the following acknowledgment would be appreciated: "This research has made use of the Vizier catalogue access tool, CDS, Strasbourg, France". The original description of the Vizier service was published in A&AS 143, 23 (2000).*

#### **A.6.10 GAVO Millennium Database Query**

**GAVO Millennium Run Query**

Window HaloSamples GalaxySamples Help

Base URL:

User:


Password:

SQL Query: 

```
select PROG.*
  from millimil..MPAHalo PROG,
       millimil..MPAHalo DES
 where DES.halold = 1
       and PROG.halold between DES.halold and DES.lastproge
```

OK

### GAVO load dialogue with an example query on the milli-Millennium database

This dialogue, selected from the Load Window's toolbar button () or the main **VO** menu, permits direct queries to the services provided by GAVO, the German Astrophysical Virtual Observatory. The main databases of general interest available through these services are the **Millennium** **Simulation** results, documented at <http://gavo.mpa-garching.mpg.de/Millennium/Help>.

To make a query, fill in the fields as required:

#### Base URL

This determines the database to be queried. The following options are available:

##### **Milli-Millennium** (<http://gavo.mpa-garching.mpg.de/Millennium>)

Public simulation results database, containing a fraction of the full simulation results. No username or password is required (the User and Password boxes will be disabled).

##### **Millennium** (<http://gavo.mpa-garching.mpg.de/MyMillennium>)

Full, protected simulation results database. The username and password must be filled in for these queries. Registration is by application to GAVO.

#### Other

Other GAVO-provided services can be accessed from this dialogue by entering the relevant base URL here -- users of these services will know what they are. They may or

may not require a username and password.

#### User

#### Password

Registration information if required (dependent on base URL as above).

#### SQL Query

The text of an SQL query on the database in question.

Then click the **OK** button, and the query will execute and load the results into TOPCAT. Clicking **Cancel** while it is in progress will stop it running.


The **HaloSamples** and **GalaxySamples** menus provide some examples of queries on the Milli-Millennium database (these have been copied from the GAVO query page). If you select one of these the **SQL Query** panel will be filled in accordingly.

Much more documentation, including tutorials and descriptions of the database schemas, is available on the GAVO website, at <http://gavo.mpa-garching.mpg.de/Millennium/Help>.

### A.6.11 BaSTI Theory Database Query

The screenshot shows the 'BaSTI Data Loader' dialog box with the 'Query' tab selected. The dialog has a title bar with 'BaSTI Data Loader' and standard window controls. Below the title bar is a menu bar with 'Window' and 'Help'. Under the menu bar are three icons: a pushpin, a question mark, and a close button. The main area contains instructions: 'Build your query using provided fields. Check the boxes to select the fields you want in the query results. Then press SUBMIT: the BaSTI tables that satisfy your query will be displayed in the Results tab. Keep the mouse over the text fields to get an hint on boundary values.' Below this are several fields for query construction. On the left, there are five rows of fields: 'Data Type' with a checked checkbox and a dropdown menu showing 'Track'; 'Scenario' with a checked checkbox and an empty dropdown; 'Type' with a checked checkbox and an empty dropdown; 'Mass Loss' with a checked checkbox and an empty dropdown; and 'Photometric System' with a checked checkbox and an empty dropdown. On the right, there are five rows of fields: 'Age [Gyr]' with an unchecked checkbox and two empty text boxes labeled 'min' and 'max'; 'Mass [MSun]' with a checked checkbox and two empty text boxes; 'Z' with a checked checkbox and two text boxes containing '.015' and '.02'; 'Y' with a checked checkbox and two empty text boxes; and '[Fe/H]' with a checked checkbox and two empty text boxes. Below these fields is a row with 'Mixture' with a checked checkbox and a dropdown menu showing 'Scaled Solar Model', and '[M/H]' with a checked checkbox and two empty text boxes. At the bottom left, it says 'Number of results (present query): 832'. At the bottom right are 'RESET' and 'SUBMIT' buttons. At the very bottom center is an 'OK' button.

#### BaSTI load dialog query tab

This dialogue, selected from the Load Window's toolbar button () or the main **VO** menu, allows direct queries to the BaSTI (*Bag of Stellar Tracks and Isochrones*) database hosted by the INAF-Teramo Astronomical Observatory. You can find more detailed documentation on the web site.

This load dialogue has two tabs: a **Query** tab to input search parameters, and a **Results** tab to display the results table with one row for each table resulting from the user's query.

The **Query** tab contains a set of parameters by which the query will be constrained, some aids to help the user while preparing the selection and two buttons. The **Reset** button simply clears query inputs and (if present) user's selections in the Results tab. The **Submit** button performs the query and switches the dialog to the Results tab. As an aid to allow a refined query without too much recursive steps, at the bottom center of the tab, a counter displays how many rows (i.e. resulting tables) the output will count. Remembering that the results will contain 30 rows at maximum, the user can then refine his/her constraints to reduce the number of results.

To do so the query helps the user in two ways mainly: automatically unselecting the unavailable parameters for a specific query (e.g. the mass range for an isochrone search) and displaying, for the ranged parameters, the value limits for that parameter, this happens just moving the mouse cursor over the input area.

Here follows a short description of the query parameters, for full details refer to the BaSTI main site.

#### **Data Type**

The type of simulation desired: isochrones, tracks or summary tables. This field is the only mandatory to submit a query.

#### **Scenario**

Stellar evolution scenario, i.e. canonical simulation or including overshooting calculations.

#### **Type**

Type of track simulation: normal or including the Asymptotic Giant Branch (AGB) simulation.

#### **Mass Loss**

Selects which value the Mass Loss parameter should have: 0.2 or 0.4 are the available choices.

#### **Photometric System**

A set of photometric systems is used to *colour* the stellar simulated tracks for comparison with observational data. This dropdown input allows for a selection within them.

#### **Mixture**

Stellar abundances mixture: scaled solar values or alpha enhanced.

#### **Age**

Age range, in values of Gyr, for the isochrones.

#### **Mass**

Mass range, in values of Solar Masses.

#### **Z**

Initial metal abundance range.

#### **Y**

Initial Helium abundance range.

#### **[Fe/H]**

Iron over Hydrogen logarithmic ratio.

#### **[M/H]**

Metals over Hydrogen logarithmic ratio.

Once the Query panel has been filled in, hit the **Submit** button. This will show the **Results** tab. This displays a table where each row refers to an available result from the BaSTI database as constrained by the user's query. On top of the table the number of results identified by the query is recalled. The user now can switch back to refine the query or selected (mouse clicking) the table/s he/she wants

to load into TOPCAT. Once the selection is ready (CTRL+click or SHIFT+click for multiple selections) pressing the **OK** button will load the tables into TOPCAT.

### **A.6.12 Example Tables**

Provided with TOPCAT are some example tables, which you can access in a number of ways. The simplest thing is to start up TOPCAT with the "-demo" flag on the command line, which will cause the program to start up with a few demonstration tables already loaded in.

You can also load examples in from the **Examples** menu in the Load Window however. This contains the following options:

#### **Load Example Table**

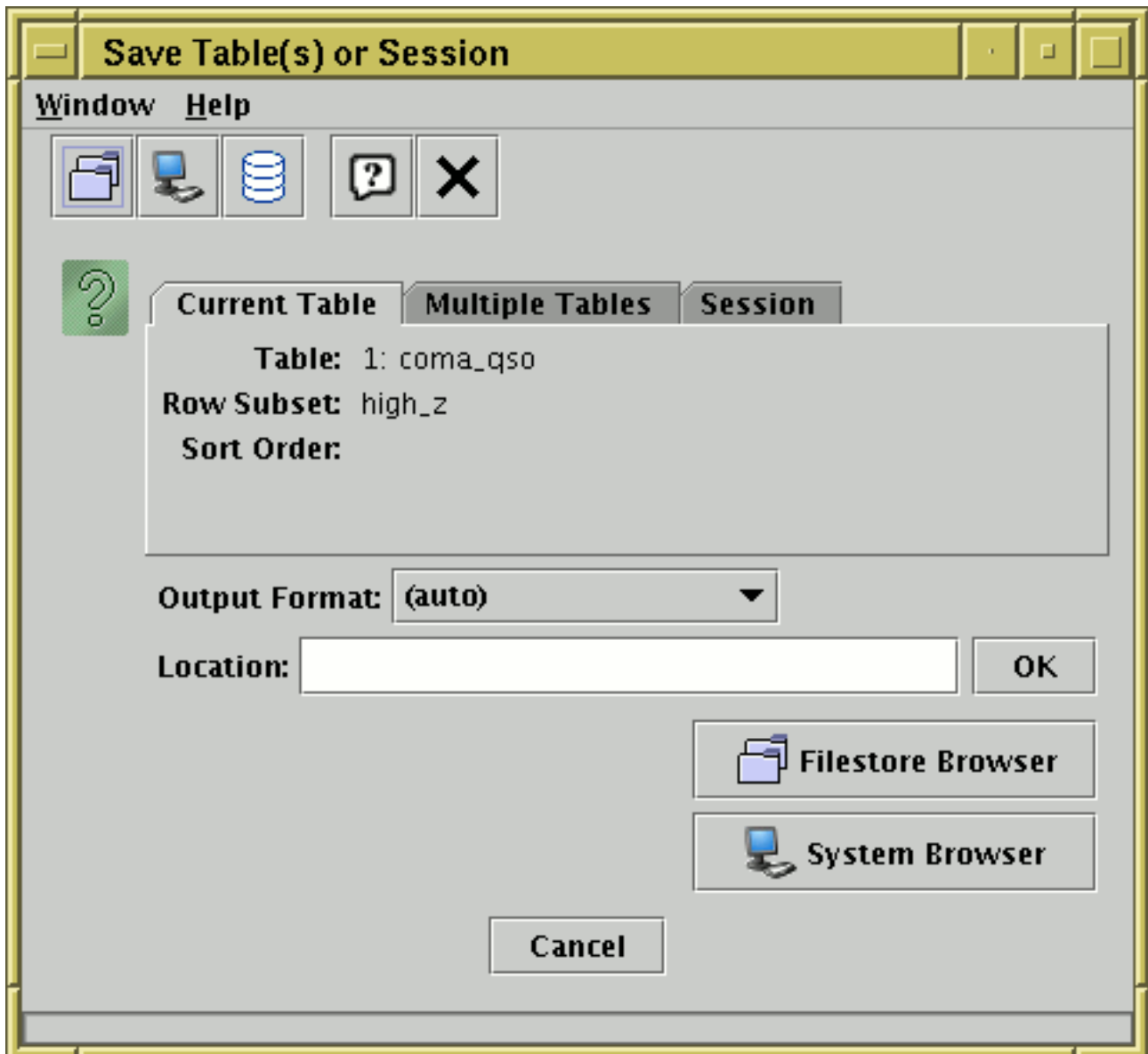
Loads in a single example table.

#### **Browse Demo Data**


Pops up a Hierarchy Browser (Appendix A.6.3) looking at a hierarchy of tables in different formats. This option is designed to show some of the organisational complexity which TOPCAT can handle when browsing tables.

Note these examples are a bit of a mixed bag, and are not all that exemplary in nature. They are just present to allow you to play around with some of TOPCAT's features if you don't have any real data to hand.

### **A.7 Save Window**



### Save Window

The Save Window is used to write tables out, and it is accessed using the **Save Table** button () in the Control Window's toolbar or **File** menu.

The window consists of two parts. At the top is the **Content Panel**, which is used to determine exactly what table or tables are going to be saved, and below it is the **Destination Panel**, which is used to determine where they will be saved to. These panels are described separately in the subsections below.

For large tables, a progress bar will be displayed indicating how near the save is to completion. It is not advisable to edit tables which are being saved during a save operation.

In some cases, saving the table to a file with the same name as it was loaded from can cause problems (e.g. an application crash which loses the data unrecoverably). In other cases, it's perfectly OK. The main case in which it's problematic is when editing an uncompressed FITS binary table on disk. TOPCAT will attempt to warn you if it thinks you are doing something which could lead to trouble; ignore this at your own risk.

If you save a table to a format other than the one from which it was loaded, or if some new kinds of




metadata have been added, it may not be possible to express all the data and metadata from the table in the new format. Some message to this effect may be output in such cases.

There is no option to compress files on output. You can of course compress them yourself once they have been written, but note that this is not usually a good idea for FITS files, since it makes them much slower to read (for TOPCAT at least).

### A.7.1 Content Panel

The Content Panel is the upper part of the save window, and it is used to select what table or tables you want to save. The options are described in the following subsections.

#### A.7.1.1 Current Table

The **Current Table** save option saves the table currently selected in the Control Window. What is written is the Apparent Table corresponding to the current table, which takes into account any modifications you have made to its data or appearance this session. The current **Row Subset** and **Sort Order** are displayed in this window as a reminder of what you're about to save. Information about Row Subsets themselves and hidden columns will not be saved, though you can save information about subset membership by creating new boolean columns based on subsets using the "To Column" button (  ) from the Subsets Window. Alternatively you could use the Session

Save option described below.

#### A.7.1.2 Multiple Tables

The **Multiple Tables** save option allows you to save multiple tables to the same file. If a FITS or VOTable based output format is used, this means that when the file is loaded back into TOPCAT, all the saved tables will be individually loaded in.

The list of loaded tables is shown, with a column of checkboxes on the left. By default these are selected, but you can select or unselect them by clicking. When the save is performed, only those tables with the checkbox selected will be saved.

As with the Current Table save, it is the Apparent Table in each case which is saved, so that only unhidden columns, and rows in the Current Subset will be included. On the right hand side of the table is information to remind you which row subset and ordering will be saved.

#### A.7.1.3 Session

The **Session** save option allows you to save much of the information about the loaded tables in your current TOPCAT session. Unlike the **Current Table** or **Multiple Tables** options, the whole of each loaded table, along with other information about your use of it, is saved, rather than just the Apparent Table. The saved items include:

- All rows, not just those in the current Row Subset
- All visible and hidden columns, along with their order and visibility status
- All defined Row Subsets
- The current Sort Order
- The current Row Subset
- The label
- Whether row selection causes a SAMP broadcast

The list of loaded tables is shown, with a column of checkboxes on the left. By default these are

selected, but you can select or unselect them by clicking. When the save is performed, only those tables with the checkbox selected will be saved.

The tables are saved as a multi-table fits-plus (recommended) or VOTable file. This is a normal multi-table file in that any FITS or VOTable reader can read the tables it contains, but it contains some specialised metadata encoding information of special significance to TOPCAT, marking it as a saved session. The upshot is that if you save a file in this way and then load it back into TOPCAT, the tables it loads will appear in very much the same state as when you saved them, in terms of defined and current subsets, row order, visible and invisible columns, and so on. If you process it with some application other than TOPCAT, it will look quite like the table you saved, but with some additional data and metadata.




Note however that the reloaded state is not identical to that before saving. One important difference is that columns and row subsets which were defined algebraically are saved and restored as fixed values, so it is no longer possible to edit the expressions. Another is that only state belonging to tables is saved, so that for instance plots will not be restored. Table activation actions are not saved either. It is possible that some of this may be changed in future releases.

### A.7.2 Destination Panel

The Destination Panel is the lower part of the save window, and is used to select where and how the table or tables should be saved.

The **Output Format** selector is used to choose the format in which the table will be written from one of the supported output formats (Section 4.1.2). The available selections are somewhat different depending on what it is you are saving; for instance some formats cannot accommodate multiple tables, so these formats are not offered for the Multiple Table save. If the "(auto)" option is used, an attempt is made to guess the format based on the filename given; for instance if you specify the name "out.fits" a FITS binary table will be written.

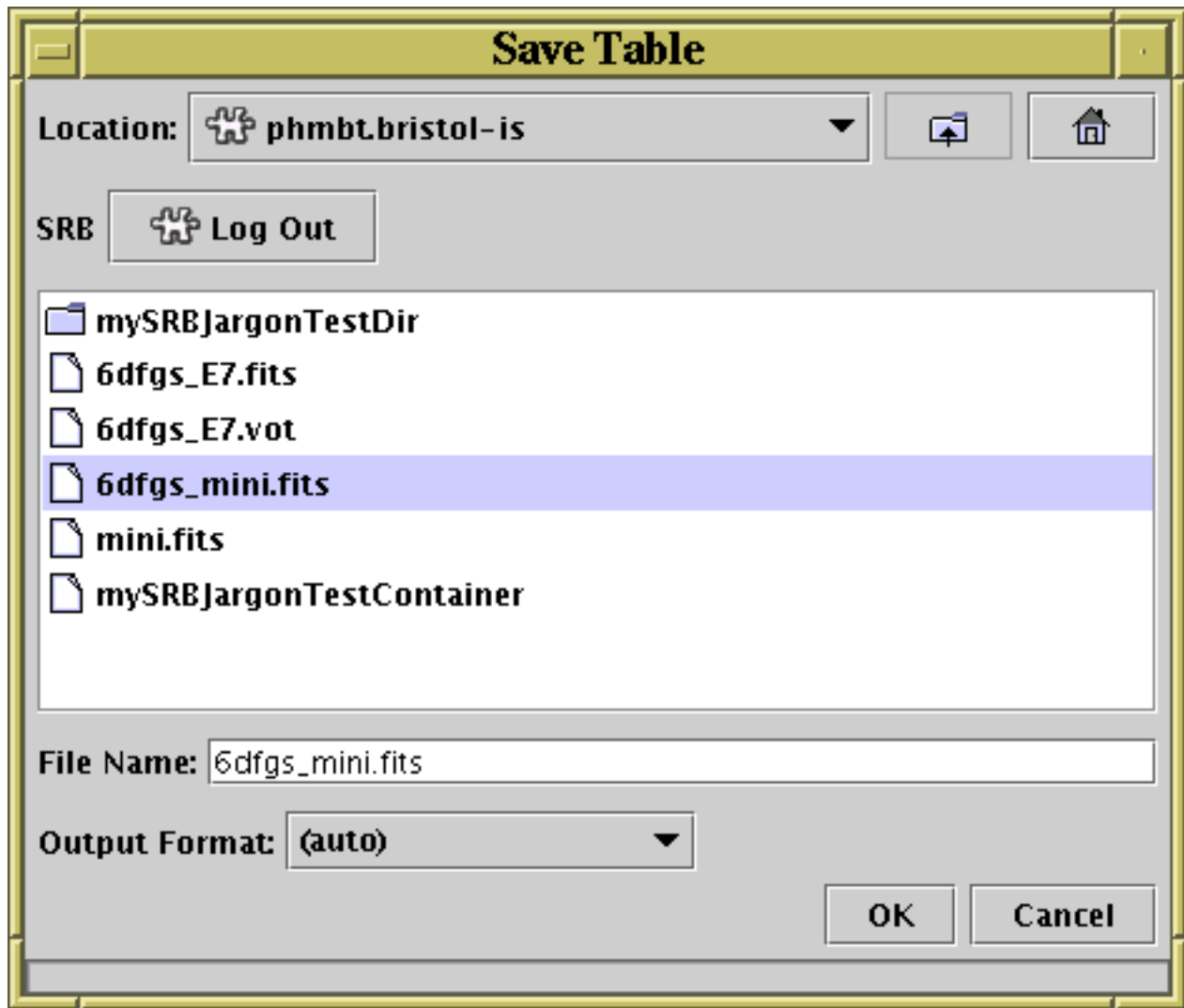
You can specify the location of the output table in these ways, which are described in the following sections:

- Type in the location directly in the Output Location field (Appendix A.7.2.1) as a filename or a URL as described in section Section 4.2
- Use the **Filestore Browser** button or toolbar button () to get a browser (Appendix A.7.2.2) that shows you local and remote files
- Use the **System Browser** button or toolbar button () to get a native browser (Appendix A.7.2.3) that uses your OS's standard file saving interface
- Use the **SQL table** toolbar button () to get the SQL Output Dialogue (Appendix A.7.2.4)

#### A.7.2.1 Enter Location

You can specify where to save a table by typing its location directly into the **Output Location** field of the Save Table window. This will usually be the name of a new file to write to, but could in principle be a URL or a SQL specifier.

#### A.7.2.2 Filestore Browser



### Filestore Browser for table saving


By clicking the **Browse Filestore** button in the **Save Table** window, you can obtain a browser which will display the files in a given directory.

The browser initially displays the current directory, but this can be changed by typing a new directory into the **File Name** field, or moving up the directory hierarchy using the selector box at the top, or navigating the file system by clicking the up-directory button or double-clicking on displayed directories. The initial default directory can be changed by setting the `user.dir` system property (Section 10.2.3).

The browser can display files in remote filestores such as on MySpace or SRB servers; see the section on the load filestore browser (Appendix A.6.1) for details.

To save to an existing file, select the file name and click the **OK** button at the bottom; this will overwrite that file. To save to a new file, type it into the **File Name** field; this will save the table under that name into the directory which is displayed. You can (re)set the format in which the file will be written using the **Output Format** selector box on the right (see Section 4.1.2 for discussion of output formats).

### A.7.2.3 System Browser

By clicking the **System Browser** button or toolbar button () in the Save Window, you can use

your Operating System's native file browser to decide where to save a file. What this looks like is entirely dependent on your OS; it may or may not contain system-specific features like shortcuts to commonly-used directories.

Since TOPCAT has no control over the way this dialogue looks, it cannot contain the **Output Format** selector, and certain other things such as save cancel may not work as well as for the other dialogue types. To select the output table format, you will need to use the selector in the **Save Window**.

#### A.7.2.4 SQL Output Dialogue

#### SQL table writing dialogue

If you want to write a table to an SQL database (Section 4.1.2.9), you can use a specialised dialogue to specify the table destination by clicking the **SQL Table** button in the **Save Table** window.

This provides you with a list of fields to fill in which define the new table to write, as follows:

##### Protocol

The name of the appropriate JDBC sub-protocol. This is defined by the JDBC driver that you are using, and is for instance "mysql" for MySQL's Connector/J driver or "postgresql" for PostgreSQL's JDBC driver.

##### Host

The hostname of the machine on which the database resides (may be "localhost" if the database is local).

##### Database name

The name of the database.

##### New table name

The name of a new table to write into the given database. Subject to user privileges, this will overwrite any existing table in the database which has the same name, so should be used with care.

##### User name

The username under which you wish to access the database. This is not strictly necessary if there is no access control for the database in question.

##### Password

The password for the given username. Again, whether this is necessary depends on the access

policy of the database.

### Write Mode

Options for writing the table to the database. Choose one of:

#### **create**

A new table is created for the output. If one with the given name already exists, an error results.

#### **dropcreate**

A new table is created for the output. If one with the given name already exists, it is dropped first.

#### **append**

The results are appended to an existing table with the given name. If the table has the wrong number or types of columns, an error results.

There are a number of criteria which must be satisfied for SQL access to work within TOPCAT (installation of appropriate drivers and so on) - see the section on JDBC configuration (Section 10.3). If you don't take these steps, this dialogue may be inaccessible.

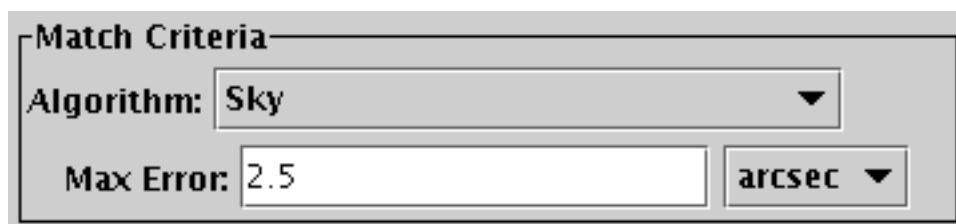
## A.8 Match Windows

This section documents the windows used to crossmatch rows either between two or more local tables or within a single table. This topic is discussed in more detail in Section 5. Windows for other kinds of joins are described elsewhere: concatenation in Appendix A.10.2 and matching with remote tables via VO services in Appendix A.9 and against VizieR and SIMBAD via the CDS X-Match service in Appendix A.10.1.

The next subsection describes the features which are common to the different types of match window, and the following subsections detail the operation of each distinct match window (internal, pair and multi-table).

### A.8.1 Common Features

#### A.8.1.1 Match Criteria



The screenshot shows a window titled "Match Criteria". Inside, there is a label "Algorithm:" followed by a dropdown menu currently showing "Sky". Below this, there is a label "Max Error:" followed by a text input field containing "2.5". To the right of the text field is a unit selector dropdown menu currently showing "arcsec".

**Match Criteria panel. The details will differ depending on what match type is chosen.**

The match criteria box allows you to specify what counts as a match between two rows. It has two parts. First, you must select one of the options in the **Algorithm** selector. This effectively selects the coordinate space in which rows will be compared with each other. Depending on the chosen value, a number of fields will be displayed below, which you must fill in with values that determine how close two rows have to be in terms of a metric on that space to count as a match.

The following match types (algorithm values) are offered:

#### **Sky**

Comparison of positions on the celestial sphere. In this case you will need to specify columns

giving Right Ascension and Declination for each table participating in the match. The Max Error value you must fill in is the maximum separation of matched points around a great circle.

### **Sky with Errors**

The matching is like that for the **Sky** option above, but an error radius (positional uncertainty) is given for each row in the input tables, rather than just a single value for the whole match. Along with the Right Ascension and Declination columns, you also specify an Error column which gives the error radius corresponding to that position. Two rows are considered to match when the separation between the two RA,Dec positions is no larger than the sum of the two Error values for the corresponding rows. The Scale value should be set to a rough average of the per-row errors. It is used only to set a sensible default for the Healpix-k tuning parameter, and its value does not affect the result. If Healpix-k is set directly (see Appendix A.8.1.3), Scale is ignored. **Note:** the semantics of this matcher have changed slightly since version 3.8 of TOPCAT and earlier. In earlier versions the single parameter was named Max Error and provided an additional constraint on the maximum accepted separation between matched objects. For most uses, the old and new behaviours are expected to give the same results, but in cases of difference, the new behaviour is more likely what you want.

### **Sky Ellipses**

Compares elliptical regions on the sky for overlap. You will need to specify columns giving the central position, major and minor radii, and position angle of the ellipse. Two rows are considered to match if there is any overlap between the ellipses. The goodness of match is a normalised generalisation of the symmetrical case used by the **Sky with Errors** option above. The position angle is measured from north to the semi-major axis, in the direction towards the positive RA axis. The Scale value should be set to a rough average of the major radii; It is used only to set a sensible default for the Healpix-k tuning parameter, and its value does not affect the result. If Healpix-k is set directly (see Appendix A.8.1.3), Scale is ignored. The calculations are approximate since in some cases they rely on projecting the ellipses onto a Cartesian tangent plane before evaluating the match, so for larger ellipses the criterion will be less exact. For objects the size of most observed stars or galaxies, this approximation is not expected to be problematic.

### **Sky 3D**

Comparison of positions in the sky taking account of distance from the observer. In this case you will need to specify columns giving Right Ascension and Declination in angular units, as well as distance from the origin in arbitrary units for each table participating in the match. The Error value is a maximum separation in Cartesian space of matched points in the same units as the radial distance.

### **Exact Value**

Requires exact matching of values. In this case you will need to specify the column containing the match key for each table participating in the match; this might typically be an object name or index number. Two rows count as matching if they have exactly the same entry in the specified field, except rows with a null value in that column, which don't match any other row. Note that the values must also be of the same type, so for instance a Long (64-bit) integer value will not match an Integer (32-bit) value.

### **N-dimensional Cartesian**

Comparison of positions in an isotropic N-dimensional Cartesian space. In this case you will need to specify N columns giving coordinates for each table participating in the match. The Error value is the maximum spatial separation of matched points. Currently the highest dimensionality you can select is 3-d - does anyone want a higher number?

### **N-dimensional Cartesian, Anisotropic**

Comparison of positions in an N-dimensional Cartesian space with an anisotropic metric. In this case you will need to specify N columns giving coordinates for each table participating in the match, and an error distance for each of these dimensions. Points P1 and P2 are considered to match if P2 falls within the ellipsoid with radii given by the error distances, centered on P1.

This kind of match will typically be used for non-'spatial' spaces, for instance (magnitude,redshift) space, in which the metrics in different dimensions are not related to each other. Currently the highest dimensionality you can select is 4-d - does anyone want a higher number?

### **N-dimensional Cuboids**

This matching is like that for **N-dimensional Cartesian** above, but points are considered to match if they fall within the same cuboid rather than ellipsoid. The error values are the half-axis lengths of the cuboid, like rectangular "radii". This kind of match is suitable for grouping items into regularly-spaced pixels, though it's not a very efficient way of doing that.

### **N-dimensional Cartesian with Errors**

The matching is like that for the **N-dimensional Cartesian** option above, but an error radius (positional uncertainty) is given for each row in the input tables, rather than just a single value for the whole match. Along with the Cartesian coordinate columns, you also specify an Error column which gives the error radius corresponding to that position. Two rows are considered to match when the separation between the two positions is no larger than the sum of the two Error values for the corresponding rows. The Scale parameter should be approximately the characteristic size of the per-object error values. Its value, in conjunction with the Bin Factor tuning parameter, affects the performance of the match but not the result.

### **2-d Cartesian Ellipses**

Compares elliptical regions in the plane for overlap. You will need to specify columns giving the central position, major and minor radii, and orientation angle of the ellipse. Two rows are considered to overlap if there is any overlap between the ellipses. The goodness of match is a generalisation of the symmetrical case used by the **2-d Cartesian with Errors** option above. The orientation angle is measured anticlockwise from the X-axis to the ellipse major axis. The Scale parameter should be set to a rough average of the major radii. Its value, in conjunction with the Bin Factor tuning parameter, affects the performance of the match but not the result.

### **Sky + X**

Comparison of positions on the celestial sphere with an additional numeric constraint. This is a combination of the Sky and 1-d Cartesian matches above, so the columns you need to supply are RA, Dec and one extra, and the errors are angular separation on the sky and the error in the extra column. A match is registered if it matches in both of the constituent tests. You could use this for instance to match objects which are both close on the sky and have similar luminosities.

### **Sky + XY**

Comparison of positions on the celestial sphere with an additional 2-d anisotropic Cartesian constraint. This is a combination of the Sky and 2-d Anisotropic Cartesian matches above, so the columns you need to supply are RA, Dec and two extra, and the errors are angular separation on the sky and the error radii corresponding to the extra columns. A match is registered if it matches in both of the constituent tests. You could use this for instance to match objects which are both close on the sky and have similar luminosities and redshifts.

### **HTM**

Performs sky matching in just the same way as the **Sky** option above, but using a different algorithm (pixelisation of the celestial sphere is performed using the Hierarchical Triangular Mesh rather than the HEALPix scheme). The results in both cases should be identical, but HTM is much slower. Hence, this option is only useful for debugging. It may be withdrawn in future releases.

Depending on the match type, the units of the error value(s) you enter may be significant. In this case, there will be a unit selector displayed alongside the entry box. You must choose units which are correct for the number you enter.

More information is available in the GUI as a tooltip by hovering with the mouse pointer over the

field in question.

See Appendix A.8.1.3 for information on optional *tuning parameters* which are sometimes displayed in this panel.

The matching framework is capable of performing even more complicated types of match, for instance Sky + 6-dimensional anisotropic Cartesian. These are not offered purely to avoid complicating the GUI. More flexible matching requirements in this in and other respects can be achieved by using the matching commands in STILTS instead.

### A.8.1.2 Column Selection Boxes

**Table 1**

Table: 8: mgc.fits ▼

RA column: MGC\_ALPHA\_J2000 ▼ degrees ▼

Dec column: MGC\_DELTA\_J2000 ▼ degrees ▼

**Column Selection Box.** The details will differ depending on what match type is chosen.

The column selection boxes allow you to select which of the columns in the input tables will provide the data (the coordinates which have to match). For each table you must select the names of the required columns; the ones you need to select will depend on the match criteria you have chosen.

For some columns, such as Right Ascension and Declination in sky matches, units are important for the columns you select. In this case, there will be a selector box for the units alongside the selector box for the column itself. You must ensure that the correct units have been selected, or the results of the match will be rubbish.

In some cases these column and/or unit selectors may have a value filled in automatically (if the program thinks it can guess appropriate ones) but you should ensure that it has guessed correctly in this case. Only suitable columns are available for choosing from these column selectors; in most cases this means numeric ones.

Instead of selecting a column name from the list provided in these boxes, you may type in an expression. This can be a column name, or an algebraic expression based on column names, or even a constant. The expression language syntax is described in Section 7.

### A.8.1.3 Tuning

This subsection describes the use of two toolbar buttons available in the match windows:



#### Tuning Parameters

Displays *tuning parameters* alongside match parameters in the Match Criteria panel.



#### Full Profiling

Increases the amount of timing and memory use information displayed in the Logging Panel.


The parameters such as Max Error visible by default in the Match Criteria panel define what counts



as a match and must be filled in for the match to proceed.

Optionally, you can see and adjust another set of parameters known as *Tuning parameters*. These will not affect the result of the match, but may affect its performance, in terms of how much CPU time and memory it takes. Most of the time, you can forget about this option, since TOPCAT attempts to pick reasonable defaults, but if your match is very slow (especially if it's unexpectedly slow given the sizes of the tables involved), or if it's failing with messages about running out of memory, then adjusting the tuning parameters may help.

To view the tuning parameters, use the **Tuning Parameters** () toolbar button or menu item.

This will add display of tuning parameters to the match parameters in the Match Criteria panel. Suggested values are filled in by default, and may be affected by the match parameters that you fill in; you can play around with different values and see the effect on match performance. If you do this, it is useful to use also the **Full Profiling** () toolbar button to turn on full profiling. This

will cause additional information on the amount of CPU time and memory used by each step to be displayed in the logging panel at the bottom. There is a small penalty in CPU time required to gather this information, which is one reason it is not turned on by default.

What tuning parameters are available depends on the match type you have chosen. Some additional description is available as tooltips if you hover over the query field. In most cases, they are one or other of the following:

#### **HEALPix $k$**

Used for sky-like matches.  $k$  is an integer value which determines the size of pixels into which the celestial sphere is decomposed for binning rows. The legal range is between 0 (corresponding to a pixel size around 60 degrees) and 20 (around 0.2 arcsec). In HEALPix language,  $k$  is  $\log_2(N_{\text{side}})$ .

#### **Scale Factor**

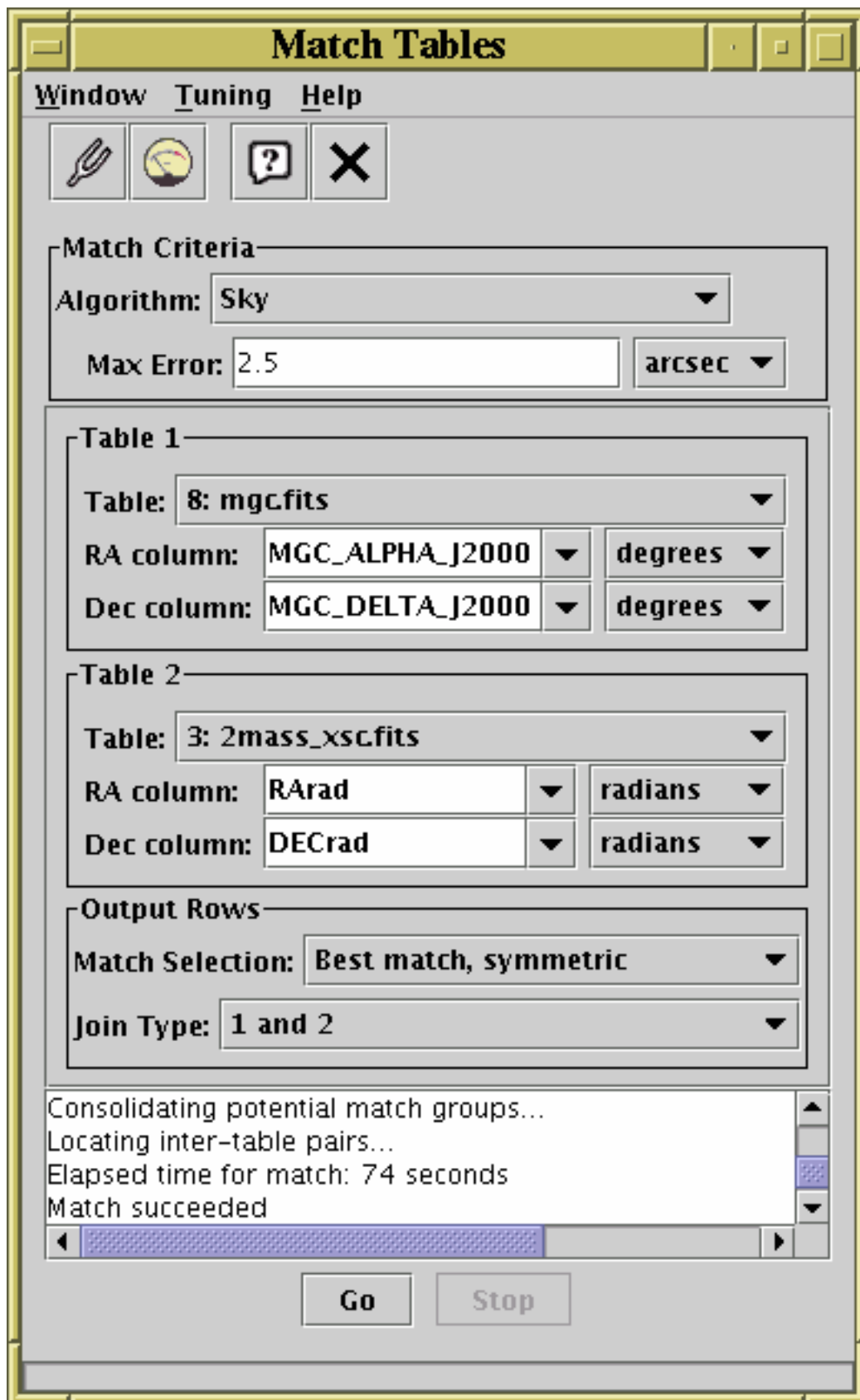
Used for Cartesian-like matches. The scale factor is a floating point value which determines the size of the notional N-dimensional pixels into which the space is decomposed for binning rows, as a multiple of the match error. The smallest legal value is 1.

In either case, the number of rows per bin should be not too large, and not too small (though exactly what that means in quantitative terms is another matter). Larger bins/pixels generally mean less memory use and more CPU time, though that's not always the case.


Even if you happen to have a good understanding of how the matching algorithm works (and you probably don't), this kind of tuning is a bit of a black art, and depends considerably on the details of the particular match. In some cases however it is quite possible to improve the time taken by a match, or the size of table which can be matched in a given amount of memory, by a factor of several. If you want to try to improve performance, try the default, adjust the tuning parameters slightly, look at how the performance changes by examining the logging output, and maybe repeat to taste.

Another thing which can affect speed and memory usage is whether your Java Virtual Machine is running in 32-bit or 64-bit mode. There are pros and cons of each - sometimes one will be better, sometimes the other. If you need to improve performance, experiment!

### **A.8.2 Pair Match Window**



### Pair Match Window

The Pair Match Window allows you to join two tables together side-by-side, aligning rows by matching values in some of their columns between the tables. It can be obtained using the **Pair Match** (  ) button in the Control Window toolbar or **Joins** menu.

In a typical scenario you might have two tables each representing a catalogue of astronomical objects, and you want a third table with one row for each object which has an entry in both of the original tables. An object is defined as being the same one in both tables if the co-ordinates in both rows are "similar", for instance if the difference between the positions indicated by RA and Dec columns differ by no more than a specified angle on the sky. Matching rows to produce the join requires you to specify the criteria for rows in both tables to refer to the same object and what to do when one is found - the options are discussed in more detail in Section 5.2.

The result is created from the Apparent versions of the tables being joined, so that any row subsets, hidden columns, or sorts currently in force will be reflected in the output. Progress information on the match, which may take a little while, is provided in the logging window and by a progress bar at the bottom of the window. When it is completed, you will be informed by a popup window which indicates that a new table has been created. This table will be added to the list in the Control Window and can be examined, manipulated and saved like any other. In some cases, some additional columns will be added to the output table which give you more information about how it has progressed (see Appendix A.8.2.1).

The Match Window provides a set of controls which allow you to choose how the match is done and what the results will look like. It consists of these main parts:

#### Match Criteria box

Allows you to define what counts as a match between two rows.

#### Column Selection boxes

Allows you to select which tables are to be joined and which columns in them supply the matching coordinates.

#### Output Rows selector



Allows selection of which rows are to be included in the output table (for instance whether only those rows matching in both tables should be output or not).

#### Log window

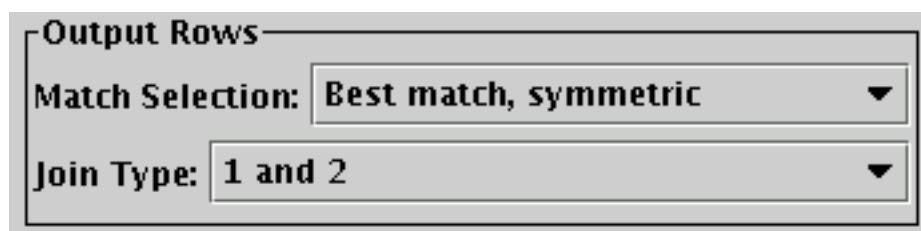
Reports on progress as the match is taking place. The progress bar at the bottom of the window also provides an indication of how far through each stage processing has got.

#### Control buttons

The **Go** button starts the search when you are happy with the selections that you have made, and the **Stop** button interrupts it midway if you decide you no longer want the results (closing the Match Window also interrupts the calculation).

For information on the **Tuning Parameters** () and **Full Profiling** () toolbar buttons, see Appendix A.8.1.3.

### A.8.2.1 Output Rows Selector Box



#### Pair match Output Rows Selector Box

When the match is complete a new table will be created which contains rows determined by the matches which have taken place. The **Output Rows** selector box allows you to choose on what

basis the rows will be included in the output table as a function of the matches that were found.

In all cases each row will refer to only one matched (or possibly unmatched) "object", so that any non-blank columns in a given row come from only rows in the input tables which match according to the specified criteria. However, you have two choices to make about which rows are produced.

The **Match Selection** selector allows you to choose what happens when a row in one table can be matched by more than one row in the other table. There are four options:

**All matches**

Every match between the two tables is included in the result. Rows from both of the input tables may appear multiple times in the result.

**Best match, symmetric**

The best pairs are selected in a way which treats the two tables symmetrically. Any input row which appears in one result pair is disqualified from appearing in any other result pair, so each row from both input tables will appear in at most one row in the result.

**Best match for each Table 1 row**

For each row in table 1, only the best match from table 2 will appear in the result. Each row from table 1 will appear a maximum of once in the result, but rows from table 2 may appear multiple times.

**Best match for each Table 2 row**

For each row in table 2, only the best match from table 1 will appear in the result. Each row from table 2 will appear a maximum of once in the result, but rows from table 1 may appear multiple times.

The "best" match in these options generally means "closest" - it is the one with the lowest match score, where the definition of this score is determined by the match criteria you have selected. The differences between the various **Best match...** options are a bit subtle. In cases where it's obvious which object in each table is the best match for which object in the other, choosing between these options will not affect the result. However, in crowded fields (where the distance between objects within one or both tables is typically similar to or smaller than the specified match radius) it will make a difference. In this case one of the asymmetric options is usually most appropriate, but you'll need to think about which of them suits your requirements. The performance (time and memory usage) of the match may also differ between these options, especially if one table is much larger than the other.

The **Join Type** selector allows you to choose what output rows result from a match in the input tables.

**1 and 2**

The output table contains only rows which have an entry from both of the input tables, so that every output row represents an actual matched pair. This corresponds to an SQL *inner join*.

**All from 1**

All of the matched rows are present in the output as for **1 and 2**, but additionally the unmatched rows from the first table are present with the columns from the second table blank. This corresponds to an SQL *left outer join*.

**All from 2**

As for **All from 1** but the other way round. This corresponds to an SQL *right outer join*.

**1 or 2**

Every row, matched and unmatched, from both of the input tables appears in the output. This is the union of rows from **All from 1** and **All from 2**. This corresponds to an SQL *full outer join*.

**1 not 2**

This presents all the rows in the first table which do *not* have matches in the second table.

Consequently, it only contains columns from the first table, since all the entries from the second one would be blank in any case.

## 2 not 1

The same as **1 not 2** but the other way round.

## 1 xor 2

The "exclusive or" of the match - the output only contains rows from the first table which don't have matches in the second table and vice versa. It is the union of **1 not 2** and **2 not 1**.

Note that the choices of which **Match Selection** and **Join Type** to use are somewhat interlinked, and some combinations may not be very meaningful.

In most cases (all the above except for **1 not 2** and **2 not 1**), the set of columns in the output table contains all the columns from the first table followed by all the columns from the second table. If this causes a clash of column names, offending columns will be renamed with a trailing "\_1" or "\_2". Depending on the details of the match however, some additional useful columns may be added:

### Match Score

For rows that represent a match, a numeric value representing how good the match was will usually be present. This is typically a separation in real or notional space - for instance for a **Sky** match it is the distance between the two matched celestial positions in arcseconds along a great circle. It will always be greater than or equal to zero, and a smaller value represents a better match. The name and exact meaning of this column depends on the match criteria - examine its description in the Columns Window for details.

### GroupSize, GroupID

If some of the rows match more than once (which may happen for any of the Match Selection options above apart from **BEST**), two columns named **GroupID** and **GroupSize** will be added. These allow you to identify which matches are multiple. In the case of rows which represent a unique match, they are blank. But for rows which represent a set of multiple matches, the GroupSize value tells you how many rows participate in this match, and the GroupID value is an integer which is the same for all the rows which participate in the same match. So if you do a sort on the GroupID value, you'll see all the rows in the first non-unique match group together, followed by all the rows in the second non-unique group... and after them all the unique matches.

Here is an example. If your input tables are these:

X	Y	Vmag
---	---	----
1134.822	599.247	13.8
659.68	1046.874	17.2
909.613	543.293	9.3

and

X	Y	Bmag
---	---	----
909.523	543.800	10.1
1832.114	409.567	12.3
1135.201	600.100	14.6
702.622	1004.972	19.0

then a Cartesian match of the two sets of X and Y values with an error of 1.0 using the **1 and 2** option would give you a result like this:

X_1	Y_1	Vmag	X_2	Y_2	Bmag	Separation
---	---	----	---	---	----	-----

1134.822	599.247	13.8	1135.201	600.100	14.6	0.933
909.613	543.293	9.3	909.523	543.800	10.1	0.515

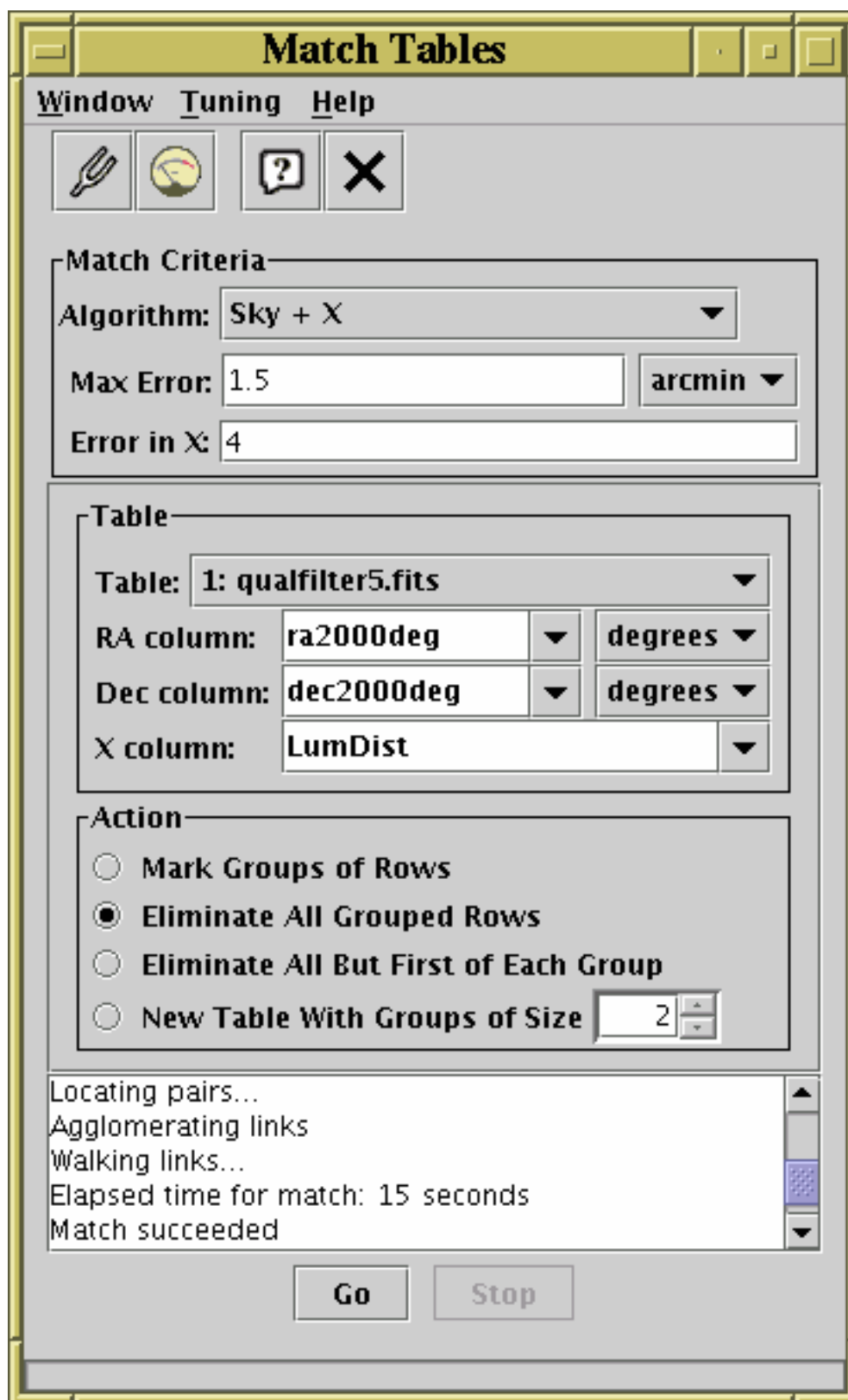
using **All from 1** would give you this:

X_1	Y_1	Vmag	X_2	Y_2	Bmag	Separation
---	---	----	---	---	----	-----
1134.822	599.247	13.8	1135.201	600.100	14.6	0.933
659.68	1046.874	17.2				
909.613	543.293	9.3	909.523	543.800	10.1	0.515


and **1 not 2** would give you this:

X	Y	Vmag
-	-	----
659.68	1046.874	17.2

### A.8.3 Internal Match Window



### Internal Match Window

The Internal Match Window allows you to perform matching between rows of the same table, grouping rows that have the same or similar values in specified columns and producing a new table as a result. It can be obtained by using the **Internal Match** (  ) item in the Control Window's

**Joins** menu.

You might want to use this functionality to remove all rows which refer to the same object from an

object catalogue, or to ensure that only one entry exists for each object, or to identify groups of several "nearby" objects in some way.

The result is created from the Apparent versions of the tables being joined, so that any row subsets, hidden columns, or sorts currently in force will be reflected in the output. Progress information on the match, which may take some time, is provided in the logging window and by a progress bar at the bottom of the window. When it is completed, you will be informed by a popup window which indicates that a new table has been created. This table will be added to the list in the Control Window and can be examined, manipulated and saved like any other.

The window has the following parts:

#### **Match Criteria box**

Allows you to define what counts as a match between two rows.

#### **Column Selection box**

Allows you to select which table to operate on and which columns supply the matching coordinates.

#### **Match Action box**



Allows you to select what will be done (what new table will be created) when the matching groups of rows have been identified.

#### **Log Panel**

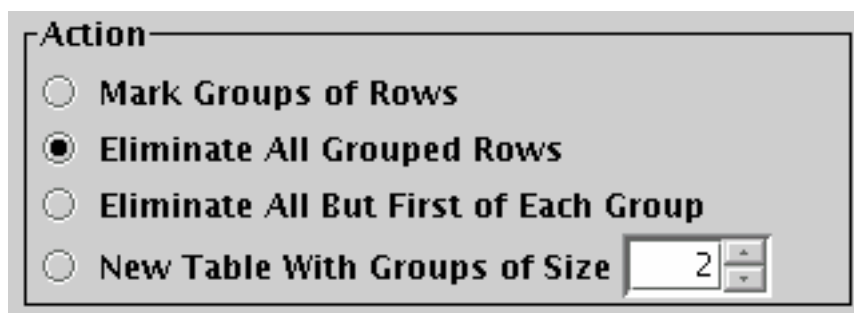
Displays progress as the match is taking place. The progress bar at the bottom of the window also provides an indication of how far through each stage processing has got.

#### **Control buttons**

The **Go** button starts the search when you are happy with the selections that you have made, and the **Stop** button interrupts it midway if you decide you no longer want the results (closing the Match Window also interrupts the calculation).

For information on the **Tuning Parameters** () and **Full Profiling** () toolbar buttons, see Appendix A.8.1.3.

### **A.8.3.1 Internal Match Action Box**



#### **Internal Match Action Box**

The Internal Match Action box gives a list of options for what will happen when an internal match calculation has completed. In each case a new table will be created as a result of the match. The options for what it will look like are these:

#### **Mark Groups of Rows**

The result is a table the same as the input table but with two additional columns: **GroupID** and **GroupSize**. Each group of rows which matched is assigned a unique integer, recorded in the GroupId column, and the size of each group is recorded in the GroupSize column. Rows which



don't match any others (singles) have null values in both these columns. So for example by sorting the resulting table on GroupID you can group rows that match next to each other; or by sorting on GroupSize you can see all the pairs, followed by all the triples, ...

You can use this information in other ways, for instance if you create a new Row Subset using the expression `"GroupSize == 5"` you could select only those rows which form part of 5-object clusters.

### **Eliminate All Grouped Rows**

The result is a new table containing only "single" rows, that is ones which don't match any other rows in the table according to the match criteria. Any rows which match are thrown out.


### **Eliminate All But First of Each Group**

The result is a new table in which only one row (the first in the input table order) from each group of matching ones is retained. A subsequent internal match with the same criteria would therefore show no matches.

### **New Table With Groups of Size N**

The result is a new "wide" table consisting of matched rows in the input table stacked next to each other. Only groups of exactly N rows in the input table are used to form the output table; each row of the output table consists of the columns of the first group member, followed by the columns of the second group member and so on. The output table therefore has N times as many columns as the input table. The column names in the new table have "\_1", "\_2", ... appended to them to avoid duplication.

## **A.8.4 Multiple Match Window**

The Multiple Match Window allows you to perform matching between more than two tables. It can be obtained by using the **Triple Match**, **Quadruple Match** etc (  ) items in the Control

Window's **Joins** menu.

The result is created from the Apparent versions of the tables being joined, so that any row subsets, hidden columns, or sorts currently in force will be reflected in the output. Progress information on the match, which may take some time, is provided in the logging window and by a progress bar at the bottom of the window. When it is completed, you will be informed by a popup window which indicates that a new table has been created. This table will be added to the list in the Control Window and can be examined, manipulated and saved like any other.

The window has the following parts:

#### **Match Criteria box**

Allows you to define what counts as a match between two rows.

#### **Column Selection box**

Allows you to select which table to operate on and which columns supply the matching coordinates.

#### **Match Action box**

Allows selection of which rows are to be included in the output table (for instance whether only those matches which appear in all input tables should result in output rows).



#### **Log window**

Reports on progress as the match is taking place. The progress bar at the bottom of the window also provides an indication of how far through each stage processing has got.

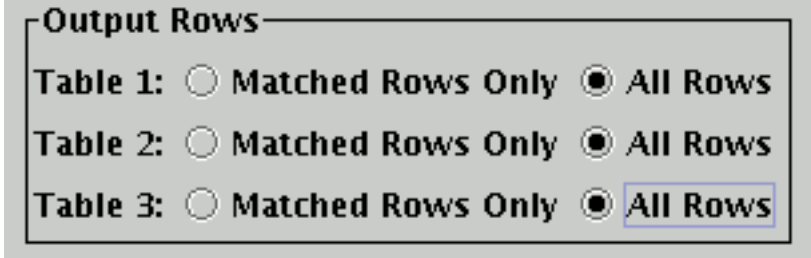
#### **Control buttons**

The **Go** button starts the search when you are happy with the selections that you have made, and the **Stop** button interrupts it midway if you decide you no longer want the results (closing

the Match Window also interrupts the calculation).

For information on the **Tuning Parameters** () and **Full Profiling** () toolbar buttons, see Appendix A.8.1.3.

#### A.8.4.1 Multiple Match Action Box



**Output Rows**

Table 1: ☐ Matched Rows Only ☒ All Rows

Table 2: ☐ Matched Rows Only ☒ All Rows

Table 3: ☐ Matched Rows Only ☒ All Rows

#### Multiple Match Action Box

The Multiple Match Action Box allows you to choose which rows appear in the output table following a successful multi-table match. For each input table you have two options to choose from:

##### Matched Rows Only

Rows will appear in the output table only if they contain an entry from the table in question (but see below).

##### All Rows

Rows will appear in the output table if they contain an entry from the table in question, regardless of constraints on other input tables (overrides the previous option).

At present the multi-table options available within TOPCAT are not so comprehensive as those provided by the corresponding STILTS command `tmatchn`. This may be improved in future.

### A.9 VO Data Access Windows

Several windows in TOPCAT allow access to standard Virtual Observatory (VO) services. These share the idea of querying the *Registry* for suitable data services, selecting one of the services thus returned, and then submitting one or more queries to that data service to generate a tabular result. These ideas are explained in more detail in Section 6, while the current section describes the windows. The next subsection describes the components from which these windows are put together. All the windows contain a registry query panel to select a suitable data access service. The positional query windows also contain either a single or a multiple search panel, and the TAP window contains its own custom panel, to define the actual query or queries to submit to the selected service. The windows themselves are described in the following subsections.

TOPCAT also provides access to two important services provided by the CDS (Centre de Données astronomiques de Strasbourg): the VizieR Catalogue Service Query and CDS X-Match Upload Window. These are not strictly speaking Virtual Observatory services, but since they provide access to the comprehensive VizieR archive of astronomical tables they can provide similar or superior functionality to their VO counterparts, so it's worth being aware of them.

#### A.9.1 Common Features

This section describes the graphical components which make up the VO data access windows.

### A.9.1.1 Registry Query Panel

**Window Columns Registry Interop Help**

**Available Cone Services**

Registry:  ☐ RegTAP ☐

Keywords:

Match Fields: ☒ Short Name ☒ Title ☐ Subjects ☐ ID ☐ Publisher ☐ Description

☒ Accept Resource Lists

Short Name	Title
J/AJ/140/1907	SDSS-GALAXY QSO Catalog (Hatchings+, 2010)
J/ApJ/658/99	Pairs of QSO in SDSS-DR4 (Myers+, 2007)
J/ApJ/194/45	QSO properties from SDSS-DR7 (Shen+, 2011)
J/MNRAS/354/L31	SDSS DR2 QSO and DLA properties (Murphy+, 2004)
J/MNRAS/392/19	The 2dF-SDSS QSO survey (Croom+, 2009)
SDSS(QSO)	Sloan Digital Sky Survey Quasar Catalog (7th Data Release)
SDSSBALQSO	Sloan Digital Sky Survey Broad Absorption Line Quasars Catalog: 3rd Data Release
SDSSCXOQSO	Sloan Digital Sky Survey Quasars Detected by Chandra

AccessURL	Description	Version
http://heasarc.gsfc.nasa.gov/...		

**Resource Count: 14**

### Registry Query Panel

The Registry Query Panel appears in all VO query windows (though by default the TAP window uses an alternative registry search panel) and allows you to search for entries in the Registry representing data services of the type appropriate to the window. Note that if you know the service URL for the service that you want to use, you don't need to use this panel at all; you can simply fill in the URL in the lower part of the window.

The top part of this panel contains the fields you can fill in to query the registry for services matching your constraints. This consists of the following parts:

#### Registry

This determines the registry that will be queried for suitable data services. You can select from the available options, or enter a URL corresponding to some other registry service that you know about. The default is currently the GAVO registry, which at time of writing usually works pretty well, but other registries may have somewhat different data holdings. Some registries don't work very well, so you may have trouble if you use other ones.

To the right of this box is a selector that allows you to choose between two registry protocol options: **RegTAP** and **RI1.0**, corresponding to the IVOA Relational Registry and Registry Interface 1.0 standards respectively. At the time of writing (June 2014), RI1.0 is being phased out, and RegTAP is recommended. It is in any case considerably faster and more reliable than RI1.0. Unless you are a registry enthusiast, sticking to RegTAP is recommended.

The **Update Registry List** item in the **Registry** menu queries the currently selected registry for other registries, and adds more entries to the selector box as appropriate; the little round indicator between the URL and protocol selectors indicates progress of this query.

#### Keywords

Fill this in with keywords appropriate to the services you want to find. The format is space-separated words, and there is no wildcarding. So for instance if you want to find services relating to quasars in Sloan data you could enter the terms "SDSS QSO". These terms are matched against certain fields of each registry entry (as determined by the Match Fields checkboxes below) of each registry entry to see if it matches. Usually one or two short words is the appropriate level of detail. If no keywords are filled in, all the appropriate services are retrieved. Some experimentation may be required to specify a query with a reasonable number of results; if your query returns no results, try fewer or more general terms, and if it returns too many, try more or more specific ones.

### And/Or button

This determines whether a registry entry is considered to match the chosen keywords if it matches all of them, or at least one of them. Clicking it toggles between "And" and "Or". The default is "And", meaning that all keywords must be matched (against any of the selected match fields) for a selection.

### Match Fields

This line contains a number of checkboxes which allow you to define which fields in each registry record will be compared against the keywords you have entered when looking for matches. The contents of the **Short Name**, **Title**, **Subjects**, **ID** and **Publisher** fields are visible in the table of results. **Description** is a free-form, often long, description of the resource; it can be useful, but can often accidentally contain text which causes unwanted matches. By default the Short Name, Title, Subjects and ID fields are matched - click the checkboxes to change the selection

### Find Services button

When the other fields have been filled in, hit the Find button and the registry will be searched to find data services that match the constraints you have given.



In some cases, when the window first appears, it will automatically query the registry for all known services of the appropriate type. This is currently the case for SIA and SSA services, since there is a manageable number of these. There are thousands of cone searches however, so for the cone search windows you need to fill in a query and submit it yourself. In either case, you can resubmit a query (hit the **Cancel** button first if necessary) to make another selection of registry entries.

When the query completes, a table of matching entries is displayed below the Submit button. The number of resources found is displayed at the bottom, labelled **Resource Count**. The table contains one row for each matching entry, i.e. each service of the appropriate type that you can submit your data query to. The columns in the table give a short name, title, description and some other information about each service. You can sort the table rows according to the entries in different columns by clicking on the column header; click again on the same header to reverse the sort order. A little up/down arrow in a column header indicates that that column is currently determining the sort order. You can adjust the columns which are visible using the **Columns** menu, or drag the columns sideways to change their order in the usual way.

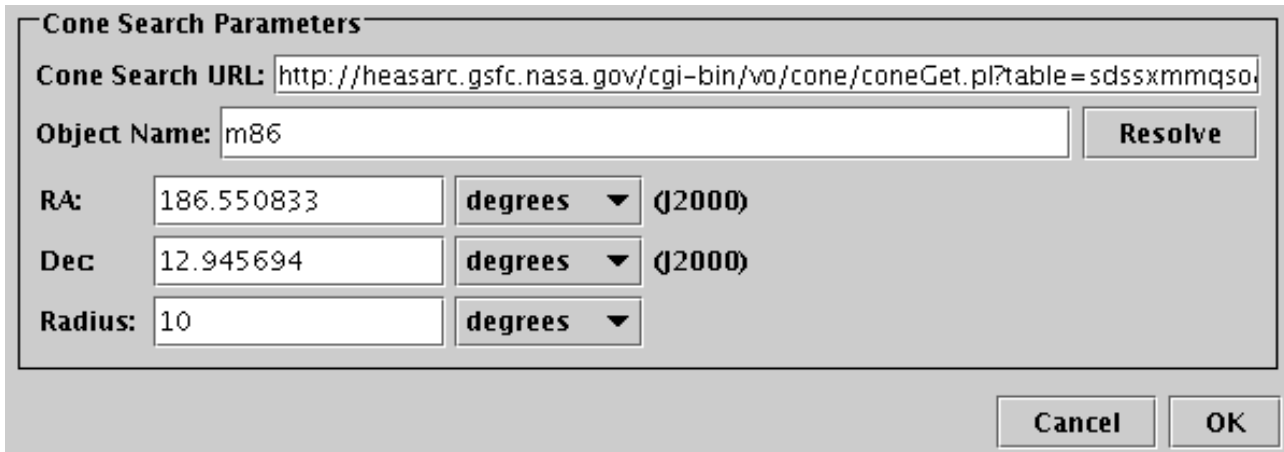
You should select the service that you want to query for data by clicking on it. When this is done, you will see one or more rows appear in a second table underneath the first one. In most cases, this contains a single row which will be automatically selected. However, some registry entries have multiple data services associated with a single record - in this case you should select the appropriate entry from the lower table as well.

Once you have selected a data service in this way, its service URL will be entered in the lower part of the window (not shown in the figure above) and you are ready to perform the actual single or multiple data query.

If a SAMP hub is running (see Section 9), it is possible to exchange lists of resources between this

window and other applications. If another application sends a suitable message to TOPCAT, and the resources are of an appropriate type, and the window is visible at the time, the received list can be loaded and displayed here, replacing any which are currently displayed. You can control whether incoming lists are accepted and displayed in this way using the **Accept Resource Lists** toggle, either using the checkbox just above the resource display panel, or using the corresponding item in the **Interop** menu. If you wish to send your selection to one or more other applications which can make use of them, use the Broadcast Resource List () or Send Resource List () options in the **Interop** menu.

### A.9.1.2 Single Positional Search Panel



#### Single Positional Search Panel

The Single Positional Search Panel appears in VO-based table load dialogues and is used to specify data queries on a single region of the sky. The purpose is to load a new table (containing entries representing catalogue entries, images or spectra). To use it you must fill in the **URL** field and the position definition, and then hit the **OK** button.

#### URL

This must contain the service URL for the data service that you are querying. Usually, this will be filled in by selecting one of the services obtained by the registry query. However, you can fill it in manually with the URL of a service you know about if you prefer. If you know what you're doing, it's also possible to doctor a service URL filled in by the registry selection, for instance by adding `&name=value` parameters to it.

#### Object Name

You can fill this in with the name of an object to be resolved by SIMBAD, and hit the **Resolve** button, which will fill in the coordinates in the RA and Dec fields below. If you enter the RA and Dec manually, you don't need to use this field.

#### RA and Dec

Fill in the central position in terms of J2000 Right Ascension and Declination that you are interested in here. You can either do it manually, or by using the **Object Name** field as above. Select the units as appropriate.

#### Radius

Indicates how large is the region you are interested in. This field has slightly different appearance and meaning for the different data service types; hover over it with the mouse to see a tooltip with the details. For SIA and SSA, but not for Cone Search, it is permissible to leave it blank (though certain services don't seem to like that). Select the units as appropriate.

When the fields are filled in, hit the **OK** button and wait for the new table defined by your query parameters to load. If you get bored waiting (the service may be down, or just very slow, or you may have submitted a query with a very large return set), you can hit the **Cancel** button, and perhaps adjust the parameters and try again.

If a SAMP hub is running (see Section 9), then other running clients may be able to fill in the RA and Dec coordinates. For instance clicking on a sky position in an image viewing tool might fill in the sky position you clicked on. Usually this will happen automatically. This is often convenient, but may not be what you want. You can control whether incoming coordinate settings are accepted in this way using the **Accept Sky Positions** toggle below the Resolve button, or using the corresponding item in the **Interop** menu.

### A.9.1.3 Multiple Positional Search Panel

**Multiple Cone Search Parameters**

Cone Search URL:

Input Table:

RA column:   ☐

Dec column:   ☐

Search Radius column:

Output Mode:

Parallelism:  Error Handling:

### Multiple Positional Search Panel

The Multiple Positional Search Panel appears in VO-based multiple query windows which specify a procedure in which one query is made for each row of an input table. Each of the input rows defines a (probably small) region on the sky. The purpose is to find data records (about catalogue entries, images or spectra) in a remote catalogue corresponding to each row of a local catalogue. As such, it is used to define a kind of join, of a local Apparent Table with a remote one, where the remote one is exposed via a VO data service.

To use it you must fill in the **URL** field to select the remote service, and fields defining how each row of the input table can be turned into a positional query. You also need to define what form the result will be returned in. These parts are described below.

The URL field specifies the data service which is to be queried:

#### URL

This must contain the service URL for the data service that you are querying. Usually, this will be filled in by selecting one of the services obtained by the registry query. However, you can fill it in manually with the URL of a service you know about if you prefer. If you know what you're doing, it's also possible to doctor a service URL filled in by the registry selection, for instance by adding `&name=value` parameters to it.

The following fields define what queries are to be sent to the service:

**Input Table**

Select one of the tables loaded into TOPCAT whose row positions you want to search around.

**RA and Dec columns**

Fill these in with the columns in the input table which contain J2000 Right Ascension and Declination (you can perform coordinate conversions using the Sky Coordinates Window if required). Select the units as appropriate. If the input table has appropriate metadata, the correct values may be filled in by default, but it may occasionally guess wrong, so it's wise to check that they are filled in correctly.

**Search Radius column**

Fill this in with a value that represents the size of search that you want to make around each row's position. You can either use a constant value, or a column name, or any numeric algebraic expression based on the table (see Section 7). Select units as appropriate. This field has slightly different appearance and meaning for the different data service types; hover over it with the mouse to see a tooltip with the details. For SIA and SSA, but not for Cone Search, it is permissible to leave it blank (though certain services don't seem to like that). Select the units as appropriate.

The Output Mode selector indicates what will be done with the result.

**Output Mode**

When the queries have been performed, there are different ways that the results can be returned. Since the operation here is basically a crossmatch between a local table and a remote one, this effectively describes what sort of join is to be done. The options are as follows:

**New joined table with best matches**

A new table will be generated, containing one row for each row in the input table which returned at least one search result from the remote table. The best (closest) match will be included; any others will be discarded. The new table will have all the columns of the input table and all the columns returned by the data service.

**New joined table with all matches**

A new table will be generated, containing one row for each search result returned. Rows from the input table will be duplicated where more than one search result was returned corresponding to that row. The new table will have all the columns of the input table and all the columns returned by the data service.

**New joined table, one row per input row**

A new table will be generated containing the same number of rows as the input table. If any search results were returned for each row, the best (closest) one is included. Any others will be discarded, and cells will be empty for rows with no search results. The new table will have all the columns of the input table and all the columns returned by the data service.

**Add subset for matched rows**

No new table will be generated, but a new subset will be added to the table which includes only those rows which returned at least one search result.

The final controls adjust the details of how the individual queries are submitted to the remote service.

**Parallelism**

Controls how many cone search queries are running at once. By making several queries to the service concurrently, the time it takes to fill in the whole table can be much quicker than making the query for the first row, waiting for the result, making the query for the second row, etc. The value here is approximately how many queries will be run at the same time.

Increasing the value might make your multiple query run faster; or it might overload the server and make you unpopular with the service administrator, or result in your query taking longer or failing altogether, or both. The default value of 5 is probably reasonable, but experiment with adjusting it if you want.

### **Error Handling**

Determines what happens if one of the queries fails with an error. The options are:

**abort**

terminate the operation; the whole multiple search fails

**ignore**

treat a failed query the same as one which returns zero rows

**retry N times**

try N times to get a non-fail result

**retry indefinitely**

keep trying for a non-fail result

The best setting for this depends on the way the service is set up; the default is **abort**, but for unreliable or poorly implemented services it may be better to continue operation in the face of a few failures.

When all of the fields are filled in (defaults can be used for many of them), hit the **Go** button, and the search will commence. It may be quite slow; a progress bar is shown at the bottom of the window. When it completes, a popup window summarising the result will be shown. If you get bored waiting, you can hit the **Cancel** button, and perhaps adjust the parameters and try again.

### **A.9.2 Cone Search**



**Cone Search**

Window Columns Registry Interop Help

Registry:  RegTAP

Keywords:  And

Match Fields: ☒ Short Name ☒ Title ☐ Subjects ☐ ID ☐ Publisher ☐ Description

☒ Accept Resource Lists Cancel Find Services

Port Name	Title
94/45	QSO properties from SDSS-DR7 (Stern+, 2011)
/354/L31	SDSS DR2 QSO and DLA properties (Murphy+, 2004)
/392/19	The 2dF-SDSS QSO survey (Croom+, 2009)
0)	Sloan Digital Sky Survey Quasar Catalog (7th Data Release)
QSO	Sloan Digital Sky Survey Broad Absorption Line Quasars Catalog: 3rd Data Release
QSO	Sloan Digital Sky Survey Quasars Detected by Chandra
QSO	Sloan Digital Sky Survey/UKIRT DSS Large Area Survey Matched Quasars Catalog
Cand.	Sloan Digital Sky Survey NBC Quasar Candidate Catalog
QSO	Sloan Digital Sky Survey (DR5)/XMM-Newton Quasar Survey Catalog

AccessURL	Description	Version
<a href="http://heasarc.gsfc.nasa.gov/...">http://heasarc.gsfc.nasa.gov/...</a>		

Resource Count: 15

**Cone Parameters**

Cone URL:

Object Name:  Resolve

RA:  degrees (J2000) ☒ Accept Sky Positions


Dec:  degrees (J2000)

Radius:  degrees

Verbosity: 2 (normal)

OK

### Cone table load dialogue

The cone search load dialogue can be opened using the **Cone Search** button (  ) from the Load

Window's toolbar or the Control Window's **VO** menu. It allows you to query one of a number of external web services for objects from a given catalogue in a given region of the sky. Data held by cone search services are effectively source catalogues with sky position columns.

The window consists of a Registry Query Panel at the top, and a Single Positional Search Panel below.

The search panel includes a **Verbosity** selector as well as the normal features: this allows you to choose options for how many columns will appear in the output table, from 1 (minimum) to 3 (maximum). Some services will take notice of this parameter and return narrower or wider tables respectively, and others will ignore it.

### A.9.3 Simple Image Access (SIA) Query

**SIA Query**

Window Columns Registry Interop Help

Available SIA Query Services

Registry:  RegTAP

Keywords:  And

Match Fields: ☒ Short Name ☒ Title ☒ Subjects ☐ ID ☐ Publisher ☐ Description

☒ Accept Resource Lists Cancel Find Services

Short Name	Title
LH	2MASS 6X Lockman Hole Ancillary Data Atlas
2MASS QL	2MASS All-Sky Quicklook Image Service
LGA	2MASS Large Galaxy Atlas
DSS2	2nd Digitized Sky Survey (Blue)
GB6	4850 MHz Survey/GB6
1420MHz	Bonn 1420 MHz Survey
CADC	CADC Image Search

AccessURL	Description	Version
<a href="http://irsa.ipac.caltech.edu/c...">http://irsa.ipac.caltech.edu/c...</a>		

SIA Query Parameters

SIA Query URL:

Object Name:  Resolve

RA:  degrees (J2000) ☒ Accept Sky Positions


Dec:  degrees (J2000)

Angular Size:  degrees

Image Format: image/fits

OK

#### SIA query load dialog

The SIA query load dialogue can be opened using the **SIA Query** button () from the Load Window's toolbar or the Control Window's **VO** menu. It allows you to query one of a number of external web services for images covering a given region of the sky.

The window consists of a Registry Query Panel at the top, and a Single Positional Search Panel below. The **Angular Size** may be set to zero; this means that any image covering the chosen position will be selected. There is additionally an **Image Format** selector which allows you to restrict the result to contain only images in certain formats (the default is set to only FITS).

The result of a successful query is a table in which each row represents an image that can be downloaded from a remote site; one of the columns contains the image URL, and the other columns contain associated metadata such as image format, size, shape and so on. Different services provide different kinds of answer to these queries; some may have a static list of observed or reduced FITS files, and others may generate images on the fly. See the SIA standard, and individual registry records, for details. A useful thing to do with the resulting table might be to set its Activation Action to the **View URL as Image** option using the Activation Window.


#### A.9.4 Simple Spectral Access (SSA) Query

The screenshot shows the 'SSA Query' window with the following components:

- Menu Bar:** Window, Columns, Registry, Interop, Help
- Available SSA Query Services:**
  - Registry:  RegTAP
  - Keywords:  And
  - Match Fields: ☒ Short Name ☒ Title ☐ Subjects ☐ ID ☐ Publisher ☐ Description
  - ☒ Accept Resource Lists Cancel Find Services
- Results Table:**

Short Name	Title	
HST Spectra	Hubble Space Telescope Spectra	UV, Optical, and Infrared A
HST/FOS/SSAP	Hubble Space Telescope Faint Object Spectrograph	Calibrated spectra of the H
- SSA Query Parameters:**
  - SSA Query URL:
  - Object Name:  Resolve
  - RA:  degrees (J2000) ☒ Accept Sky Positions
  - Dec:  degrees (J2000)
  - Diameter:  arcmin
  - Spectrum Format: votable
- Buttons:** OK

## SSA query load dialog

The SSA query load dialogue can be opened using the **SSA Query** button () from the Load Window's toolbar or the Control Window's **VO** menu. It allows you to query one of a number of external web services for spectra in a given region of the sky.

The window consists of a Registry Query Panel at the top, and a Single Positional Search Panel below. The SSA specification says that the **Diameter** field may be left blank, in which case the service "should supply a default value intended to find nearby objects". However, experience shows that this doesn't always work, so it may be best to fill this in. There is additionally a **Spectrum Format** selector which allows you to restrict the result to contain only spectra in certain formats. By default, the service chooses which formats to return.

The result of a successful query is a table in which each row represents a spectrum that can be downloaded from a remote site; one of the columns contains the spectrum download URL, and the other columns contain associated metadata such as spectrum format, WCS, provenance information and so on. See the SSA standard for details. A useful thing to do with the resulting table might be to set its Activation Action to the **View URL as Spectrum** option using the Activation Window.

### A.9.5 Multiple Cone Search

**Multiple Cone Search**

Window Columns Search Registry Interop Help

Available Cone Search Services

Registry:  RegTAP

Keywords:  And

Match Fields: ☒ Short Name ☒ Title ☐ Subjects ☐ ID ☐ Publisher ☐ Description

☒ Accept Resource Lists Cancel Find Services

Short Name	Publisher	Title
FIRST	NASA/GSFC HEASARC	Faint Images of the Radio Sky at Twenty cm (FIRST)
J/A+A/416/35	CDS	Radio Emission from VLA FIRST Survey AGN (Wadad
J/A+A/452/739	CDS	ULX sources and FIRST radio sources (Sanchez-Sutil
J/AJ/121/1872	CDS	FIRST radio-selected QSOs (Helfand+, 2001)
J/AJ/127/2565	CDS	Faint FIRST variable radio sources (De Vries+, 2004

AccessURL	Description	Version
http://vizier.u-strasbg.fr/viz-...		

Multiple Cone Search Parameters

Cone Search URL:  🔍

Input Table: 1: 6dfgs\_mini.xml.bz2 🔍

RA column:  degrees (J2000)

Dec column:  degrees (J2000)

Search Radius column:  arcsec

Verbosity: 2 (normal)

Output Mode: New joined table with best matches 🔍

Parallelism:  Error Handling: ignore

Go Stop

### Multiple cone search window

The multiple cone search window can be opened using the **Multicone** button (🔍) in the Control

Window's toolbar, or its **VO** or **Joins** menus. It allows you to select one of a number of external web services which provide remote catalogue queries, and execute a query for each row of an input table. This effectively allows you to perform a positional crossmatch between a local table and a remote one.

The window consists of a Registry Query Panel at the top, and a Multiple Positional Search Panel below.

The search panel includes a **Verbosity** selector as well as the normal features: this allows you to choose options for how many columns will appear in the output table, from 1 (minimum) to 3 (maximum). Some services will take notice of this parameter and return narrower or wider tables respectively, and others will ignore it.

The following item is available in the toolbar and **Search** menu:



#### **Use Service Coverage**

If this option is selected, then the search attempts to query the service for the sky coverage it represents, if that information is available. Then it only dispatches queries for positions in the input table which fall within that coverage region, since any queries outside that region are bound to return an empty result. Depending on the coverage, this can substantially reduce the number of queries made, and hence the time taken. If there is no overlap between the input table and service coverage, the query is bound to return no results at all, and the time will be reduced effectively to zero.

When in operation (i.e. when this option is selected and the service provides coverage information) the coverage of the service, input table and their overlap is summarised with small (Mollweide equatorial) all-sky icons in blue, red and magenta respectively, as in the screenshot above. This makes it easy to see the sky regions in which results may be obtained, and also the amount by which this option reduces useless queries.

Currently, this option effectively works only for cone searches provided by CDS's VizieR service (publisher "CDS") which provide coverage information using Multi-Order Coverage maps (MOCs). For non-VizieR services this option has no effect.

### **A.9.6 Multiple SIA Query**

**Multiple SIA**

Window Columns Registry Interop Help

Available SIA Services

Registry:  ☐ RegTAP ☐

Keywords:

Match Fields: ☒ Short Name ☒ Title ☐ Subjects ☐ ID ☒ Publisher ☐ Description

☒ Accept Resource Lists

Short Name	Title	Subjects	Identifier	
DSS ESO	Digitized Sky Survey	Surveys, Stars, Galaxies	ivo://eso.org/dss	ESO
DSS ESO	Digitized Sky Survey	Surveys, Stars, Galaxies	ivo://archive.eso.org/dss	European

AccessURL	Description	Version
http://archive.eso.org/bin/dss...		

Multiple SIA Parameters

SIA URL:

Input Table:

RA column:   (J2000)

Dec column:   (J2000)


Search Radius column:

Image Format:

Output Mode:

Parallelism:

### Multiple SIA window

The multiple Simple Image Access window can be opened using the **Multiple SIA** button () from the Control Window's **VO** or **Joins** menus. It allows you to select one of a number of external web services which provide queries on remotely accessible image holdings, and execute a query for each row of an input table.

The window consists of a Registry Query Panel at the top, and a Multiple Positional Search Panel

below. The **Search Radius column** may be set to zero; this means that any image covering the chosen position will be selected. There is additionally an **Image Format** selector which allows you to restrict the result to contain only images in certain formats (the default is set to only FITS).

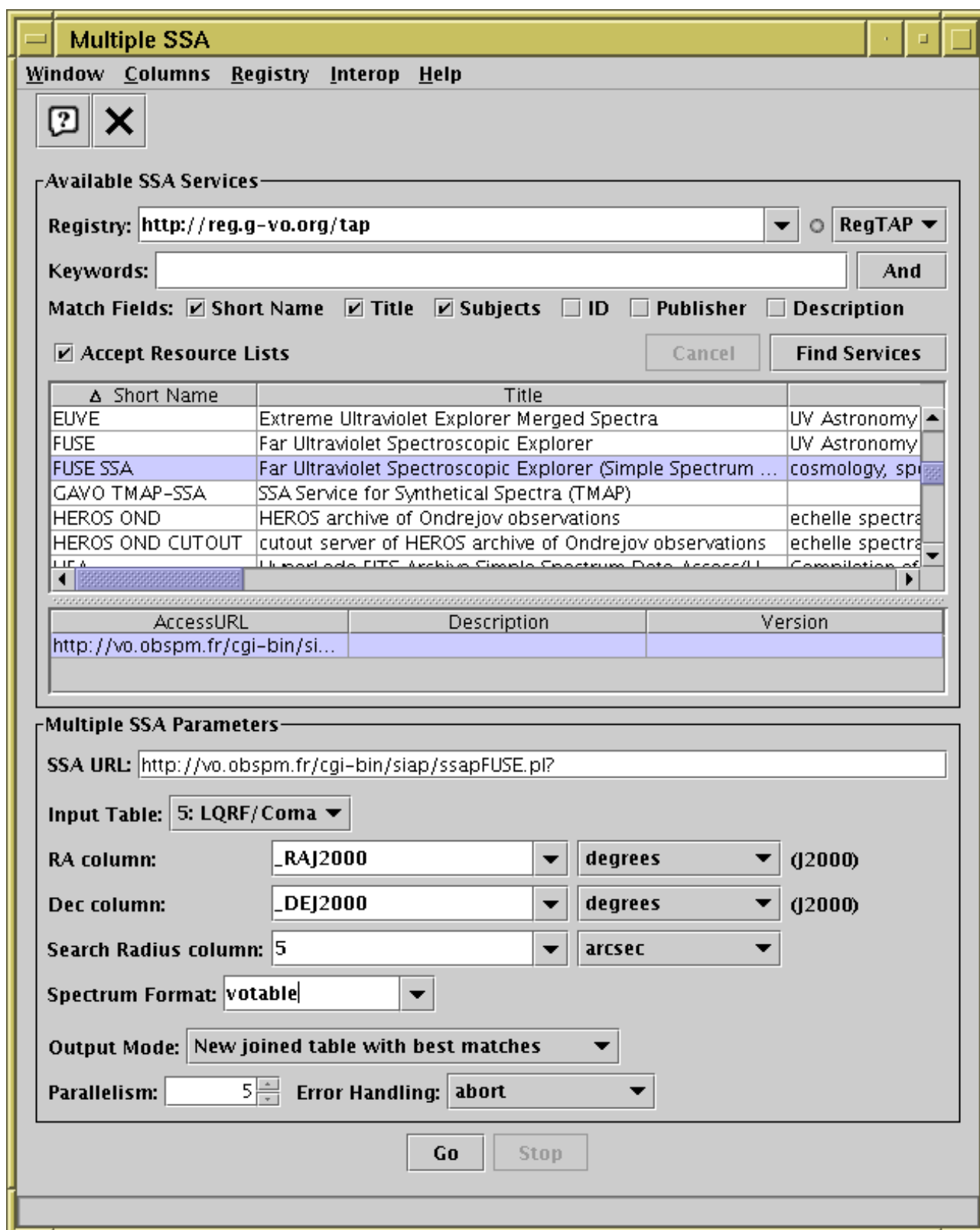
The result of each single successful query is a table in which each row represents an image that can be downloaded from a remote site; one of the columns contains the image URL, and the other columns contain associated metadata such as image format, size, shape and so on. Different services provide different kinds of answer to these queries; some may have a static list of observed or reduced FITS files, and others may generate images on the fly. See the SIA standard, and individual registry records, for details.

Note that some services return multiple images at the same positions but at different wavelengths. In this case TOPCAT's criterion for the "best" match (the one closest to the central query position) may not make much sense. For this reason, take care if using the "New joined table with best matches" or "New joined table, one row per input row" output modes.


One use of this function is to add some columns to an existing table which contain URLs of images from a given server corresponding to the sky positions of those rows. Having done this you might want to set the resulting table's Activation Action to the **View URL as Image** option using the Activation Window.

### A.9.7 Multiple SSA Query





## Multiple SSA window

The multiple Simple Spectral Access window can be opened using the **Multiple SSA** button () from the Control Window's **VO** or **Joins** menus. It allows you to select one of a number of external web services which provide queries on remotely accessible spectral data holdings, and execute a query for each row of an input table.

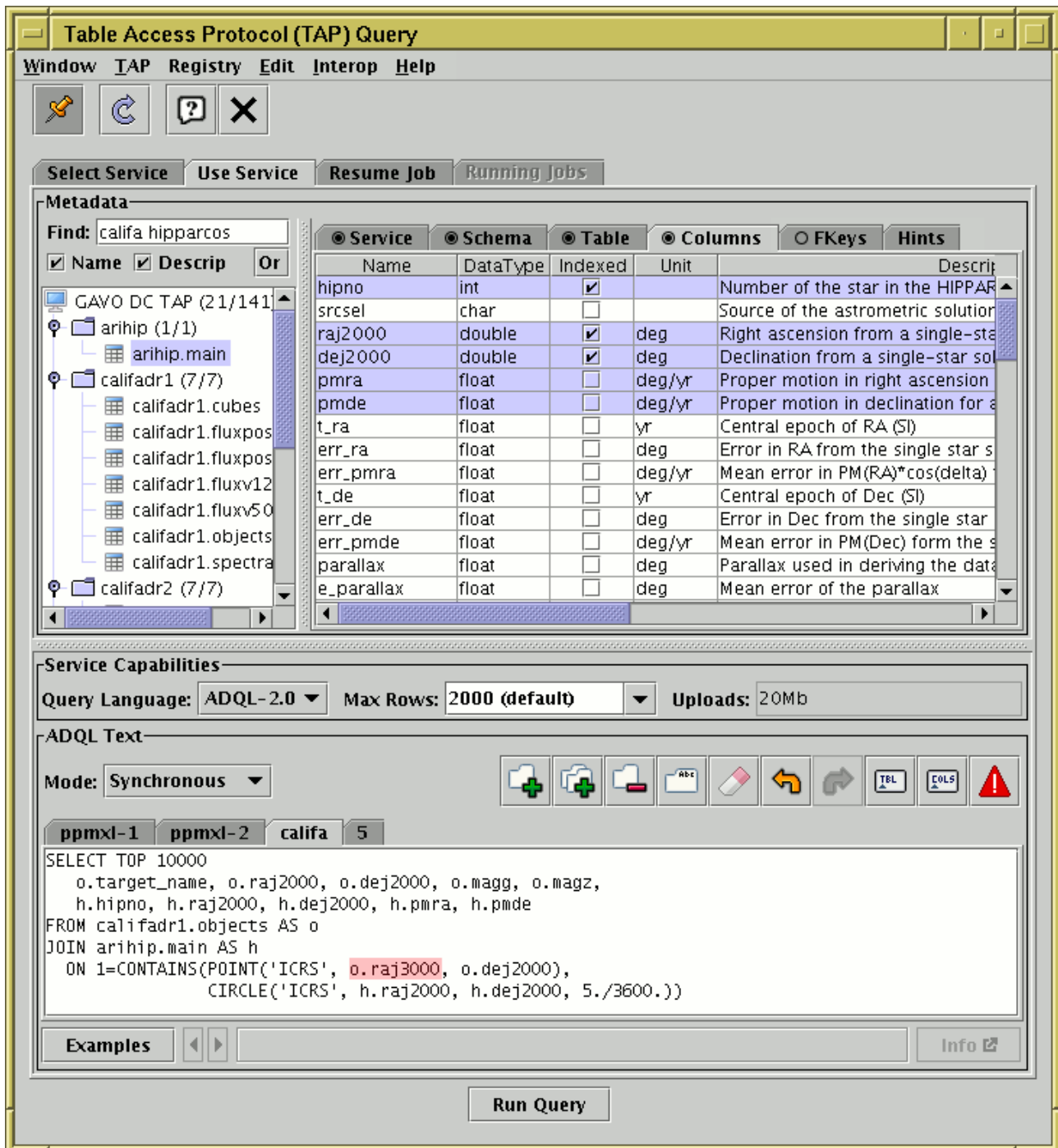
The window consists of a Registry Query Panel at the top, and a Multiple Positional Search Panel

below. The SSA specification says that the **Search Radius** column may be left blank, in which case the service "should supply a default value intended to find nearby objects". However, experience shows that this doesn't always work so it may be best to fill this in. There is additionally a **Spectrum Format** selector which allows you to restrict the result to contain only spectra in certain formats. By default, the service chooses which formats to return.


The result of each single successful query is a table in which each row represents a spectrum that can be downloaded from a remote site; one of the columns contains the spectrum URL, and the other columns contain associated metadata such as spectrum format, WCS, provenance information and so on. See the SSA standard for details.

One use of this function is to add some columns to an existing table which contain URLs of spectra from a given server corresponding to the sky positions of those rows. Having done this you might want to set the resulting table's Activation Action to the **View URL as Spectrum** option using the Activation Window.

#### A.9.8 Table Access Protocol (TAP) Query



## Table Access Protocol Window

The Table Access Protocol (TAP) load window can be opened using the **TAP Query** button ()

in the Control Window's toolbar or **VO** menu, or the Load Window's toolbar. It allows you to use the TAP protocol to make freeform queries of remote database services using an SQL-like language. This is a powerful capability, giving you full access to remote data holdings; it's similar to systems like SDSS's CasJobs, but using the IVOA-endorsed TAP standard means that you can use the same interface for many different remote datasets.

Using TAP is one of the more complicated things you can do in TOPCAT, but it's well worth getting to grips with because of the capabilities it gives you for customised access to all kinds of remote data.

In order to interrogate the remote database you will have to write queries in Astronomical Data

Query Language (ADQL). ADQL is essentially a particular dialect of SQL, so if you have used SQL before, you should be able to write ADQL queries. A tutorial on writing ADQL/SQL is beyond the scope of this manual, but the Use Service tab provides an **Examples** menu which will give you a good start in using it. You can find some introductions and tutorials elsewhere, for instance at <http://docs.g-vo.org/adql> (Markus Demleitner) or <http://is.gd/ADQLTutorial> (Simon Murphy). Some TAP services also provide tutorials or examples on their web pages.

Once your query is written, you can submit it in either *synchronous* or *asynchronous* mode. Synchronous queries begin to execute right away and should have a relatively short runtime; asynchronous queries are submitted to a server, run for a potentially long time, and retrieve the result later.

The window is composed of four tabs, of which you can see one at a time (some of these tabs themselves contain tabbed parts). You will use two or three of them when submitting a TAP query. Each tab is described in more detail in the following subsections, but an overview of their operation is like this:

#### Select Service tab

Chooses the TAP service you wish to query, either using a registry search or by entering the service URL directly. Once you have chosen a service you can move to the **Use Service** tab.

#### Use Service tab

Displays information about the service and the tables it holds, and allows you to enter the text of the query and submit the job. If the job is submitted synchronously, loading will begin directly, but in the asynchronous case, you will move to the **Running Jobs** tab.

#### Running Jobs tab

Shows a list of the asynchronous query jobs you have submitted, with details on their progress.

#### Resume Job tab

Optionally allows you to resume an asynchronous query started in some earlier session.

When you first visit this window, the **Select Service** tab will be visible, and when an asynchronous query has been submitted the **Running Jobs** tab will be visible. If you want to submit another query to the same service, or use a different service, just select **Use Service** or **Select Service** respectively, by clicking on the appropriate tab at the top.

This window offers some menus additional to the standard VO window menus:

#### TAP menu

This contains some configuration options specific to Table Access Protocol service interaction. It has the following sub-items:

##### Reload ( )

Reloads information displayed in the window from the server. This may be updating job status information or reloading table or service metadata; the exact behaviour depends on which tab is currently visible.

##### Job Deletion submenu

Allows you to choose when asynchronously submitted jobs will by default be deleted from the server. Once they are deleted they will no longer be visible in the **Running Jobs** tab. If they are not deleted in this way, they will stay on the server until they are deleted according to the server's deletion policy. Choose one of the available options:

- **On Completion:** deleted when it has completed with either success or error status
- **On Exit:** deleted when TOPCAT shuts down
- **Never:** not deleted (except according to server policy)

You can override these settings once a job has started by using the controls at the bottom of the Running Jobs tab. The settings in this menu affect the default value of the **Delete On Exit** checkbox. If you want to resume the job in a later session, you'll have to make sure that the job is not deleted on completion or exit.

### Metadata Acquisition submenu

Controls how the application loads service metadata - information about the tables and columns provided by the service. The following options are available:

- **Auto:** make a sensible decision; this generally means **TAP\_SCHEMA\_C** for services with very many tables, and **TableSet** otherwise.
- **TableSet:** reads the document at the `/tables` endpoint of the TAP service; all table metadata is loaded up front.
- **TAP\_SCHEMA\_C:** reads metadata from TAP\_SCHEMA tables; all schemas, tables and foreign keys are read at once, columns are read only when required.
- **TAP\_SCHEMA\_CF:** reads metadata from TAP\_SCHEMA tables; all schemas and tables are read at once, columns and foreign keys are read only when required.
- **TAP\_SCHEMA:** reads all metadata at once from the TAP\_SCHEMA tables.
- **VizieR:** uses the 2-stage TableSet from `/tables` as implemented (at time of writing) by the TAPVizieR service; columns are read only when required (experimental, may be withdrawn).
- **CADC:** uses the 2-stage TableSet from `/tables` as implemented (at time of writing) by CADK services; columns are read only when required (experimental, may be withdrawn).
- **VOSI11-1step** Uses VOSI 1.1-style detail-sensitive requests to the `/tables` endpoint. `detail=max` is specified initially to indicate that full metadata up-front metadata load is preferred.
- **VOSI11-2step** Uses VOSI 1.1-style detail-sensitive requests to the `/tables` endpoint. `detail=min` is specified initially to indicate that on-demand column load is preferred.
- **VOSI11** Uses VOSI 1.1-style detail-sensitive requests to the `/tables` endpoint. No `detail` specification is made initially, so the service decides whether up-front loading is done.

You will usually want to stick with the default value of **Auto**. However, if you're on a slow link, and it's taking a long time to display any metadata for medium-sized services, you might want to switch to **TAP\_SCHEMA\_C/CF**, or **VizieR** for talking to TAPVizier.

### Response Format submenu

Configures how the application requests that VOTable results are to be returned. The options (**TABLEDATA**, **BINARY**, **BINARY2**) correspond to the different defined VOTable serialization formats. The selected format is only actually requested if the service has indicated support for it, otherwise the service default is used.

### Upload Format submenu

Configures how tables are transmitted to the service for upload if an upload query (one involving a `TAP_UPLOAD.* table`) is performed. The options (**TABLEDATA**, **BINARY**, **BINARY2**) correspond to the different defined VOTable serialization formats. The choice shouldn't affect any results, but it may affect required bandwidth and some services may (though shouldn't) have non-standard requirements for serialization format.

### Service Discovery

Configures how TOPCAT searches for available TAP services. There are currently two options, **GloTS**, which uses the Global TAP Schema service list maintained by GAVO, and **Reg Prototype** which uses a proposed convention based on the IVOA Registry. You are advised to use the default value (currently **GloTS**) unless you understand what you're doing here.

### HTTP gzip

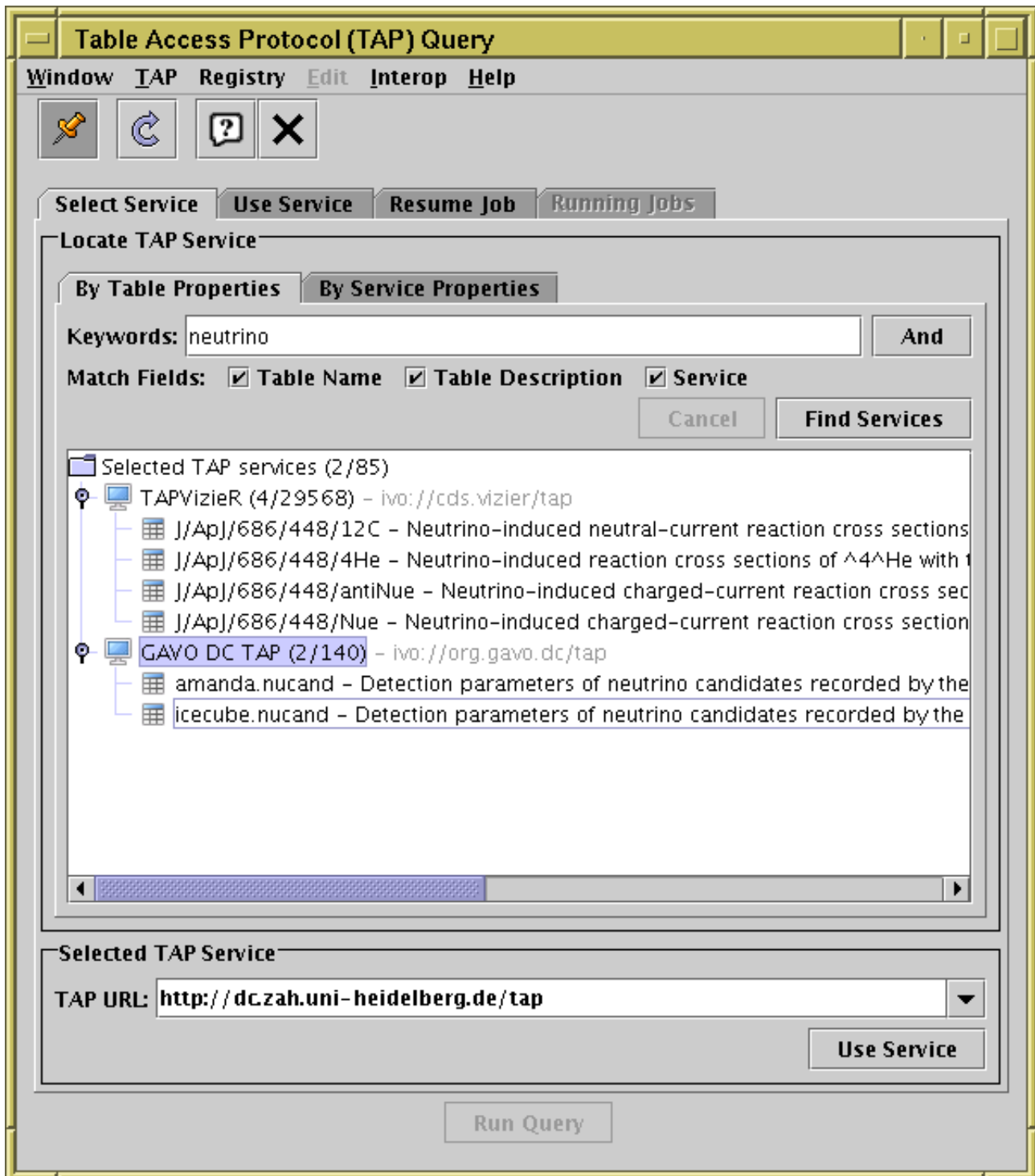
If checked, this option sends `Accept-Encoding: gzip` headers with most of the HTTP requests when communicating with the TAP service. That ought to reduce required bandwidth considerably, but only if the service chooses to honour this request. The setting should have no effect on the actual results, but may affect performance.

**Edit menu**

Provides options concerned with editing ADQL text. This is only active when the **Use Service** tab is visible, and it is described in that section.

**Note:** TAP is a complex protocol, and at time of writing, many TAP services still do not provide full and error-free implementations. Unexpected behaviour may be a result of service implementation deficiencies. Hopefully that will improve over time.

**A.9.8.1 Select Service tab**



**TAP window with Select Service tab visible**

The **Select Service** tab of the TAP load dialogue allows you to choose which TAP service you wish to query. You need to do this before submitting a new TAP query.

There are two separate tabs within the **Locate TAP Service** sub-panel, corresponding to two different ways to select the service you want to use.

#### **By Table Properties**

Use this tab to find TAP services that contain particular tables, or tables relating to particular topics. When first opened (or if no search terms are entered), it shows a list by name of all the known TAP services, ordered by the number of tables they contain. If you enter one or more search terms in the **Keywords** field and hit **Find Services** (or return), then the list will be redisplayed as a tree in which the children of each service are the tables whose name and/or

description matches the search terms you entered (click on the node handles to display/hide the children). You can choose whether to match the **Table Name** or **Table Description** using the **Match Fields** checkboxes; basic **Service** metadata can also be matched (just name, title, ID; for more complex searches by service metadata use the By Service Properties tab, described below). Clicking the **And/Or** button to toggle its value determines whether the text from the displayed tables/service text are required to match all of your search terms, or just one of them. When you've decided which service you want to use, click it, and its access URL will be filled in the **TAP URL** selector at the bottom. Clicking on a table will automatically select its parent service. If you double-click on a service it will take you directly to the Use Service tab.

### By Service Properties

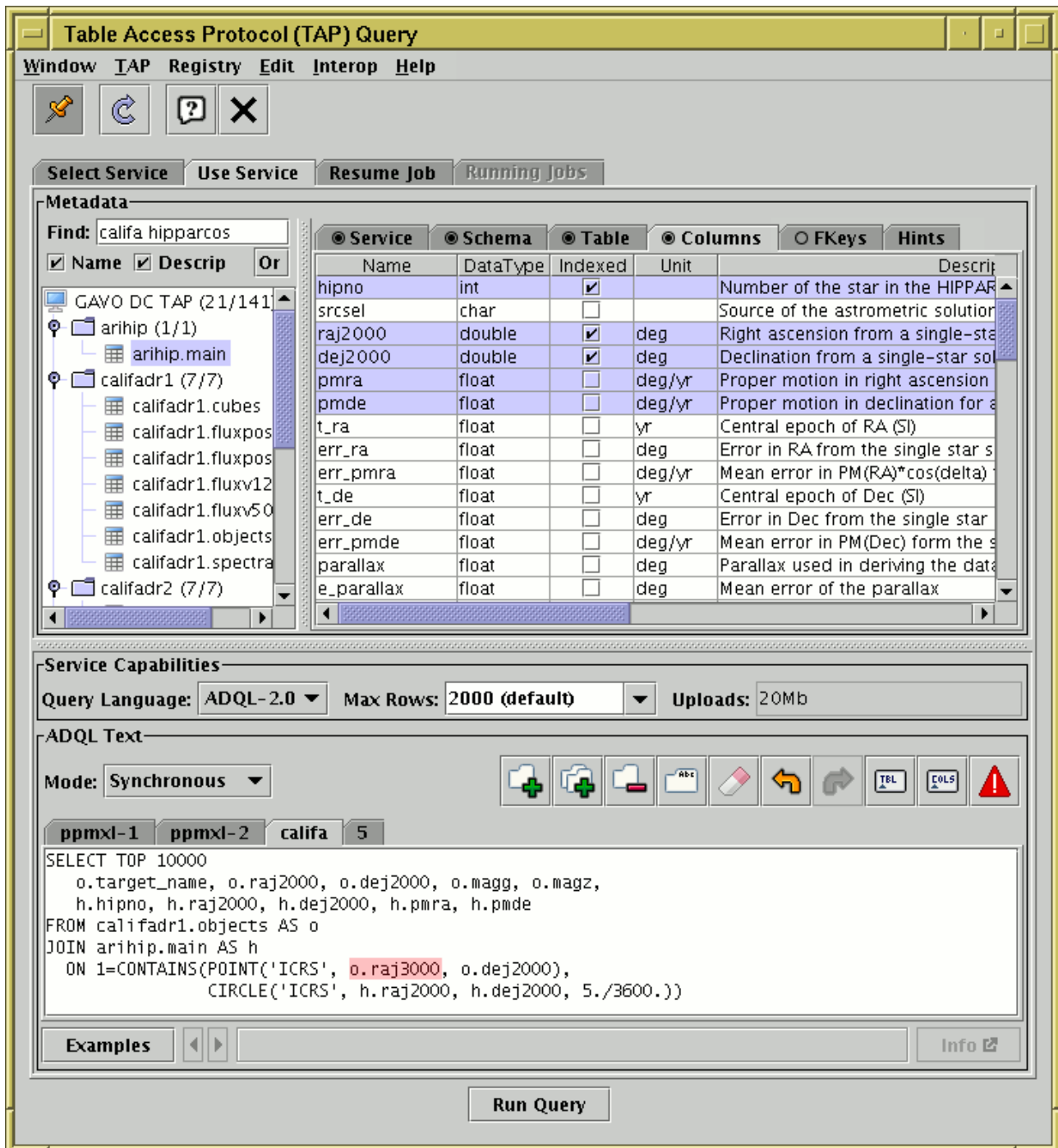
Use this tab to find TAP services based on properties of the service itself, such as the service name or publisher. It is the same as the Registry Query Panel of the other VO query windows. Having queried the registry, click on one of the rows to enter its service URL in the **TAP URL** field at the bottom. At time of writing there are only about a hundred TAP services registered, so it's feasible to query the registry for them all (hit **Find Services** without entering any keywords), and choose the one you want by eye, perhaps sorting the displayed services by one of the metadata items (click on the column header to sort).

If you know the URL of the TAP service you wish to query, you can enter it directly into the **TAP URL** field at the bottom of the window without a registry search. This field remembers URLs previously entered into it, so if you click the little down-arrow on the right, you can easily return to services you visited earlier in the same session. The state (entered ADQL etc) of those sessions is also retained.

Once a service URL has been chosen in one of these ways, you can click the **Use Service** button at the bottom (or equivalently the tab header at the top), and you will be moved to the Use Service tab.

### A.9.8.2 Use Service tab





### TAP window with Use Service tab visible

The **Use Service** tab of the TAP load dialogue displays information about the service you have selected to query, including what tables are available and what columns they contain, and allows you to enter the query text and some additional options about how the query is to be executed. The panel has three parts: **Metadata**, **Service Capabilities**, and **ADQL Text**.


The **Metadata** panel displays information about the tables held by the selected service. On the left of the panel is a searchable tree summarising the schemas and tables in the service, and on the right are some tabs containing more detail. Note all this information has to be loaded from the server, so in some cases you may have to wait before it is visible.

The tree on the left contains a listing of the service's *Schemas*, and under them the *Tables* they contain. A Schema in this context is just a subject grouping that contains one or more tables. Each schema notes in brackets the number of tables it contains; you can reveal or hide the tables by

clicking on the schema node's handle. If the service contains only a small number of tables it may be convenient just to scroll up and down to see them all. But for services with hundreds or thousands of tables, you may need to restrict the displayed tables to those you're interested in. To do this, type one or more search terms into the **Keywords** field. The nodes displayed in the tree will be filtered to include only those that match the search terms as you type them. The **Name** and **Descrip** checkboxes indicate whether you want your search terms to match tables/schemas by name and/or by description. The **And/Or** toggle button determines whether each displayed table/schema has to match all of the search terms or just one. So for instance if you want to list just two tables, you can type in both table full names, uncheck the **Descrip** control, and set the toggle button to "Or".

To the right of the tree is a tabbed panel giving detail relating to the currently selected tree node. Select a tree node by clicking on it. Each tab has a little circle next to the title which may be empty or filled, according to whether it contains information. If it's empty it's either because that information doesn't make sense for the selected node in the tree (e.g. if a schema is selected, there is no column information) or because the information has not been retrieved from the service; either it's on its way and will be filled in later, or there's an error. The tabs are as follows:

### Service

Provides various items of information about the TAP service itself, gathered from the service and its registry record if available. This includes the service's name, Service URL, Reference URL, curation and contact information, detailed description, declared data models, and listing with description of any service-specific User-Defined Functions it supports. This may provide some useful reference information for writing queries (available UDFs especially). If a Reference URL or Examples URL are provided, clicking on the URLs or the little link icon () in the panel will try to open the relevant pages in an external web browser (or you can copy and paste them yourself). The content of this tab is not sensitive to which node of the tree is currently selected.

### Schema

Provides a textual description of the currently selected schema, or the schema of the currently selected table. A schema is just a higher-level grouping of tables within the TAP service. Depending on how the service has arranged its tables this may or may not provide useful information that's not available at the table level.

### Table

Provides a textual description of the currently selected table (if any).

### Columns

Tabulates a list of all the columns in the currently selected table; column names, data types, units, descriptions, UCDs, and possibly other metadata are shown. This is generally the information you will need to know about a table before you can write queries. You can sort the entries by clicking on the column header (clicking cycles between *ascending*, *descending*, and *unsorted*), so for instance to list the columns in alphabetical order click on the **Name** column header, or to group together all the columns with units of `mag` click on the **Unit** column header. If no table is selected in the tree, this tab will be empty.

### FKeys

If the service provides information about foreign keys (links between fields in different tables in the database) they will be tabulated here for the currently selected table. If any are present, they may help you to formulate efficient queries. If no table is selected in the tree, this tab will be empty.

### Hints

This panel gives you a very basic "Cheat Sheet" for writing ADQL queries. There are just a few hints to jog your memory about the required syntax for some common operations. But the very best advice it gives you is to use the **Examples** menu, and the service-provided examples (from the Examples menu or via a link) if available.

The **Service Capabilities** panel shows capability metadata that the service provides about itself. This has to be loaded from the server and may not appear immediately. It contains the following information:

### Query Language

Shows what query languages, and what versions, are available. If there is more than one listed, you can choose which language you want to submit your query in, but if you choose something which is not a variant of ADQL it may not work.

### Max Rows

Selector which shows the maximum number of output rows that the service is willing to deliver as the result of a query under normal circumstances. Entries in this list value may be marked "(default)" or "(max)". You can change this value by typing in a different number, as long as it does not exceed the server's maximum. This value is here to protect the user, as well as the service, from inadvertently requesting an unduly large output table. Note that if the construction "TOP *nnn*" is used in the ADQL, the *nnn* limit may override the value supplied here.

### Uploads

Indicates whether table uploads are permitted, and perhaps what is the largest upload size (in terms of rows and/or bytes) which will be accepted. See below for more about table uploads.

The **ADQL Text** panel is where you can enter the actual query text for the request. It has the following parts:

### Mode selector

Provides the following options to control how your job is executed: Synchronous, Asynchronous or Quick Look.

- **Synchronous** means that TOPCAT sends the query to the server and waits for the result as the response of the same HTTP request. It's simpler and faster than async, but if the job takes a long time, the request may time out and the results be lost.
- **Asynchronous** means that TOPCAT sends the query to the server and gets a message back telling it where to look for information about the job's progress; the application then keeps polling the server to find out when it has completed.
- **Quick Look** runs the query as for Synchronous mode above, but when the result comes back it is just displayed in a new popup window rather than loading it into TOPCAT for further analysis.

Synchronous mode is suitable for short jobs (it may typically execute with a few seconds less delay), but for longer jobs asynchronous is more reliable. It is possible to submit an asynchronous job on one day from one machine, and pick up the results on a different day from another machine, using another invocation of TOPCAT or of a different application altogether. Quick look is good if you just want to see the result but don't want to do any detailed analysis of it. One example is if you just want to know the number of rows in a table (`SELECT COUNT(*) FROM table`); the result of the query is the row count in a one-column, one-row table, and there's not much point loading that table into TOPCAT.

### Edit toolbar

This contains a number of actions that affect the **Text Entry** panel:

#### Add Tab

Adds a new blank editing tab, so you can start editing a new query without affecting any existing ones.

#### Copy Tab

Adds a new editing tab with text initially copied from that in the currently visible one.



### Remove Tab

Deletes the currently visible tab.



### Title Tab

Sets the title of the currently visible tab. By default the titles are consecutive integers (the first is "1", the second "2" etc), but you can give them more meaningful names if you like.



### Clear

Erases all the text in the currently visible tab. The text can be retrieved using the Undo action.



### Undo

Undoes the most recent editing action in the currently visible tab. You can do the same thing from the keyboard using **Ctrl-Z**.



### Redo

Reverses the most recent Undo action in the currently visible tab. You can do the same thing from the keyboard using **Ctrl-Shift-Z** or **Ctrl-Y**.



### Insert Table

Inserts into the text at the current cursor position the name of the table (if any) that is currently selected in the metadata tree above. This is a convenience that may save some typing, especially for long or complicated table names.



### Insert Columns

Inserts into the text at the current cursor position a comma-separated list of the names of any columns that are currently selected in the **Columns** tab of the Metadata panel above. In order to select columns, click on the rows in the Columns panel; you may need to hold down the Ctrl key while clicking or something similar (depends on your OS) to select more than one. Columns are inserted in the text in the order they were selected. This is a convenience that may save some typing, but note that you may still need to edit the inserted text, for instance to add table name/alias prefixes.



### Parse Errors

If this button is enabled, it means that the parser has detected an error in the currently visible ADQL text. The error will usually be highlighted in pink in the editing panel. Clicking this button pops up the actual parser error message, which may give you more idea what's wrong.

These actions are also available from the **Edit** menu.

## Text Entry panel

This panel is where you type ADQL text that forms the query to send to the TAP service for execution. The **Edit** actions described above, along with Undo/Redo actions from the keyboard, can also help. You can have multiple queries on the go at once in different tabs; add/remove/retitle tabs with the toolbar actions, and move between them by clicking on the tab headers at the top. The tab you can see is the one that will be executed if you hit the **Run Query** button.

As you enter the query text, it will be checked for syntax errors. If an error is found, the place in the text which appears to have caused it will be highlighted in pink (as in the figure above).



) button. The checking itself is reasonably reliable, but the position of the highlighted error, and the text of the error message, can sometimes be a bit misleading, so don't take the details too seriously. The error highlighting is just used as a warning, you are not prevented from submitting a query marked as erroneous, since in some cases TAP services may accept ADQL which is not strictly correct, and in some cases the error checking may not be perfect. Note also that for various reasons the service may not be able to accept all queries that count as syntactically valid ADQL, so even if TOPCAT doesn't report any errors, the service may still respond with an error message.


### Examples line

Clicking on the **Examples** button at the bottom left of the panel will pop up a menu that allows you to select from a number of example ADQL queries. If you select one, the relevant ADQL text will be pasted into the editing panel; you should be able to hit the **Run Query** button straight away and execute it to get a result. In some cases the example text is influenced by the table you're currently looking at in the metadata panel. Until you are experienced with ADQL, starting from one of these examples and editing it to taste is often a good way to write a query of your own.

The Examples menu is divided into a number of sub-menus, though not all will be enabled for all services. The sub-menus are:

- **Basic:** Basic examples that should be applicable to most services.
- **Upload:** Examples that involve uploading a tables from TOPCAT to the TAP server and executing queries using it. Only enabled if the service supports table uploads and at least one table is currently loaded into TOPCAT.
- **Service-Provided:** Examples provided by the TAP service. These are likely to provide some insight into the kinds of query that make sense and make good use of the particular tables available from the service, so it is often instructive to look at these. Only enabled if the service actually provides some examples (at time of writing, not many do, but hopefully this will improve). *Note to TAP service implementers: at time of writing, there are two variants of the standard way for services to provide examples (DALI sec 2.3 and TAPNotes sec 4.2.1), and it's not clear which is preferred. For now, TOPCAT will recognise either.*
- **TAP\_SCHEMA:** Examples using the tables in the TAP\_SCHEMA schema, which describes metadata about the tables in the service. You can use these to do things like count tables or columns in the current service, or locate tables with columns having particular physical meaning.
- **ObsTAP:** Examples using the IVOA ObsCore data model (`ivoa.obscore` table), which describes astronomical observations in a standardised way. You can use these to do things like search for all observations at a particular sky position, or observation time, or waveband etc. Only available if the service declares support for the ObsCore data model.
- **RegTAP:** Examples using the IVOA RegTAP data model (tables in the `rr` schema). You can use these to do things like search for all TAP services containing tables having columns with particular physical meanings, or all cone search services relating to particular wavebands, etc. You can use queries like this to perform more sophisticated searches for TAP services than you can do using the Select Service tab. Only available if the TAP service declares support for the RegTAP data model (i.e. if it hosts an IVOA searchable registry).

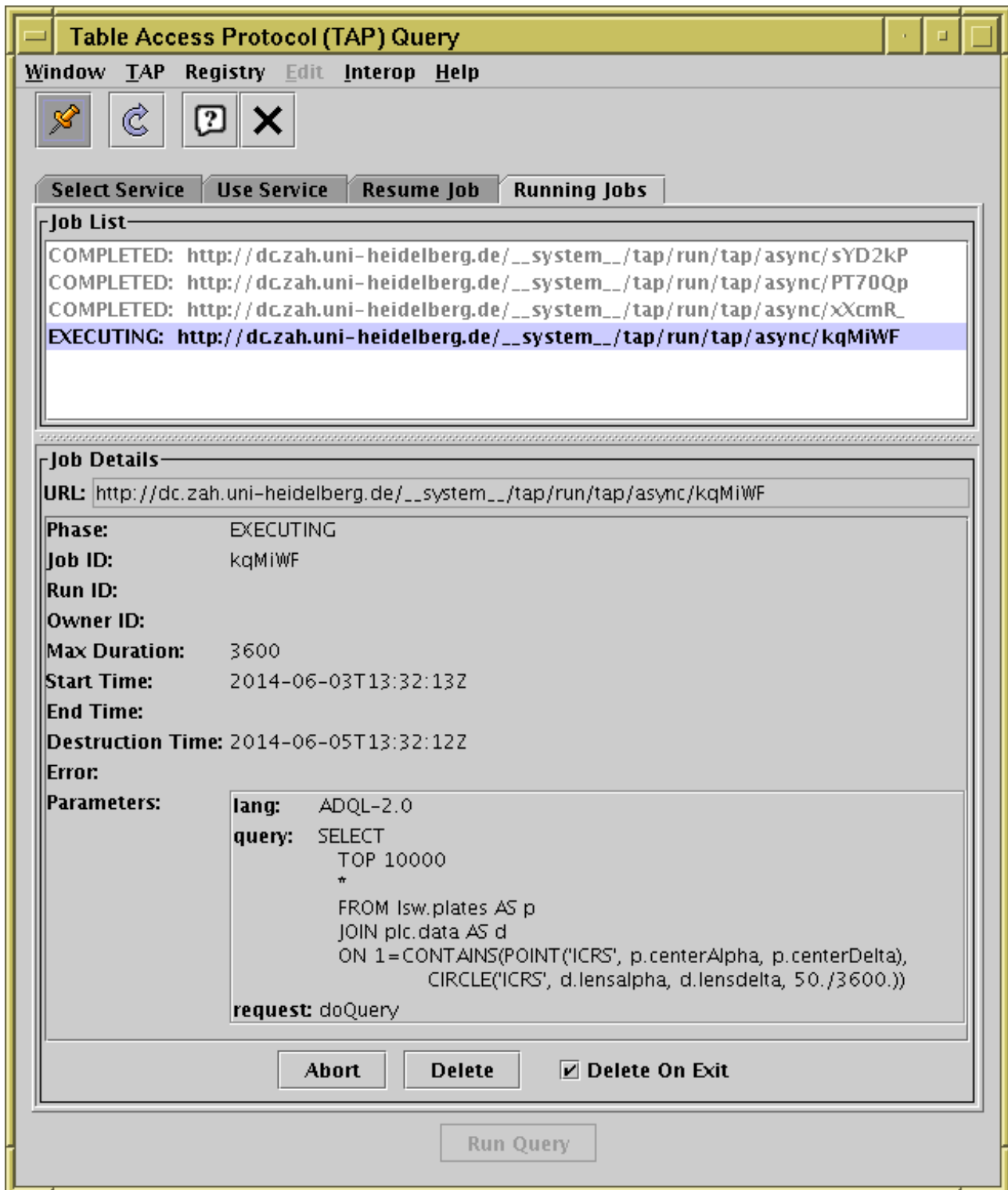
When you select an example from one of the sub-menus, its name and title will be displayed to the right of the Examples button. Little left and right arrow buttons allow you to cycle through the examples in the current submenu so you can browse what's available.

Some examples come with additional explanation on the web. If the example you're currently looking at has such additional documentation then the **Info**  button to the right will be enabled, and clicking on that should open the relevant page in a browser on your desktop.

Some TAP services permit **Table Uploads**. What this means is that you can upload tables from TOPCAT into the remote database and perform queries on your table directly. In this way you can perform joins between your own tables (anything loaded into TOPCAT) and any of the tables in the remote database. This is an extremely powerful feature. To take advantage of it, you just need to use a special form for the table name to reference a TOPCAT-based table: you write "TAP\_UPLOAD.t<n>", where <n> is the ID number of the table in TOPCAT, that is the number before the colon in the Control Window Table List. So, if a table identified as "1: messier.xml" in the table list, you can reference it in ADQL as "TAP\_UPLOAD.t1" - see the **Upload** sub-menu of the **Examples** menu described above for some examples. Note the table uploaded in this way is the Apparent Table corresponding to the given ID number, i.e. current subset and column selections apply. It's a good idea to ensure that any table you are uploading has columns with sensible names, otherwise the service may rename the columns or otherwise have trouble handling it.

Having entered suitable query text, click the **Run Query** button at the bottom of the window to submit the job. What happens then depends on the **Query Mode** described above: in **Synchronous** mode, the Load Window will appear with a load progress indicator, and when it's done the table will be loaded into TOPCAT in the usual way. In **Asynchronous** mode, you will be taken to the Running Jobs tab where you can see the progress of your submitted job; when it's ready it will be loaded into TOPCAT. In **Quick Look** mode, a window will pop up to display the results when they are ready.

#### A.9.8.3 Running Jobs tab



### TAP window with Running Jobs tab visible

The **Running Jobs** tab of the TAP load dialogue shows a list of the asynchronous jobs you have submitted, with information about their current progress.

The upper part of the screen shows a list of the jobs submitted, along with a one-word summary of their status (their *Phase*). If one of these jobs is selected, its details will be shown in the lower part. By default the most recently submitted job will be shown, but you can select a job by clicking on it.

The information displayed is retrieved from the server, and can only be seen as long as the job has not been deleted, either by you the user, or as a result of server policy (for instance a maximum job



lifetime). The following items, if provided by the service, are shown:

**URL**

The URL at which the job information is held on the server. This value can be used later in the Resume Job tab, or entered directly into a web browser to see job information.

**Phase**

The current stage of job processing. This usually progresses in the sequence QUEUED, EXECUTING, COMPLETED (or ERROR).

**Job ID**

Job identifier, unique to the server. Forms part of the URL.

**Run ID**

Run identifier, often blank.

**Owner Id**

Owner identifier, blank unless user authentication is in use.

**Max Duration**

The maximum amount of time in seconds that the server will allow the job to run for. This is "wall-clock" time rather than CPU time.

**Start Time**

Time at which actual processing of the job (as opposed to queuing) started.

**End Time**

Time at which processing of the job finished.

**Destruction Time**

Time in the future at which the server will delete the job, if it has not been manually removed by then.

**Error**

Message and possibly other information about error status if an error occurred during processing. This will normally be blank unless **Phase** has the value "ERROR".

**Parameters**

Displays parameters which were supplied to define the job to the server. This will normally include the text of the ADQL being executed.

You can cut and paste from these items, so for instance if you want to take the URL and paste it into a web browser you can do.

There are three buttons at the bottom of the screen which affect the status of the currently displayed job:

**Abort**

This button stops the job from executing. It cannot be restarted. The job still exists on the server however, so it can be examined in this window.

**Delete**

This button stops the job as for Abort, but also removes all trace of it from the server. No details can be seen any more.

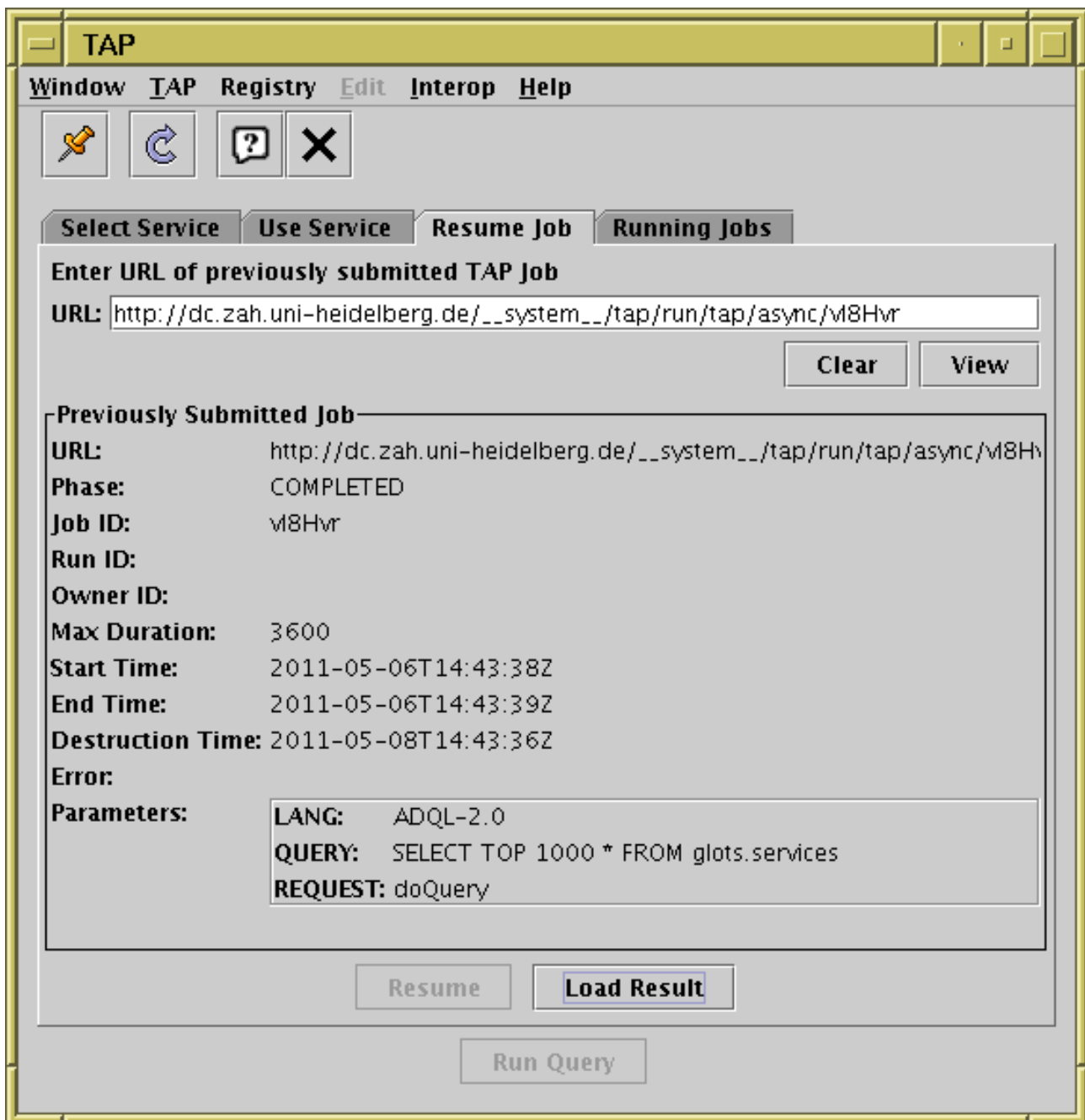
**Delete On Exit**

This checkbox determines whether the job will be deleted from the server when TOPCAT exits. If you want to come back to a long-running job in a later session, you should uncheck this so that the job remains on the server. The default action is normally that jobs are deleted on exit (i.e. this box starts off checked), but you can change the default by using the **Deletion** menu.



If no jobs are currently visible (none have been submitted or they have all been deleted), this tab is inaccessible.

#### A.9.8.4 Resume Job tab



**TAP window with Resume Job tab visible**

The **Resume Job** tab of the TAP load dialogue allows you to continue execution of an asynchronous job which was started outside of this run of TOPCAT, either during an earlier TOPCAT run, or using some other mechanism. This may be useful if you have submitted a long-running job on an earlier day or on a different machine and want to monitor its progress or pick up the results.

To use it, enter the Job URL into the **URL** field at the top and click the **View** button. If a job with that URL exists, its details will be displayed in the panel below, in the same format as for the Running Jobs tab.




Depending on whether the job has completed or not, either the **Resume** or **Load Result** button at the bottom of the window will be enabled. You can click the appropriate one either to monitor its progress further or to pick up the completed result.

## A.10 Other Windows

### A.10.1 CDS Upload X-Match Window

**CDS Upload X-Match**

Window Search Help

**Remote Table**

VizieR Table ID/Alias: 2mass6x

Name: ll/281/2mass6x

Alias: 2MASS6X

Description: 2MASS-6X catalog

Row Count: 24,023,702

Coverage: 0.018615723 (order 6)

**Local Table**

Input Table: 1: firstfits

RA column: POS\_EQ\_RA degrees (J2000)

Dec column: POS\_EQ\_DEC degrees (J2000)

**Match Parameters**

Radius: 3.5 arcsec


Find mode: Best

Rename columns: Duplicates Suffix: \_f

Block size: 20000

Go Stop

## CDS Upload X-Match Window

The CDS Upload X-Match Window allows you to join a local table with any table provided by the VizieR database of astronomical tables or with SIMBAD. You can access the window from the main Control Window using the **CDS Upload Match** button () in the toolbar, or the **Joins** or

**VO** menus. This window is an interface to the excellent CDS X-Match service provided by the Centre de Données astronomiques de Strasbourg (CDS). The service is very fast, and in most cases it is the best way to match a local table against a large external table hosted by a service. In particular, it is almost certainly much better than using the Multi-Cone window, though it's less flexible than TAP.

The local table is uploaded to the X-Match service in blocks, and the matches for each block are retrieved in turn and eventually stitched together to form the final result. The tool only uploads sky position and an identifier for each row of the input table, but all columns of the input table are reinstated in the result for reference. These details are mostly transparent when using the service, though it may help to understand how the progress bar moves.

The window has three panels, described below.

The **Remote Table** panel allows you to select the CDS table against which to match. The table must be entered by name in the **VizieR Table ID/Alias** field, and may be in one of the following forms:

- "simbad", to indicate the SIMBAD database of known astronomical objects.
- A VizieR table identifier, for instance "II/246/out". *Note* these must name the table not just the catalogue (which may contain several tables), so for instance "II/246" will not work. No help is currently provided in this window to search for these IDs, you have to use one of the services or web pages provide by CDS to locate them. A good tip is the table searching facility at the TAPVizieR service, <http://tapvizier.u-strasbg.fr/adql/>.
- A VizieR table alias, for instance "2mass". A few tens of these are defined for the largest and most-used tables in VizieR, and these can be selected by clicking on the selection box.

When a table name has been entered, additional information (Name, Alias, Description and Row Count) will be downloaded from CDS and displayed in the lower part of the Remote Table panel. If nothing shows up here, then the table name is not legal for the X-Match service, and the **Go** button will be disabled.

Sky coverage information is also displayed, including the proportion of the sky covered by the table and a graphical indication of the covered regions on the sky.

The **Local Table** panel allows you to indicate the local table (the one loaded into TOPCAT) that you want to match. You must select the required table and indicate its Right Ascension and Declination columns.

The **Match Parameters** panel supplies the other information about how the match will operate:

### Radius

The maximum distance between a local table and remote table position to count as a match. There is a maximum enforced on this value by the CDS X-Match service, currently 2 arcminutes.

### Find mode

Determines in what form the result is generated and used. The following options are available:

#### Best

Load a new table with one row for each local row that matches a remote row, giving the

closest match. Unmatched local rows are not included.

#### **All**

Load a new table with one row for each remote row that matches a local row. The match is symmetric between local and remote tables.

#### **Each**

Load a new table with the same number of rows as the local table, in the same order. The best remote match, if any, appears alongside the local row, but if there is no match the remote columns are blank.

#### **Best Remote**

Load a new table with one row for each remote table row that matches the local table. *Note* there is currently a bug in this mode; if the match is done in N blocks then a remote row may appear up to N times rather than just one. This may be fixed in a future release, but to avoid it you can make sure that the **Block size** is greater than the row count.

#### **Add Subset**

No new table is loaded, but a new Row Subset is added to the local table indicating which rows had at least one match to the remote table.

#### **Rename columns**

If columns from two input tables are combined, it's possible that some may share the same name. Multiple columns in a table with the same name can cause confusion, so this selector allows you to determine whether and how such name clashes are dealt with. Columns from the remote table can be renamed by appending a specified suffix, either always or only in case of duplicate names.

#### **Block size**

The number of rows uploaded to the X-Match service at a time. This should not affect the result, though it may affect performance.

Large block sizes tend to be good (up to a point) for reducing the total amount of time a large xmatch operation takes, but they can make it harder to see the job progressing. There is also the danger (for ALL-type find modes) of exceeding the return size limit, which will result in truncation of the returned result. At time of writing, the upload limit is 100Mbyte (about 3Mrow), and the maximum return size is 2Mrow.

When all the fields have been filled in, hit the **Go** button and the match will start; progress will be logged in the progress bar at the bottom of the window. To stop the match before it has completed, hit the **Stop** button. Note that the **Go** button is only enabled when a legal table name has been entered at the top.

The remote table in most cases contains only a subset of the the columns in the relevant VizieR table, including the most useful ones. The service currently provides no straightforward way to acquire columns which are not returned by default.

The following item is available in the toolbar and **Search** menu:

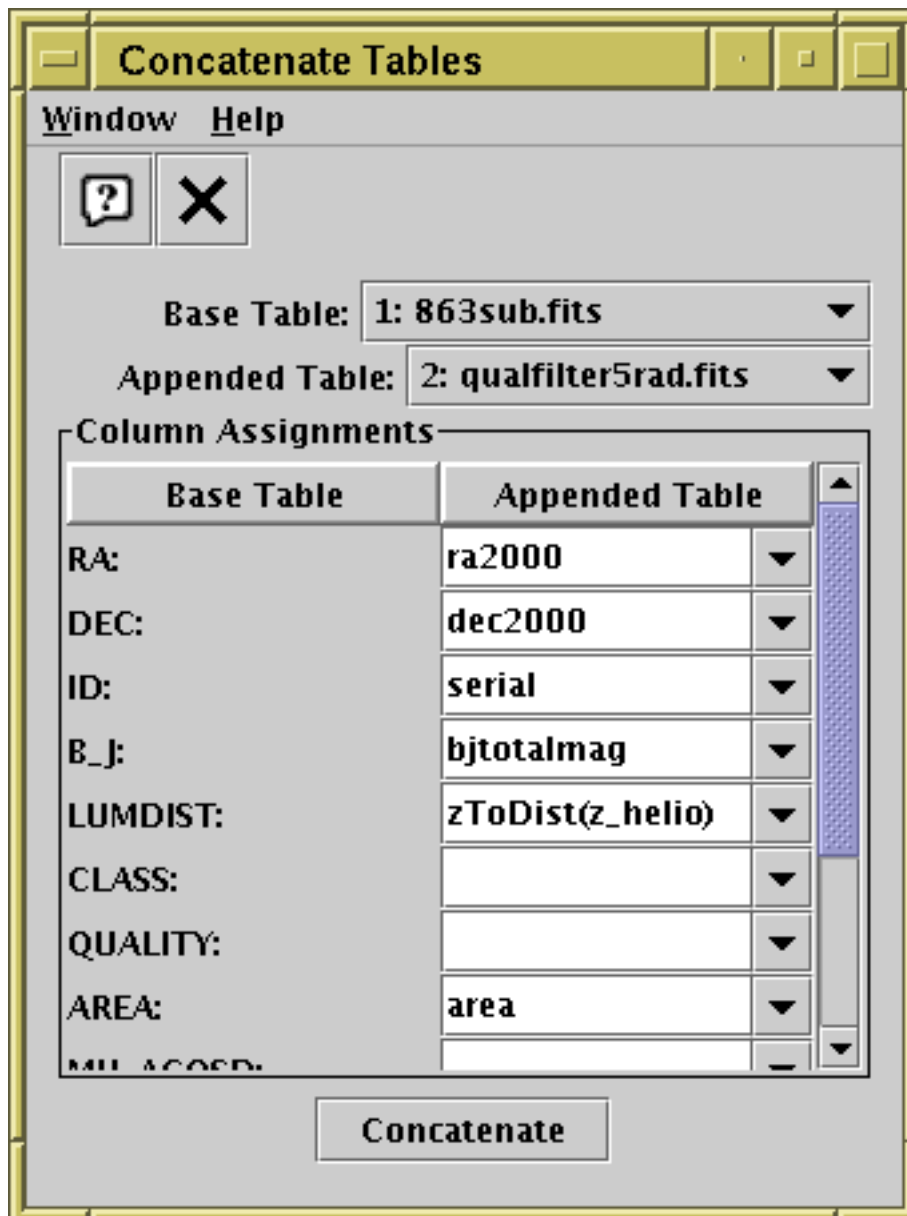
#### **Use Service Coverage**

If this option is selected, then the sky coverage of the service is used to pre-filter local table rows. In this case those rows that fall outside the coverage area of the remote table are not uploaded for matching, since it is known that they will not match. This option should therefore not affect the result, but improve performance, and as such should generally be turned on (the default).


**Acknowledgement:** CDS note that if the use of the X-Match service is useful to your research, they would appreciate the following acknowledgement:

*"This research made use of the cross-match service provided by CDS, Strasbourg."*

### A.10.2 Concatenation Window



### Concatenation Window

The Concatenation Window allows you to join two tables together top-to-bottom. It can be obtained using the **Concatenate Tables** item (  ) in the Control Window **Joins** menu.

When two windows are concatenated all the rows of the first ("base") table are followed by all the rows of the second ("appended") table. The result is a new table which has a number of rows equal to the sum of the two it has been made from. The columns in the resulting table are the same as those of the base table. To perform the concatenation, you have to specify which columns from the appended table correspond to which ones in the base table. Of course, this sort of operation only makes sense if at least some of the columns in both tables have the same meaning. This process is discussed in more detail in Section 5.1.

The concatenation window allows you to select the base and appended tables, and for each column in the base table to specify what quantity in the appended table corresponds to it. You can either select a column from the appended table from the selection box, or type in an expression in the expression language referring to columns from the appended table. If you leave the field blank, the column in question will have all null entries in the resulting table. Only suitable columns are available for choosing from these column selectors, that is ones matching the data type of the base column.

In some cases these column selectors may have a value filled in automatically if the program thinks it can guess appropriate ones, but you should ensure that it has guessed correctly in this case. If the two tables have essentially the same structure, the corresponding columns should all get filled in automatically.

When you have filled in the fields to your satisfaction, hit the **Concatenate** button at the bottom of the window, and a new table will be created and added to the table list in the Control Window (a popup window will inform you this has happened).

The result is created from the Apparent (Section 3) versions of the base and appended tables, so that any row subsets, hidden columns, or sorts currently in force will be reflected in the output.

### A.10.3 Activation Window

**Set Activation Action**

Window Help

$f(x)$  ? X

? ☐ No Action

☒ Display Cutout Image

☐ View URL as Image

☐ View URL as Spectrum

☐ View URL as Web Page

☐ Transmit Row

☐ Transmit Coordinates

☐ Execute Custom Code

Cutout Service: SuperCOSMOS All-Sky Red

RA column: RA degrees

Dec column: DEC degrees

Width/Height in Pixels: 150 (0.67 arcsec)

Image Location column:

Image Format: FITS

Image Viewer: Basic viewer (internal)

Spectrum Location column:

Spectrum Viewer: All Clients

Web Page Location column:

Browser Type: basic browser

Target Application: All Clients

RA Column: RA degrees

Dec Column: DEC degrees

Target Application: All Clients

Executable Expression:

OK Cancel

### Activation Window

The Activation Window allows you to configure an action to perform when a table row is activated (Section 8) by clicking on a row in the Data Window or a point in the Plot Window. It can be obtained by clicking the **Activation Action** selector at the bottom of the properties panel in the Control Window.

You have various options for how to define the action. On the left of the window is a list of options;

you have to choose one of these to determine what kind of action will take place. When you click on one of these options the corresponding controls on the right hand side will become enabled: use these to select the details of the action and then click the **OK** button so that subsequent activation events will cause the action you have defined (or **Cancel** so that they won't). When you click OK the **Activation Action** in the control window will indicate the action you have configured.

The available options are as follows:

### No Action

If this is selected, no special action will take place when a row is activated. This is the default.

### Display Cutout Image

This option presents an easy-to-use way of popping up a cutout image from an image server displaying a region of sky around an activated row. You need to select the columns in your table which represent Right Ascension and Declination, including the units in which they are entered in the table (TOPCAT may be able to guess some or all of this information based on column names, UCDs and unit values, in which case it will enter its guesses in the selectors for you to accept or change). You also need to select the size in pixels of the image you want to see and the name of the survey which will supply the image from one of the listed ones:

- SuperCOSMOS All-Sky Blue
- SuperCOSMOS All-Sky Red
- 2MASS Quick-Look J-band
- 2MASS Quick-Look H-band
- 2MASS Quick-Look K-band
- SDSS Colour Images (note does not provide all-sky coverage)

When you activate the row, the program will attempt to contact the web server which provides these images, retrieve the image, and display it in an image viewer window. Unfortunately at present the viewer window is an internal one; it does not display in an external SAMP viewer.

### View URL as Image

This option is suitable if one of the columns in your table gives the location (filename or URL) of an image file. The image may be in FITS, GIF, JPEG or PNG format, optionally compressed using gzip, Unix compress, or bzip2 format. Select the column which contains the location, and activating a row will pop up an image viewer to display it (if the table is the result of a single or multiple SIA query, the correct column will probably be filled in automatically). The **Image Viewer** selector allows you to choose one of the options in Appendix A.10.3.1 or, if you are registered with a SAMP or PLASTIC hub, any SAMP/PLASTIC image viewer which is also registered - see Section 9.

### View URL as Spectrum

This option is suitable if one of the columns in your table gives the location (filename or URL) of a spectrum file, and you wish to display the spectra in an external SAMP- or PLASTIC-compatible spectrum viewer, such as SPLAT or VOSpec. First select the column which contains the **Spectrum Location** (if the table is the result of a single or multiple SSA query, the correct column will probably be filled in automatically). Then choose the **Spectrum Viewer** appropriately. Once this is done, activating a row will try to send the corresponding spectrum to the selected viewer application(s). The details of what data formats are acceptable depend on the viewer application, but FITS, VOTable and some ASCII table variants are usually OK. Note that TOPCAT, as well as a suitable spectrum viewer, must be connected to a running SAMP or PLASTIC hub for this to work (see Section 9) - there is no internal spectrum viewer.

### View URL as Web Page

This option is suitable if one of the columns in your table gives the location (filename or URL) of a web page; this should normally be in HTML or plain text, but depending on what browser you use other kinds of document may be supported. Select the column which contains the location and the browser which you would like to use for display, and activating a row will try



to pop up a browser window to display it. See Appendix A.10.3.3 for more information about browsers.

### **Transmit Row**

This option causes the row number of an activated row to be transmitted to one or more other applications using an appropriate SAMP or PLASTIC message. This will only work if TOPCAT is registered with a SAMP/PLASTIC hub, and so are one or more other applications which understand that message, and the other application(s) know about the table in question, for instance because it has previously been sent to/from TOPCAT. You may specify that either a single application, or all appropriate ones, will receive the messages. See Section 9 for more explanation.

### **Transmit Coordinates**

This option causes the sky position of an activated row to be transmitted to one or more other applications using an appropriate SAMP or PLASTIC message. This will only work if TOPCAT is registered with a SAMP/PLASTIC hub, and so are one or more other applications which understand that message. An example might be a sky viewing application such as Aladin which can point to a particular region of sky whenever you activate a point. You need to specify the columns which represent (J2000) Right Ascension and Declination, and optionally a particular listener to receive the messages (otherwise all registered ones will). See Section 9 for more explanation.

### **Execute Custom Code**

This option must be used if none of the others (which are fairly restrictive) do what you want. It is highly flexible, but not so easy to use. What you have to do is to write an expression following the rules in Section 7 involving some of the column names which will be invoked when a row is activated. This expression will typically have the effect of popping up an image or a spectrum in a viewer, but, especially if you link in your own functions (see Section 7.9) it can do pretty much anything.

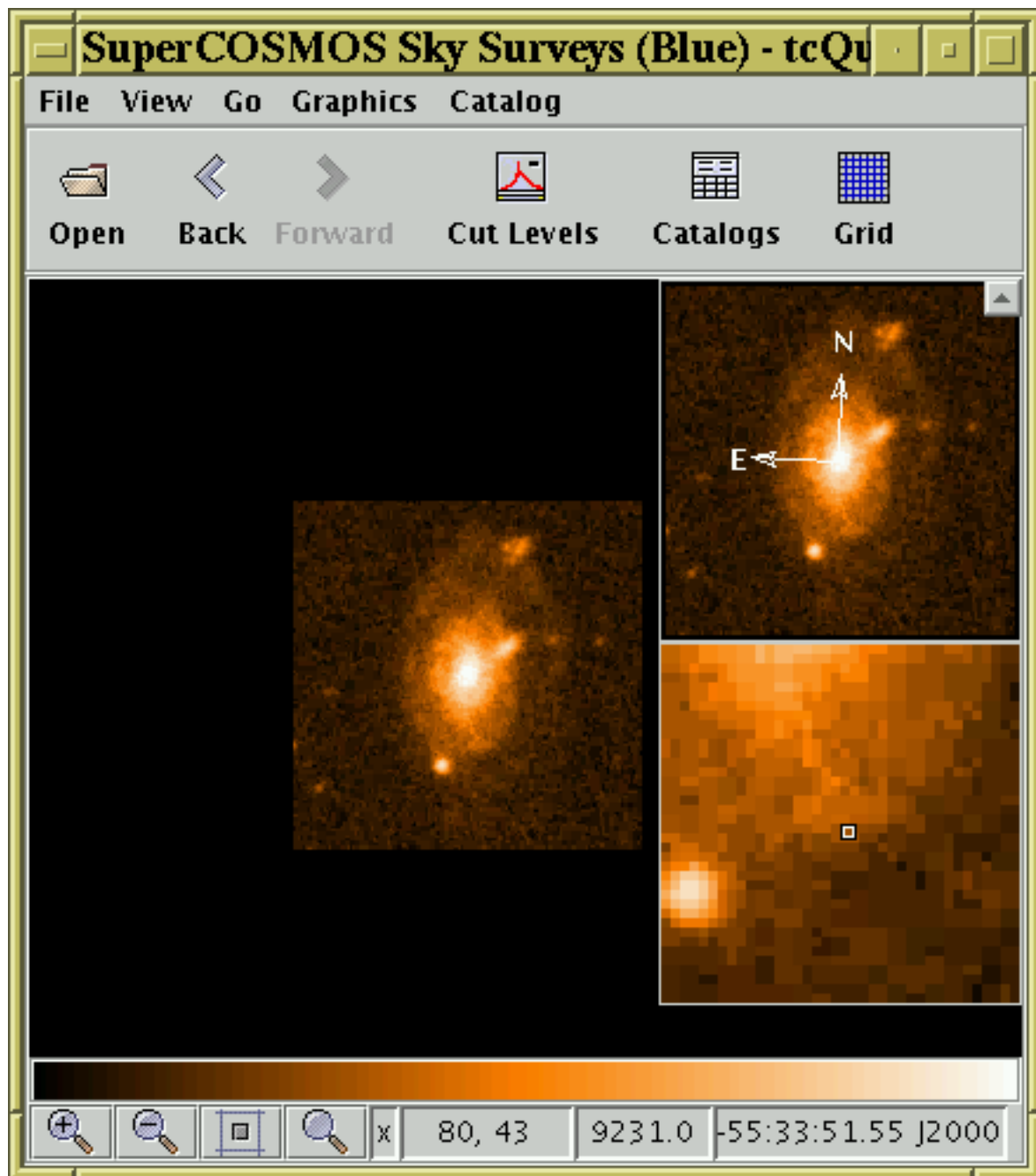
Functions which are expected to be useful for activation actions are described in Appendix B.2 and include some general-purpose ones (`displayImage` and `displaySpectrum` to display an image or spectrum in an external viewer) as well as a few which are relevant to particular survey data, for instance the `spectra2QZ()` function, which will pop up a spectrum viewer displaying all the spectra related to a given row of 2QZ survey data based on the contents of its NAME column.

As the above list shows, most of the activation actions you can define result in a viewer window of some kind popping up. Exactly what kind of viewer is used depends on how TOPCAT is set up and in some cases on your choices. More details of the viewer programs available are given in the following subsections. If these don't do what you want, you can use the **Execute Custom Code** option, perhaps in conjunction with user-defined functions or the `System.exec()` functions described in Appendix B.2, to invoke your own.

## **A.10.3.1 Image Viewer Applications**

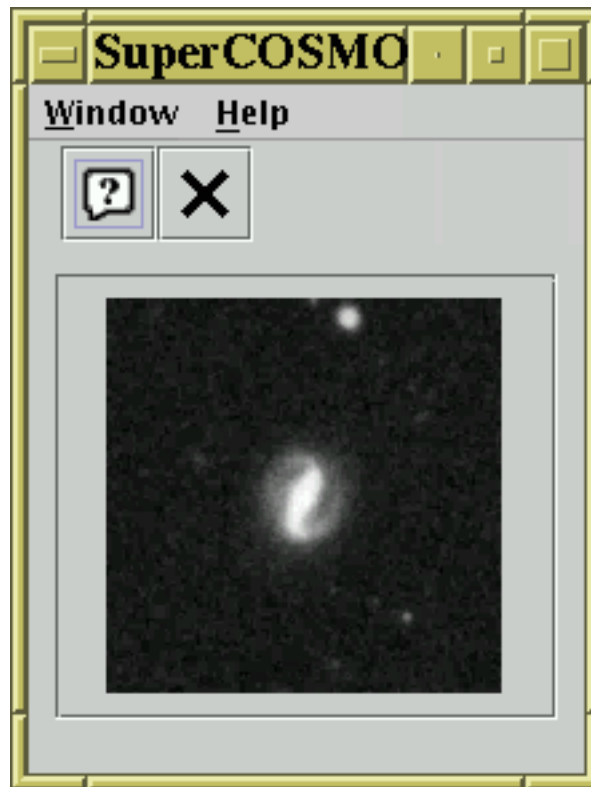
If you choose the **Display Cutout Image** or **View URL as Image** option in the Activation Window, then activating a row will display an image in an image viewer.

The default image viewer is SoG (<http://www.starlink.ac.uk/sog/>), an astronomical image viewer based on JSky, which offers colourmap manipulation, image zooming, graphics overlays, and other features. For this to work JAI, otherwise known as Java Advanced Imaging (<http://java.sun.com/products/java-media/jai/>) must be installed. JAI is a free component available from Sun, but not a part of the Java 2 Standard Edition by default. In operation, SoG looks like this:



### SoG Image Viewer

If JAI or the SoG classes themselves are absent, a fallback viewer which just displays the given image in a basic graphics window with no manipulation facilities is used. The fallback image viewer looks like this:

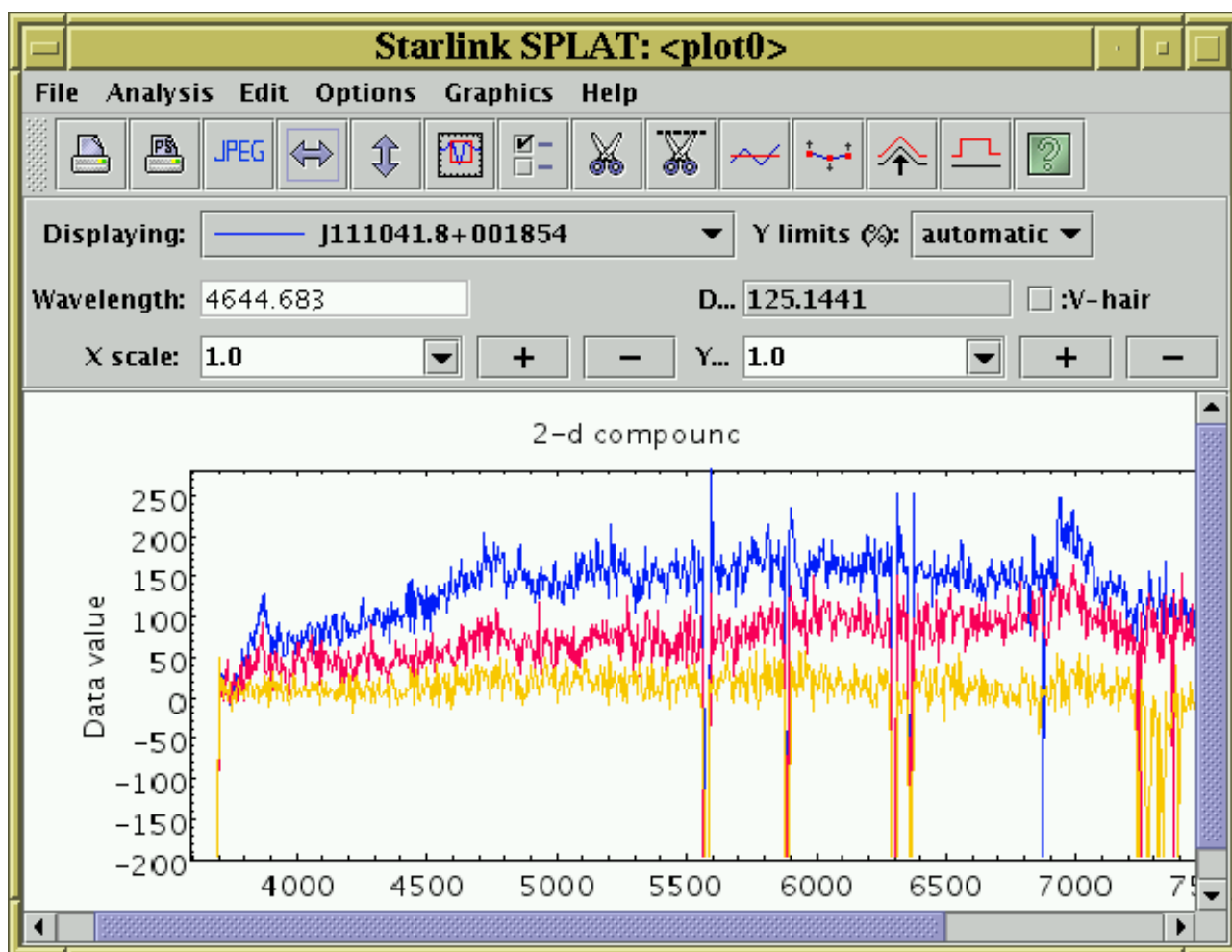


### Fallback Image Viewer

#### A.10.3.2 Spectrum Viewers

If you choose the **View URL as Spectrum** option in the Activation Window, then activating a row will display a spectrum in a spectrum viewer.

The default spectrum viewer is SPLAT (<http://www.starlink.ac.uk/splat/>), a sophisticated multi-spectrum analysis program. This requires the presence of a component named JNIAST, which may or may not have been installed with TOPCAT (it depends on some non-Java, i.e. platform-specific code). There is currently no fallback spectrum viewer, so if JNIAST is not present, then spectra cannot be displayed. In this case it will not be possible to select the **Display Named Spectrum** item in the Activation Window. An example of SPLAT display of multiple spectra is shown below.



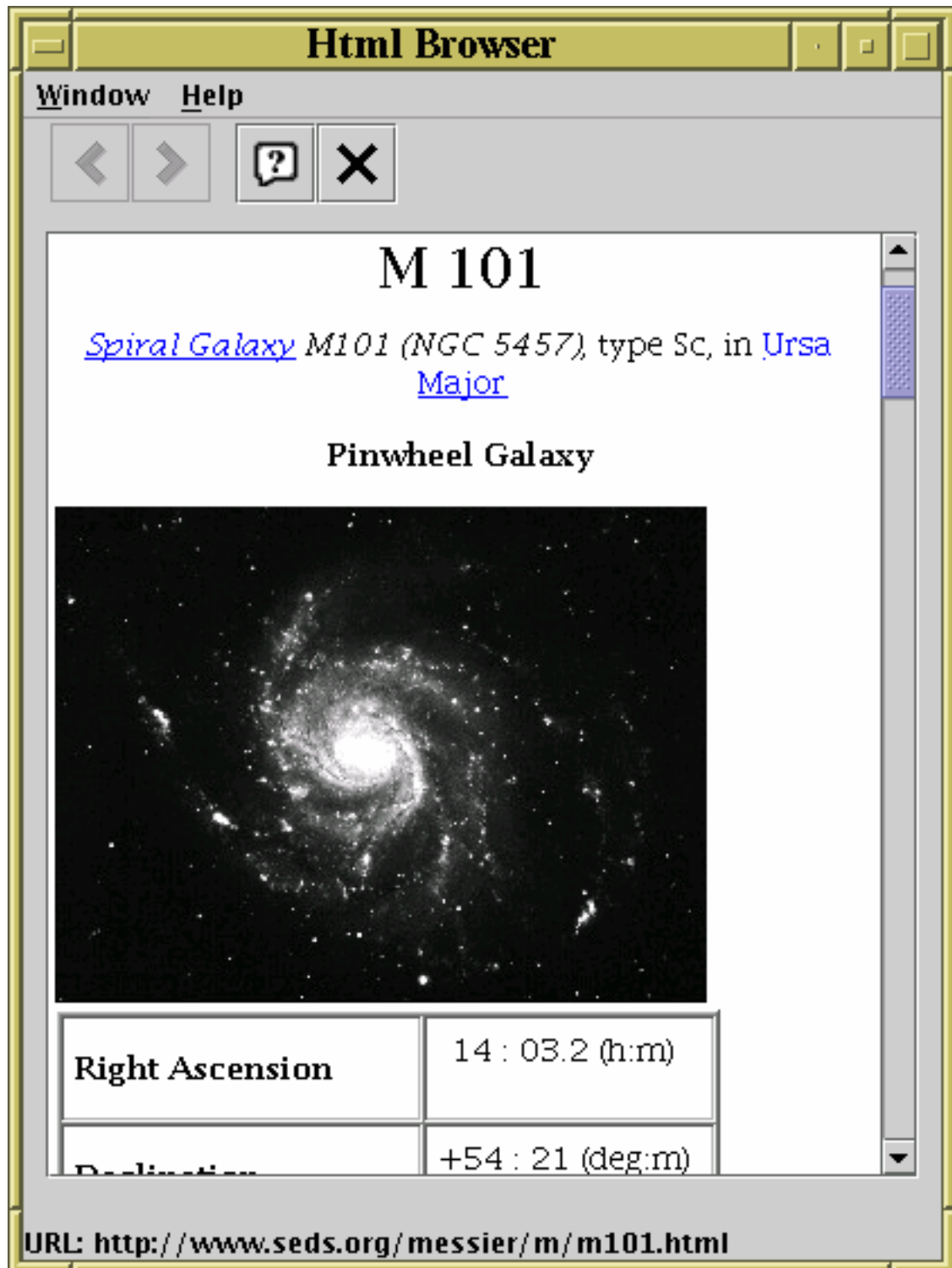
### SPLAT Spectrum Viewer

Full documentation for SPLAT is available on-line within the program, or in SUN/243.

#### A.10.3.3 Web Browsers

If you choose the **View URL as Web Page** option in the **Activation Window**, then activating a row will display the web page whose URL is in one of the columns in a web browser. You are given the option of what browser you would like to use in this case.

The default **basic browser** option uses a simple browser which can view HTML or plain text pages and has forward and back buttons which work as you'd expect. In many cases this is fine for viewing HTML pages, and it is available regardless of the system that you are running TOPCAT on. It looks like this:



### Basic HTML browser

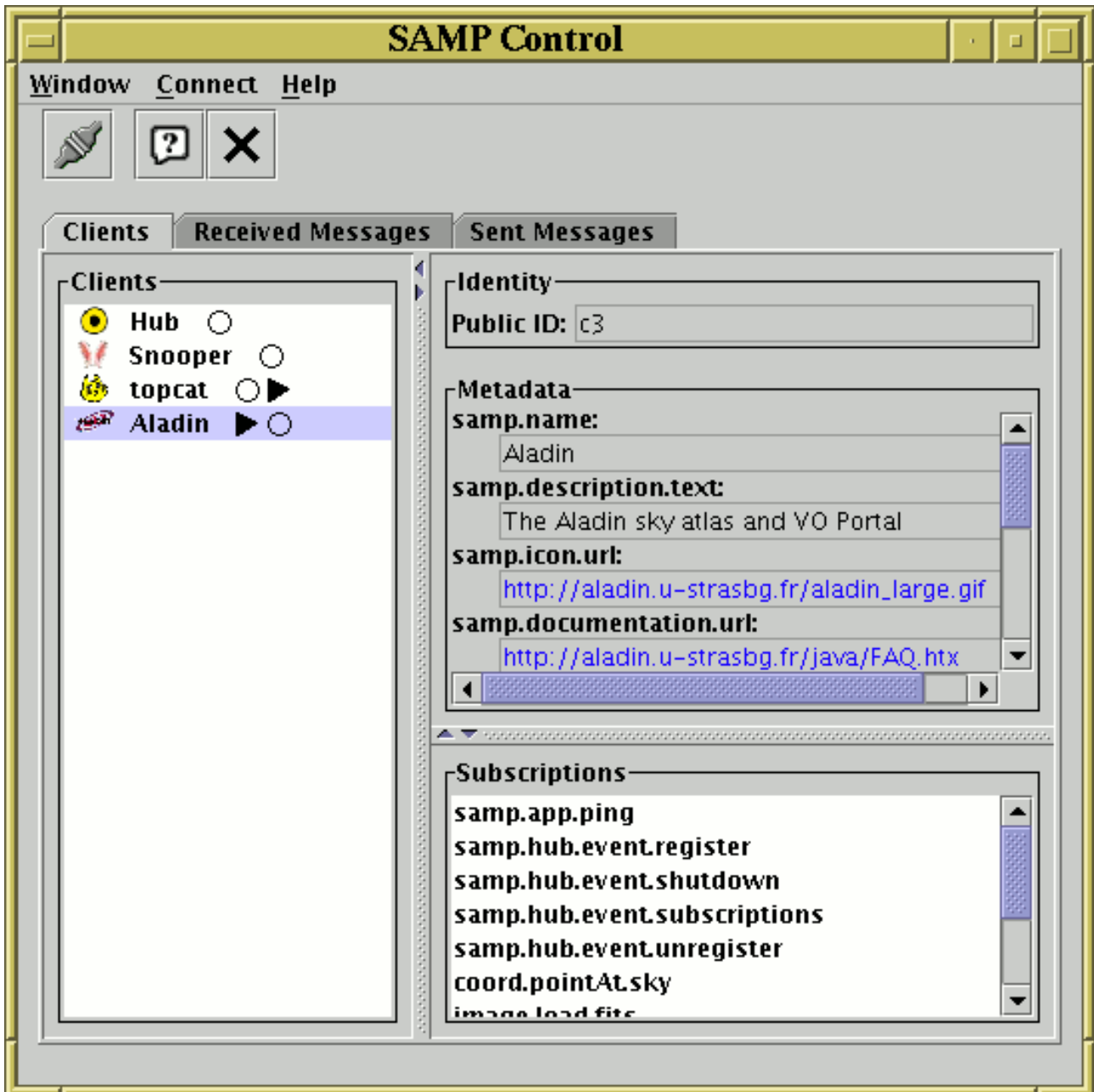
In some circumstances, it's possible to use your normal web browser for web page display instead. The list of browsers currently includes Firefox, Mozilla and Netscape as well as the basic one. Selecting these will generally only work if (1) the browser you select is installed and on your path, (2) you're on some Unix-like operating system, (3) the browser is already running when the action is invoked. In this case, the selected URL should be displayed in an existing browser window rather than opening a new one. Doing it this way has the advantage that your browser can probably display many types of document (perhaps using plugins) as well as HTML.

### A.10.4 SAMP Window

The SAMP Window displays the status of SAMP messaging and allows some control over its configuration. SAMP is a messaging protocol which allows TOPCAT to exchange information with other desktop applications - see Section 9 for details.

The main part of the window is a tabbed panel which displays the status of the SAMP connection in detail.

The first tab is labelled **Clients**:

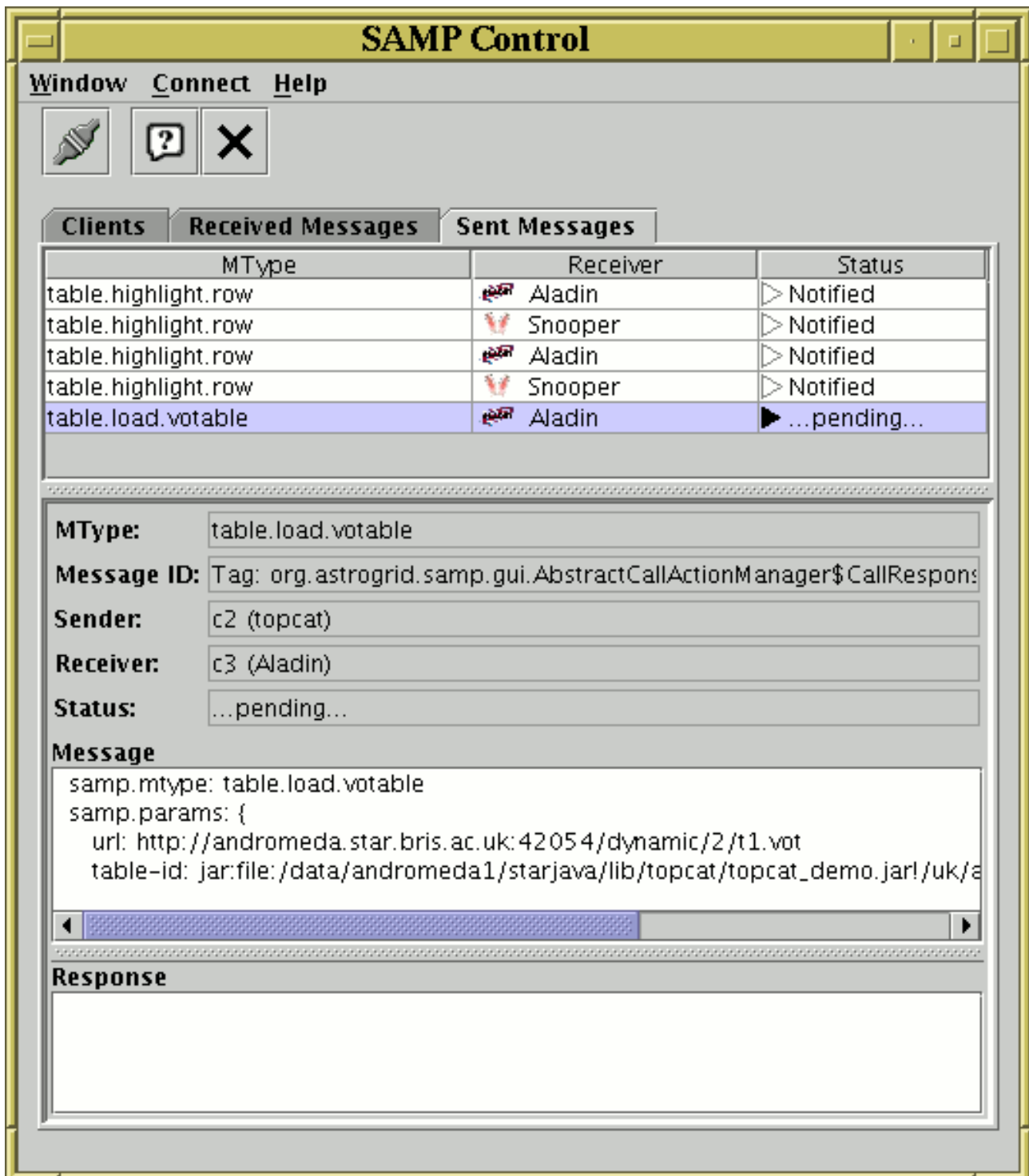


**SAMP Window Clients tab**

This displays details of all applications (including TOPCAT itself) which are currently registered with the hub. If you select a client in the list on the left hand side of the panel, you will see its details on the right hand side. This includes its **Metadata** (which may include things like its name, author, documentation URL etc) and its **Subscriptions** which describes what types of messages (MTypes) it will respond to. Also in the list on the left is for each application a graphical indication of messages recently received by/sent from it, in which TOPCAT is the other partner. These are represented by little arrows before/after the little circle following the application name. In the

figure, a message is currently in progress from TOPCAT to Aladin. More detail on this is shown in the other tabs.

The other two tabs are labelled **Received Messages** and **Sent Messages**. They look very similar to each other:



### SAMP Window Sent Messages tab

They display details of SAMP messages recently sent from or received by TOPCAT. A table near the top contains one row for each message. Shown in different columns are the message's MType (which describes the message semantics), the application receiving/sending it from/to TOPCAT, and a summary of its status. If you click on the row, all the details of the message, including its content and that of the response if one has been received, are visible in the lower panel. Messages

remain in this table for a few seconds after they have completed, but are eventually removed to avoid clutter.

The following toolbar button is available:



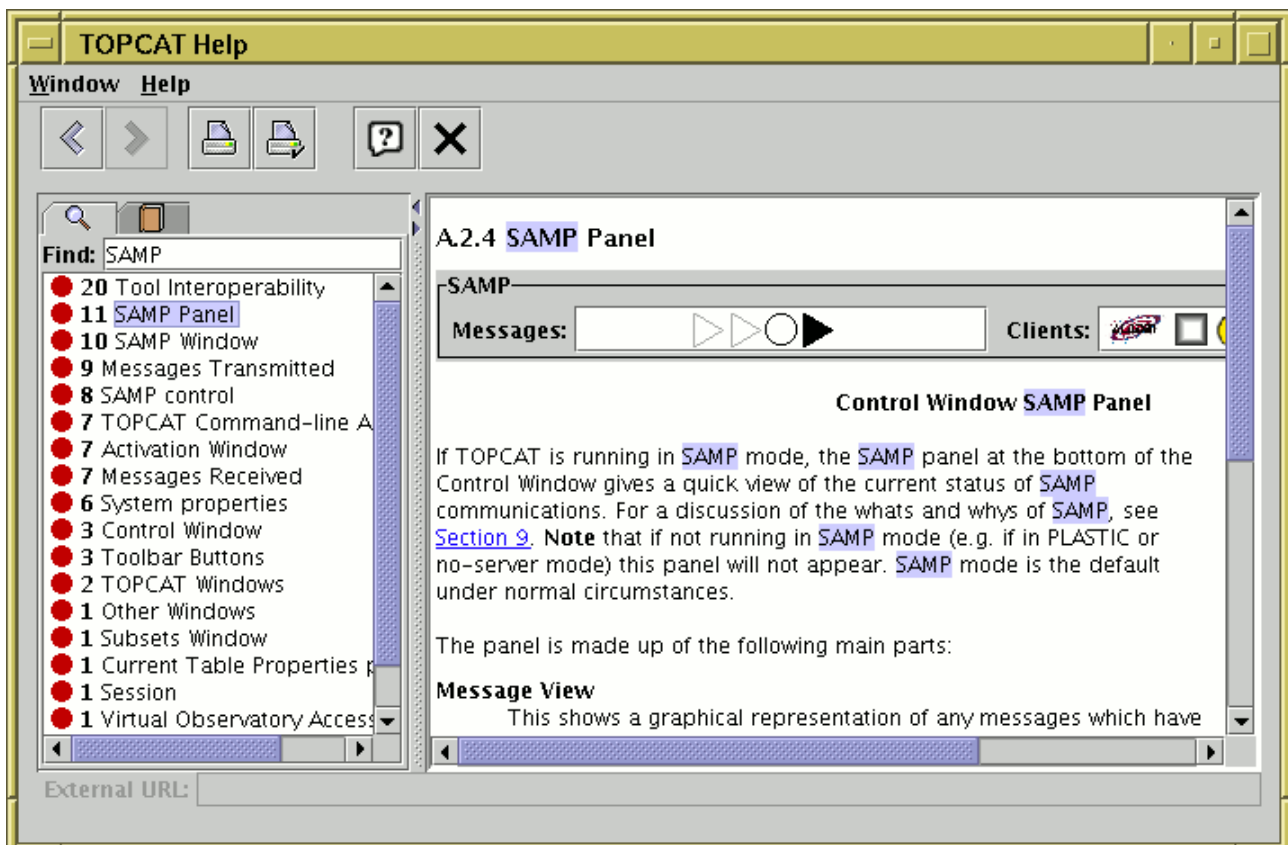
#### Connect/Disconnect

This button controls connection to the SAMP hub; without a connection no SAMP communication can take place. If a hub is running, clicking this button toggles whether TOPCAT is connected or not.

If no hub is running, clicking this button will pop up a dialogue box which allows you to start a hub. You are given the option to start either an **internal** or an **external** hub. An internal one runs in the same Java Virtual Machine as TOPCAT itself, and an external one runs in a separate process. The main important difference is that an internal hub will shut down when TOPCAT does, while an external hub will keep running until it is shut down explicitly. If an external hub is started, a window which shows its status is displayed on the screen; closing this window (which does not belong to TOPCAT) will shut down the hub.

More facilities may be introduced into this window in future releases.

### A.10.5 Help Window



### Help Window

The help window is a browser for displaying help information on TOPCAT; it can be obtained by clicking the **Help** (?) button that appears in the toolbar of most windows. It views the text

contained in this document, so it may be what you are looking at now.

The panel on the right hand side displays the currently selected section of help text itself. The panel



on the left gives a couple of ways of selecting this view:

#### **Text Search**

Allows you to search for words in the manual. Enter a word or words as search terms, and a list of sections which fully or partially match your search terms will be displayed. You can click on one of these to take you to the section that has been found.

#### **Table of Contents**

A hierarchical view of the available help topics. This is the table of contents of the manual; clicking on an entry will take you to the relevant section.

The bar in between the two panels can be dragged with the mouse to affect the relative sizes of these windows.

The toolbar contains these extra buttons:

#### **Back**

Moves backward through the list of topics in the order you have looked at them.

#### **Forward**

Moves forward through the list of topics in the order you have looked at them.

#### **Print**

Pops up a dialogue to permit printing of the current page to a file or printer (but see below).

#### **Page Setup**

Pops up a dialogue to do printer setup.

#### **To Browser**

Displays the currently displayed help page in your normal WWW browser.

Although the printing buttons work, if you want to print out the whole of this document rather than just a few sections you may be better off printing the PDF version, or printing the single-page HTML version through a web browser. The most recent version of these should be available on the web at <http://www.starlink.ac.uk/topcat/sun253/sun253.html> and <http://www.starlink.ac.uk/topcat/sun253.pdf>; you can also find the HTML version in the topcat jar file at [uk/ac/starlink/topcat/help/sun253.html](http://uk.ac/starlink/topcat/help/sun253.html) or, if you have a full TOPCAT installation, in [docs/topcat/sun253/sun253.html](http://docs/topcat/sun253/sun253.html) and [docs/topcat/sun253.pdf](http://docs/topcat/sun253/sun253.pdf) (the single-page HTML version is available here in the HTML version).

The help browser is an HTML browser and some of the hyperlinks in the help document point to locations outside of the help document itself. Selecting these links will go to the external documents. When the viewer is displaying an external document, its URL will be displayed in a line at the bottom of the window. You can cut and paste from this using your platform's usual mechanisms for this.

### **A.10.6 New Parameter Window**

**Add New Parameter**

Window Help

Parameter Name: Epoch

Type: Double

Value: 2003.81

Units: years

Description: Epoch of observations

UCD: TIME\_EPOCH

OK Cancel

### New Parameter dialogue window

The New Parameter window allows you to enter a new table parameter to be added to a table. It can be obtained by clicking the **New Parameter** (+) button in the Appendix A.3.2. A parameter is

simply a fixed value attached to a table and can contain information which is a string, a scalar, an array... in fact exactly the same sorts of values which can appear in table cells.

The window is pretty straightforward to use: fill in the fields and click **OK** to complete the addition. The **Type** selector allows you to select what kind of value you have input. The only compulsory field is **Parameter Name**; any of the others may be left blank, though you will usually want to fill in at least the **Value** field as well. Often, the parameter will have a string value, in which case the **Units** field is not very relevant.

### A.10.7 Synthetic Column Window

### Synthetic Column dialogue window

The Synthetic Column Window allows you to define a new "Synthetic" column, that is one whose values are defined using an algebraic expression based on the values of other columns in the same row. The idea is that the value of the cells in a given row in this column will be calculated on demand as a function of the values of cells of other columns in that row. You can think of this as providing functionality like that of a column-oriented spreadsheet. You can activate the dialogue using the **Add Column** (+) or **Replace Column** (↩) buttons in the Columns Window or from the (right-click) popup menu in the Data Window.

The window consists of a number of fields you must fill in to define the new column:

#### Name

The name of the new column. This should preferably be unique (different from all the other column names). It will be easier to use it in algebraic expressions if it is also:

- Different from other columns even apart from upper/lower case distinctions
- In the form of a java identifier (starts with a letter, continues with alphanumerics including underscore)
- Not too long

#### Expression

This is the algebraic expression which defines the values that the cells in the new column of the table will have. The rules for writing algebraic expressions are described in Section 7, and detailed documentation of the functions you can use can be seen in the Available Functions Window, which you can see by clicking the **Show Functions** ( $f(x)$ ) button on the toolbar.

#### Units

The units of the column. If the quantity it represents is dimensionless or you don't know the

units, this can be left blank. It would be a good idea to use a similar format for the units to that used for the existing columns in the table.

### Description

A short textual description of what the values contained by this column are. May be left blank.

### UCD

A Unified Content Descriptor for the column; a UCD is a semantic label attached to the column indicating what kind of quantity it contains by picking one option from a list defined by the CDS. The list of known UCDs is available via a selection box, or you can type a UCD in by hand. You may leave this blank if the you do not wish to assign a UCD to the column. A brief description of the UCD selected is visible below selection box itself.

### Index

Determines the position in the displayed table at which the new column will initially appear.

Of these, the **Expression** is the only one which must be filled in.

Having filled in the form to your satisfaction, hit the **OK** button at the bottom and the new column will be added to the table. If you have made some mistake in filling in the fields, a popup window will give you a message describing the problem. This message may be a bit arcane - try not to panic and see if you can rephrase the expression in a way that the parser might be happier with. If you can't work out the problem, it's time to consult your friendly local Java programmer (failing that, your friendly local C programmer may be able to help) or, by all means, contact the author.


If you wish to add more metadata items you can edit the appropriate cells in the Columns Window. You can edit the expression of an existing synthetic column in the same way.

Once created, a synthetic column is added to the Apparent Table and behaves just like any other; it can be moved, hidden/revealed, used in expressions for other synthetic columns and so on. If the table is saved the new column and its contents will be written to the new output table.

## A.10.8 Sky Coordinates Window

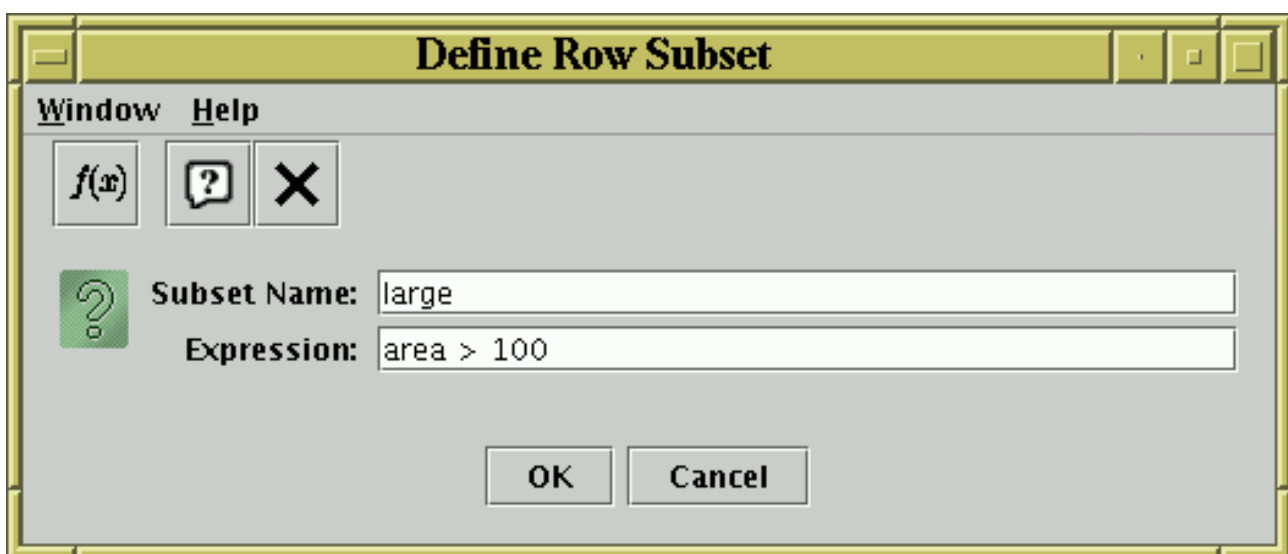
### Sky Coordinates Window

The Sky Coordinates Window allows you to add new columns to a table, representing coordinates in a chosen sky coordinate system. The table must already contain columns which represent sky


coordinates; by describing the systems of the existing and of the new coordinates, you provide enough information to calculate the values in the new columns. You can activate this dialogue using the **New Sky Coordinate Columns** () button in the Columns Window.

The dialogue window has two halves; on the left you give the existing columns which represent sky coordinates, their coordinate system (ICRS, fk5, fk4, galactic, supergalactic or ecliptic) and the units (degrees, radians or sexagesimal) that they are in. Note that the columns available for selection will depend on the units you have selected; for degrees or radians only numeric columns will be selectable, while for sexagesimal (dms/hms) units only string columns will be selectable. On the right you make the coordinate system and units selections as before, but enter the names of the new columns in the text fields. Then just hit the **OK** button, and the new columns will be appended at the right of the table.

### A.10.9 Algebraic Subset Window



**Algebraic Subset dialogue window**

The Algebraic Subset Window allows you to define a new Row Subset which uses an algebraic expression to define which rows are included. The expression must be a boolean one, i.e. its value is either true or false for each row of the table. You can activate this dialogue using the **Add Subset** () button in the Subsets Window.

The window consists of two fields which must be filled in to define the new subset:

#### Subset Name

The name of the new subset. This should preferably be unique (different from existing subset names). It will be easier to use it in other expressions if it is also:

- Different from other columns even apart from upper/lower case distinctions
- In the form of a java identifier (starts with a letter, continues with alphanumerics including underscore)
- Not too long

#### Expression

This is a boolean expression which defines the subset; it is a function of the values of any combination of the columns; only rows for which it evaluates to true will be included in the subset. The values of the other columns in the same row are referenced using their names or \$ID identifiers, and other subsets may be referenced using their names or \_ID identifiers. The

rules for expression syntax are described in Section 7, and detailed documentation of the functions you can use can be seen in the Available Functions Window, which you can see by clicking the **Show Functions** ( $f(x)$ ) button on the toolbar.

Having filled in the form to your satisfaction, hit the **OK** button at the bottom and the new subset will be added to the list that can be seen in the Subsets Window where it behaves like any other. If you have made some mistake in filling in the fields, a popup window will give you a message describing the problem.

#### A.10.10 Column Classification Window

**Column Classification**

Window Help

Query

Classification value:  ▼

Results

Number of categories:  / 35

Subset name prefix:

Count	Value	Subset Name	Add Subset?
15	Sgr	<input type="text" value="Con_Sgr"/>	<input checked="" type="checkbox"/>
11	Vir	<input type="text" value="Con_Vir"/>	<input checked="" type="checkbox"/>
8	Com	<input type="text" value="Con_Com"/>	<input checked="" type="checkbox"/>
7	Oph	<input type="text" value="Con_Oph"/>	<input checked="" type="checkbox"/>
69	other	<input type="text" value="Con_other"/>	<input type="checkbox"/>

Column Classification Window

The Column Classification Window takes a column or other algebraic expression, and generates a number of mutually exclusive Row Subsets based on its contents. For instance if you have a column with different integer values representing different object types, this window will let you add a new subset to the table for each of the distinct values (object types) appearing in the table. Each subset contains rows with a single value of the classification expression you supply. In the above example, a new subset will be created for each of the four most commonly-occurring constellations in the table.

You can activate this dialogue using the **Classify By Column** () button in the Subsets Window.

In the **Query** panel, you specify the expression you want to classify on. You can select a simple column name from the drop-down list, or type in an algebraic expression as described in Section 7. If you want to classify on ranges of values rather than exact equivalence you can use an expression that rounds to an integer, for instance `"toInteger(RMAG/2.0)*2"` would give you subsets corresponding to bins of width 2 in magnitude. When you have entered the expression, you may need to click the **Classify** button to start the classification (or it may happen automatically). The classification will often complete straight away, but for large tables it could take a noticeable amount of time, in which case a progress bar is shown at the bottom of the window. You can stop a long-running classification in progress with the **Stop** button. If you have a very large table with many distinct categories in the given column, the process can take a lot of memory - if the application runs out of memory, a warning will pop up and the classification will not complete.

When the classification has run, the results are displayed in the **Results** panel. Two fields control the way these results are displayed:

#### Number of Categories

Since in general there may be a large number of different values in the column of interest (as many as there are rows), only the few most popular ones are shown. This field controls the maximum number shown, you can adjust it as required. The fixed value after the entry field shows the total number of distinct values discovered in the data; increasing the field's value beyond this value will have no further effect.

#### Subset name prefix

Defines the prefix added to the column values to give the default name for each subset that will be added to the table. The application tries to come up with something sensible based on the classification expression, but doesn't always succeed. If you type in a new prefix here, all the subsets listed below are given new default names accordingly.

The results are displayed in a table, each row corresponding to a Row Subset that can be added to the table based on the classification. The subsets are shown in decreasing order of popularity (the subset containing the most rows is listed first). The final row, labelled "**other**", groups all the rows which are not in any of the other currently selected and displayed subsets. The following columns are shown:

#### Count

Number of rows in the subset.

#### Value

The classification value shared by all rows in the subset.

#### Subset Name

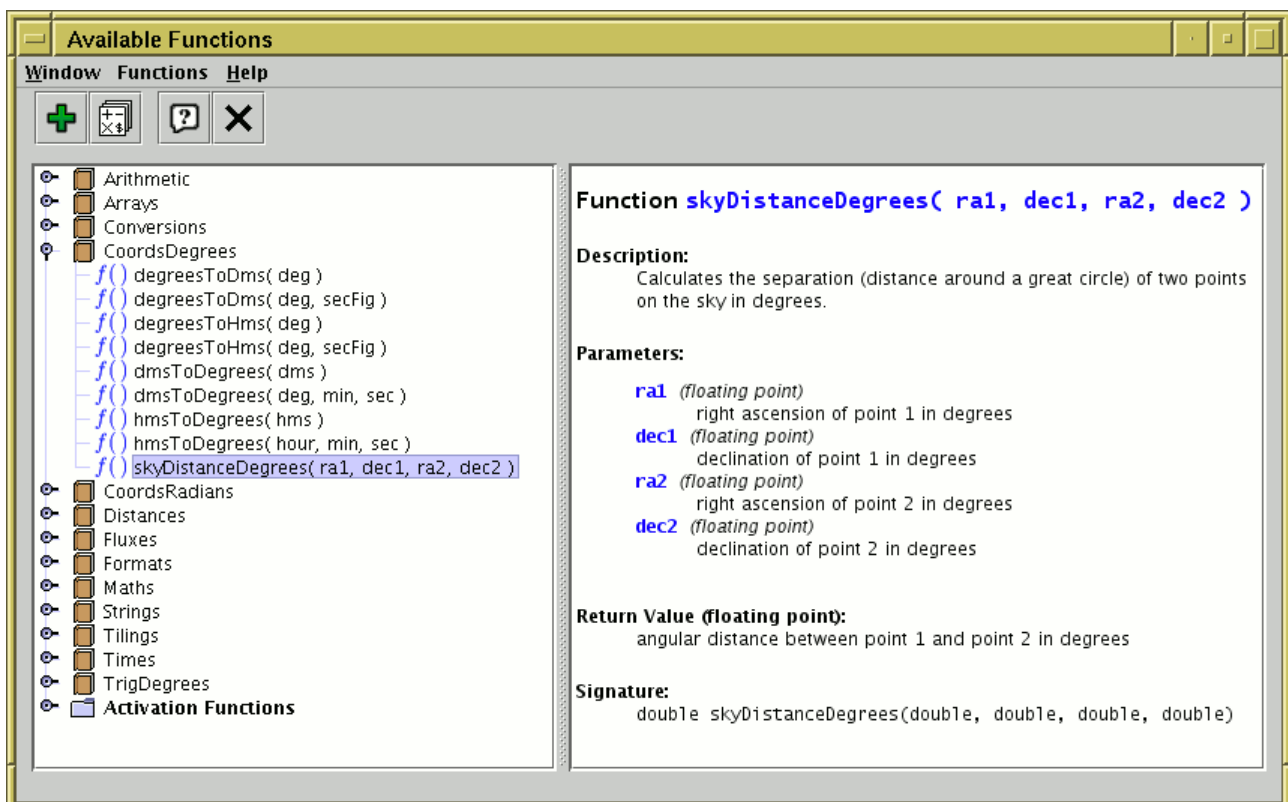
The name of the subset that will be added. The application tries to come up with sensible default names, but doesn't always succeed - you can edit this field as required. If the name matches the name of any subset already present in the table, when the new subset is added it will replace the old one.

### Add Subset?

If this box is checked, the **Add Subsets** button below will add a subset corresponding to this row; if not, it will be ignored. If unchecked, the rows from this subset are considered part of the **other** subset listed at the bottom (you can see the **Count** field in the **other** row adjusting when you include or exclude subsets by checking/unchecking this box).

When you have adjusted subset names and selected the ones you want to add, click the **Add Subsets** button at the bottom, and one subset will be added to the table for each of the items with the **Add Subset?** checkbox ticked. If you don't want to add any of these subsets after all, just hit **Cancel** or close the window.

### A.10.11 Available Functions Window



### Available Functions Window

This window displays all the functions (Java methods) which are available for use when writing algebraic expressions (Section 7). This includes both the built-in expressions and any extended ones you might have added. You can find this window by using the **Show Functions** ( $f(x)$ ) button in the Synthetic Column or Algebraic Subset window toolbars.

On the left hand side of the window is a tree-like representation of the functions you can use. Each item in this tree is one of the following:

#### Folder

A group of classes. There's only one of these, marked "Activation Functions", and it contains functions which are only available for use in Activation Actions (Section 8). When defining a new synthetic columns or algebraic subsets they are not used.

#### Class

A set of functions and/or constants; it doesn't matter what class a function is in when you use it, but since the functions in a class are usually related this makes it easier to find the one




you're looking for in this window.

### **Function**

A function that you can use in an expression.

### **Constant**

A constant value which you can refer to by name in an expression (as long as it doesn't clash with a column name).

Of these, the Folder and Class items have a 'handle' () , which means that they contain other items (classes and functions/constants respectively). By clicking on the handle (or equivalently double-clicking on the name) you can toggle whether the item is open (so you can see its contents) or closed (so you can't). So to see the functions in a class, click on its handle and they will be revealed.

You can click on any of these items and information about it will appear in the right hand panel. In the case of functions this describes the function, its arguments, what it does, and how to use it. The explanations should be fairly self-explanatory; for instance the description in the figure above indicates that you could use the invocation `skyDistanceDegrees(RA1,DEC1,RA2,DEC2)` as the expression for a new table column which gives the angular distance between two sky positions represented by columns with the names RA1, DEC1 and RA2, DEC2. Examples of a number of these functions are given in Section 7.8.

There are two menu/toolbar actions:



### **Syntax Help**

Displays the section of the manual which discusses the general syntax used for algebraic expressions, including what operators are available, what the rules are for referring to columns, etc. The help is displayed in the Help Window (and can from there be passed to a browser if that's more convenient). The displayed material is simply a shortcut to Section 7 in this document.

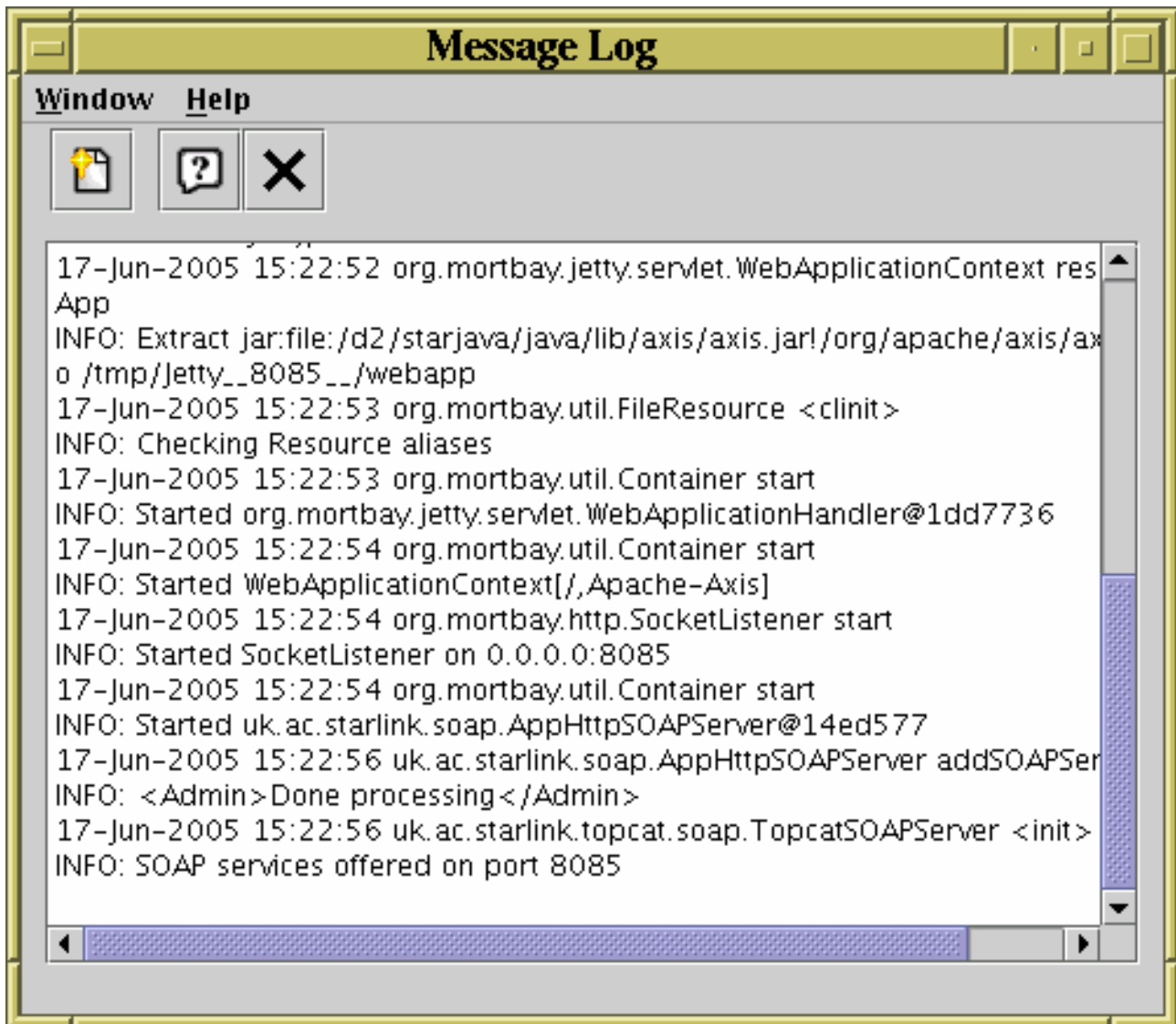


### **Add Class**

Allows you to add the name of a class to those available. You should enter the fully-qualified class name (i.e. including the dot-separated package path). The class that you specify must be on the class path which was current when TOPCAT was started, as explained in Section 10.2.1. Note however it would be more usual to specify these using the system property `jel.classes` or `jel.classes.activation` at startup, as described in Section 7.9.

Classes added either using the **Add Class** action above or the `jel.classes` or `jel.classes.activation` system properties will be visible in the tree in this window, but they may not have proper documentation (clicking on them may not reveal a description in the right hand panel).


## **A.10.12 Log Window**



## Log Window

The log window can be obtained using the **View Log** option on the **File** menu of the Control Window.


This window displays any log messages which the application has generated. Depending on whether the `-verbose` flag has been specified, some or all of these messages may have been written to console as well (if there is a console - this depends on how you have invoked TOPCAT). Under some circumstances, messages way back in the list may not be displayed.

To clear the display of all the existing messages you can use the **Clear Log** button (  ).

The messages displayed here are those written through Java's logging system - in general they are intended for debugging purposes and not for users to read, but if something unexpected is happening, or if you are filing a bug report, it may provide some clues about what's going on. Although it tries not to disturb things too much, TOPCAT's manipulation of the logging infrastructure affects how it is set up, so if you have customised your logging setup using, e.g., the `java.util.logging.config.*` system properties, you may find that it's not behaving exactly as you expected. Sorry.

## B Algebraic Functions

This appendix lists the functions which can be used in algebraic expressions (see Section 7). They are listed in two sections: the first gives the functions available for use anywhere an expression can be used, and the second gives those only for use in defining custom Activation Actions (Section 8).

Note that although all the available functions are listed here with short descriptions, their full explanation, including parameter descriptions and examples, is only available from the Available Functions Window (Appendix A.10.11), obtained using the  toolbar button.

### B.1 General Functions

The following functions can be used anywhere that you can write an algebraic expression in TOPCAT. They will typically be used for defining new synthetic columns or algebraically-defined row subsets. More complete documentation of them is available from within TOPCAT in the Available Functions Window.

#### B.1.1 Tilings

Pixel tiling functions for the celestial sphere.

The  $k$  parameter for the HEALPix functions is the HEALPix order, which can be in the range  $0 \leq k \leq 29$ . This is the logarithm to base 2 of the HEALPix NSIDE parameter. At order  $k$ , there are  $12 \cdot 4^k$  pixels on the sphere.

**htmIndex( level, lon, lat )**

Gives the HTM (Hierarchical Triangular Mesh) pixel index for a given sky position.

**healpixNestIndex( k, lon, lat )**

Gives the pixel index for a given sky position in the HEALPix NEST scheme.

**healpixRingIndex( k, lon, lat )**

Gives the pixel index for a given sky position in the HEALPix RING scheme.

**healpixNestLon( k, index )**

Returns the longitude of the approximate center of the tile with a given index in the HEALPix NEST scheme.

**healpixNestLat( k, index )**

Returns the latitude of the approximate center of the tile with a given index in the HEALPix NEST scheme.

**healpixRingLon( k, index )**

Returns the longitude of the approximate center of the tile with a given index in the HEALPix RING scheme.

**healpixRingLat( k, index )**

Returns the latitude of the approximate center of the tile with a given index in the HEALPix NEST scheme.

**healpixNestToRing( k, nestIndex )**

Converts a healpix ring index from the NEST to the RING scheme at a given order.

**healpixRingToNest( k, ringIndex )**

Converts a healpix ring index from the RING to the NEST scheme at a given order.

**healpixK( pixsize )**

Gives the HEALPix resolution parameter suitable for a given pixel size. This  $k$  value is the

logarithm to base 2 of the Nside parameter.

**healpixResolution( k )**

Gives the approximate resolution in degrees for a given HEALPix resolution parameter *k*. This *k* value is the logarithm to base 2 of the Nside parameter.

**healpixSteradians( k )**

Returns the solid angle in steradians of each HEALPix pixel at a given order.

**healpixSqdeg( k )**

Returns the solid angle in square degrees of each HEALPix pixel at a given order.

**steradiansToSqdeg( sr )**

Converts a solid angle from steradians to square degrees.

The unit sphere is  $4\pi$  steradians =  $360 \times 360 / \pi$  square degrees.

**sqdegToSteradians( sqdeg )**

Converts a solid angle from square degrees to steradians.

The unit sphere is  $4\pi$  steradians =  $360 \times 360 / \pi$  square degrees.

**htmLevel( pixelsize )**

Gives the HTM `level` parameter suitable for a given pixel size.

**htmResolution( level )**

Gives the approximate resolution in degrees for a given HTM depth level.

**SQDEG**

Solid angle in steradians corresponding to 1 square degree.

## B.1.2 Arithmetic

Standard arithmetic functions including things like rounding, sign manipulation, and maximum/minimum functions.

**roundUp( x )**

Rounds a value up to an integer value. Formally, returns the smallest (closest to negative infinity) integer value that is not less than the argument.

**roundDown( x )**

Rounds a value down to an integer value. Formally, returns the largest (closest to positive infinity) integer value that is not greater than the argument.

**round( x )**

Rounds a value to the nearest integer. Formally, returns the integer that is closest in value to the argument. If two integers are equally close, the result is the even one.

**roundDecimal( x, dp )**

Rounds a value to a given number of decimal places. The result is a `float` (32-bit floating point value), so this is only suitable for relatively low-precision values. It's intended for truncating the number of apparent significant figures represented by a value which you know has been obtained by combining other values of limited precision. For more control, see the functions in the `Formats` class.

**abs( x )**

Returns the absolute value of an integer value. If the argument is not negative, the argument is returned. If the argument is negative, the negation of the argument is returned.

**abs( x )**

Returns the absolute value of a floating point value. If the argument is not negative, the argument is returned. If the argument is negative, the negation of the argument is returned.

**max( a, b )**

Returns the greater of two integer values. If the arguments have the same value, the result is that same value.

Multiple-argument maximum functions are also provided in the `Arrays` and `Lists` packages.

**maxNaN( a, b )**

Returns the greater of two floating point values. If the arguments have the same value, the result is that same value. If either value is blank, then the result is blank.

**maxReal( a, b )**

Returns the greater of two floating point values, ignoring blanks. If the arguments have the same value, the result is that same value. If one argument is blank, the result is the other one. If both arguments are blank, the result is blank.

Multiple-argument maximum functions are also provided in the `Arrays` and `Lists` packages.

**min( a, b )**

Returns the smaller of two integer values. If the arguments have the same value, the result is that same value.

Multiple-argument minimum functions are also provided in the `Arrays` and `Lists` packages.

**minNaN( a, b )**

Returns the smaller of two floating point values. If the arguments have the same value, the result is that same value. If either value is blank, then the result is blank.

**minReal( a, b )**

Returns the smaller of two floating point values, ignoring blanks. If the arguments have the same value, the result is that same value. If one argument is blank, the result is the other one. If both arguments are blank, the result is blank.

Multiple-argument minimum functions are also provided in the `Arrays` and `Lists` packages.

### B.1.3 Conversions

Functions for converting between strings and numeric values.

**toString( fpVal )**

Turns a numeric value into a string.

**toString( intVal )**

Turns an integer numeric value into a string.

**toString( charVal )**

Turns a single character value into a string.

**toString( byteVal )**

Turns a byte value into a string.

**toString( booleanVal )**

Turns a boolean value into a string.

**toString( objVal )**

Turns any object value into a string. As applied to existing string values this isn't really useful, but it means that you can apply `toString` to any object value without knowing its type and get a useful return from it.

**parseByte( str )**

Attempts to interpret a string as a byte (8-bit signed integer) value. If the input string can't be interpreted in this way, a blank value will result.

**parseShort( str )**

Attempts to interpret a string as a short (16-bit signed integer) value. If the input string can't be interpreted in this way, a blank value will result.

**parseInt( str )**

Attempts to interpret a string as an int (32-bit signed integer) value. If the input string can't be interpreted in this way, a blank value will result.

**parseLong( str )**

Attempts to interpret a string as a long (64-bit signed integer) value. If the input string can't be interpreted in this way, a blank value will result.

**parseFloat( str )**

Attempts to interpret a string as a float (32-bit floating point) value. If the input string can't be interpreted in this way, a blank value will result.

**parseDouble( str )**

Attempts to interpret a string as a double (64-bit signed integer) value. If the input string can't be interpreted in this way, a blank value will result.

**toByte( value )**

Attempts to convert the numeric argument to a byte (8-bit signed integer) result. If it is out of range, a blank value will result.

**toShort( value )**

Attempts to convert the numeric argument to a short (16-bit signed integer) result. If it is out of range, a blank value will result.

**toInteger( value )**

Attempts to convert the numeric argument to an int (32-bit signed integer) result. If it is out of range, a blank value will result.

**toLong( value )**

Attempts to convert the numeric argument to a long (64-bit signed integer) result. If it is out of range, a blank value will result.

**toFloat( value )**

Attempts to convert the numeric argument to a float (32-bit floating point) result. If it is out of range, a blank value will result.

**toDouble( value )**

Converts the numeric argument to a double (64-bit signed integer) result.

**toHex( value )**

Converts the integer argument to hexadecimal form.

**fromHex( hexVal )**

Converts a string representing a hexadecimal number to its integer value.

**B.1.4 Distances**

Functions for converting between different measures of cosmological distance.

The following parameters are used:

- **z**: redshift
- **H0**: Hubble constant in km/sec/Mpc (example value ~70)
- **omegaM**: Density ratio of the universe (example value 0.3)
- **omegaLambda**: Normalised cosmological constant (example value 0.7)

For a flat universe,  $\text{omegaM} + \text{omegaLambda} = 1$

The terms and formulae used here are taken from the paper by D.W.Hogg, *Distance measures in cosmology*, astro-ph/9905116 (<http://arxiv.org/abs/astro-ph/9905116>) v4 (2000).

### **MpcToM( distMpc )**

Converts from MegaParsecs to metres.

### **mToMpc( distM )**

Converts from metres to MegaParsecs.

### **zToDist( z )**

Quick and dirty function for converting from redshift to distance.

**Warning:** this makes some reasonable assumptions about the cosmology and returns the luminosity distance. It is only intended for approximate use. If you care about the details, use one of the more specific functions here.

### **zToAge( z )**

Quick and dirty function for converting from redshift to time.

**Warning:** this makes some reasonable assumptions about the cosmology. It is only intended for approximate use. If you care about the details use one of the more specific functions here.

### **comovingDistanceL( z, H0, omegaM, omegaLambda )**

Line-of-sight comoving distance.

### **comovingDistanceT( z, H0, omegaM, omegaLambda )**

Transverse comoving distance.

### **angularDiameterDistance( z, H0, omegaM, omegaLambda )**

Angular diameter distance.

### **luminosityDistance( z, H0, omegaM, omegaLambda )**

Luminosity distance.

### **lookbackTime( z, H0, omegaM, omegaLambda )**

Lookback time. This returns the difference between the age of the universe at time of observation (now) and the age of the universe at the time when photons of redshift  $z$  were emitted.

### **comovingVolume( z, H0, omegaM, omegaLambda )**

Comoving volume. This returns the all-sky total comoving volume out to a given redshift  $z$ .

### **SPEED\_OF\_LIGHT**

Speed of light in m/s.

### **METRE\_PER\_PARSEC**

Number of metres in a parsec.

### **SEC\_PER\_YEAR**

Number of seconds in a year.

## **B.1.5 KCorrections**

Functions for calculating K-corrections.

### **kCorr( filter, redshift, colorType, colorValue )**

Calculates K-corrections. This allows you to determine K-corrections for a galaxy, given its redshift and a colour. Filters for GALEX, SDSS, UKIDSS, Johnson, Cousins and 2MASS are covered.

To define the calculation you must choose both a filter, specified as a `KCF_*` constant, and a

colour (filter pair) specified as a `KCC_*` constant. For each available filter, only certain colours are available, as described in the documentation of the relevant `KCF_*` constant.

The algorithm used is described at <http://kcor.sai.msu.ru/> (<http://kcor.sai.msu.ru/>). This is based on the paper "*Analytical Approximations of K-corrections in Optical and Near-Infrared Bands*" by I.Chilingarian, A.-L.Melchior and I.Zolotukhin (2010MNRAS.405.1409C) (<http://adsabs.harvard.edu/abs/2010MNRAS.405.1409C>)), but extended to include GALEX UV bands and with redshift coverage up to 0.5 as described in "*Universal UV-optical Colour-Colour-Magnitude Relation of Galaxies*" by I.Chilingarian and I.Zolotukhin (2012MNRAS.419.1727C) (<http://adsabs.harvard.edu/abs/2012MNRAS.419.1727C>)).

#### **KCF\_FUV**

GALEX FUV filter (AB). Use with `KCC_FUVNUV` or `KCC_FUVu`.

#### **KCF\_NUV**

GALEX NUV filter (AB). Use with `KCC_NUVg` or `KCC_NUVr`.

#### **KCF\_u**

SDSS u filter (AB). Use with `KCC_ur`, `KCC_ui` or `KCC_uz`.

#### **KCF\_g**

SDSS g filter (AB). Use with `KCC_gr`, `KCC_gi` or `KCC_gz`.

#### **KCF\_r**

SDSS r filter (AB). Use with `KCC_gr` or `KCC_ur`.

#### **KCF\_i**

SDSS i filter (AB). Use with `KCC_gi` or `KCC_ui`.

#### **KCF\_z**

SDSS z filter (AB). Use with `KCC_rz`, `KCC_gz` or `KCC_uz`.

#### **KCF\_Y**

UKIDSS Y filter (AB). Use with `KCC_YH` or `KCC_YK`.

#### **KCF\_J**

UKIDSS J filter (AB). Use with `KCC_JK` or `KCC_JH`.

#### **KCF\_H**

UKIDSS H filter (AB). Use with `KCC_HK` or `KCC_JH`.

#### **KCF\_K**

UKIDSS K filter (AB). Use with `KCC_JK` or `KCC_HK`.

#### **KCF\_U**

Johnson U filter (Vega). Use with `KCC_URc`.

#### **KCF\_B**

Johnson B filter (Vega). Use with `KCC_BRc` or `KCC_BIc`.

#### **KCF\_V**

Johnson V filter (Vega). Use with `KCC_VIc` or `KCC_VRc`.

#### **KCF\_Rc**

Cousins Rc filter (Vega). Use with `KCC_BRc` or `KCC_VRc`.

#### **KCF\_Ic**

Cousins Ic filter (Vega). Use with `KCC_VIc`.

#### **KCF\_J2**

2MASS J filter (Vega). Use with `KCC_J2Ks2` or `KCC_J2H2`.

#### **KCF\_H2**

2MASS H filter (Vega). Use with `KCC_H2Ks2` or `KCC_J2H2`.

#### **KCF\_Ks2**



2MASS Ks filter (Vega). Use with KCC\_J2Ks2 or KCC\_H2Ks2.

**KCC\_BIc**

Johnson B - Cousins Ic colour.

**KCC\_BRc**

Johnson B - Cousins Rc colour.

**KCC\_FUVNUV**

GALEX FUV - NUV colour.

**KCC\_FUVu**

GALEX FUV - SDSS u colour.

**KCC\_gi**

SDSS g - i colour.

**KCC\_gr**

SDSS g - r colour.

**KCC\_gz**

SDSS g - z colour.

**KCC\_H2Ks2**

2MASS H - Ks colour.

**KCC\_HK**

UKIDSS H - K colour.

**KCC\_J2H2**

2MASS J - H colour.

**KCC\_J2Ks2**

2MASS J - Ks colour.

**KCC\_JH**

UKIDSS J - H colour.

**KCC\_JK**

UKIDSS J - K colour.

**KCC\_NUVg**

GALEX NUV - SDSS g colour.

**KCC\_NUVr**

GALEX NUV - SDSS r colour.

**KCC\_rz**

SDSS r - SDSS z colour.

**KCC\_ui**

SDSS u - SDSS i colour.

**KCC\_URc**

Johnson U - Cousins Rc colour.

**KCC\_ur**

SDSS u - r colour.

**KCC\_uz**

SDSS u - z colour.

**KCC\_VIc**

Johnson V - Cousins Ic colour.

**KCC\_VRc**

Johnson V - Cousins Rc colour.

**KCC\_YH**

UKIDSS Y - H colour.

**KCC\_YK**

UKIDSS Y - K colour.

**B.1.6 Times**

Functions for conversion of time values between various forms. The forms used are

**Modified Julian Date (MJD)**

A continuous measure in days since midnight at the start of 17 November 1858. Based on UTC.

**ISO 8601**

A string representation of the form `yyyy-mm-ddThh:mm:ss.s`, where the `T` is a literal character (a space character may be used instead). Based on UTC.

**Julian Epoch**

A continuous measure based on a Julian year of exactly 365.25 days. For approximate purposes this resembles the fractional number of years AD represented by the date. Sometimes (but not here) represented by prefixing a 'J'; J2000.0 is defined as 2000 January 1.5 in the TT timescale.

**Besselian Epoch**

A continuous measure based on a tropical year of about 365.2422 days. For approximate purposes this resembles the fractional number of years AD represented by the date. Sometimes (but not here) represented by prefixing a 'B'.

**Decimal Year**

Fractional number of years AD represented by the date. 2000.0, or equivalently 1999.99 recurring, is midnight at the start of the first of January 2000. Because of leap years, the size of a unit depends on what year it is in.

Therefore midday on the 25th of October 2004 is `2004-10-25T12:00:00` in ISO 8601 format, 53303.5 as an MJD value, 2004.81588 as a Julian Epoch and 2004.81726 as a Besselian Epoch.

Currently this implementation cannot be relied upon to better than a millisecond.

**isoToMjd( isoDate )**

Converts an ISO8601 date string to Modified Julian Date. The basic format of the `isoDate` argument is `yyyy-mm-ddThh:mm:ss.s`, though some deviations from this form are permitted:

- The 't' which separates date from time can be replaced by a space
- The seconds, minutes and/or hours can be omitted
- The decimal part of the seconds can be any length, and is optional
- A 'z' (which indicates UTC) may be appended to the time

Some legal examples are therefore: `"1994-12-21T14:18:23.2"`, `"1968-01-14"`, and `"2112-05-25 16:45Z"`.

**dateToMjd( year, month, day, hour, min, sec )**

Converts a calendar date and time to Modified Julian Date.

**dateToMjd( year, month, day )**

Converts a calendar date to Modified Julian Date.

**decYearToMjd( decYear )**

Converts a Decimal Year to a Modified Julian Date.

**mjdToIso( mjd )**

Converts a Modified Julian Date value to an ISO 8601-format date-time string. The output format is `yyyy-mm-ddThh:mm:ss`.

**mjdToDate( mjd )**

Converts a Modified Julian Date value to an ISO 8601-format date string. The output format is `yyyy-mm-dd`.

**mjdToTime( mjd )**

Converts a Modified Julian Date value to an ISO 8601-format time-only string. The output format is `hh:mm:ss`.

**mjdToDecYear( mjd )**

Converts a Modified Julian Date to Decimal Year.

**formatMjd( mjd, format )**

Converts a Modified Julian Date value to a date using a customisable date format. The format is as defined by the `java.text.SimpleDateFormat` (<http://docs.oracle.com/javase/6/docs/api/java/text/SimpleDateFormat.html>) class. The default output corresponds to the string `"yyyy-MM-dd 'T' HH:mm:ss"`

**mjdToJulian( mjd )**

Converts a Modified Julian Date to Julian Epoch. For approximate purposes, the result of this routine consists of an integral part which gives the year AD and a fractional part which represents the distance through that year, so that for instance 2000.5 is approximately 1 July 2000.

**julianToMjd( julianEpoch )**

Converts a Julian Epoch to Modified Julian Date. For approximate purposes, the argument of this routine consists of an integral part which gives the year AD and a fractional part which represents the distance through that year, so that for instance 2000.5 is approximately 1 July 2000.

**mjdToBesselian( mjd )**

Converts Modified Julian Date to Besselian Epoch. For approximate purposes, the result of this routine consists of an integral part which gives the year AD and a fractional part which represents the distance through that year, so that for instance 1950.5 is approximately 1 July 1950.

**besselianToMjd( besselianEpoch )**

Converts Besselian Epoch to Modified Julian Date. For approximate purposes, the argument of this routine consists of an integral part which gives the year AD and a fractional part which represents the distance through that year, so that for instance 1950.5 is approximately 1 July 1950.

**unixMillisToMjd( unixMillis )**

Converts from milliseconds since the Unix epoch (1970-01-01T00:00:00) to a modified Julian date value

**mjdToUnixMillis( mjd )**

Converts from modified Julian date to milliseconds since the Unix epoch (1970-01-01T00:00:00).

**B.1.7 TrigDegrees**

Standard trigonometric functions with angles in degrees.

**sinDeg( theta )**

Sine of an angle.

**cosDeg( theta )**

Cosine of an angle.

**tanDeg( theta )**

Tangent of an angle.

**asinDeg( x )**

Arc sine. The result is in the range of -90 through 90.

**acosDeg( x )**

Arc cosine. The result is in the range of 0.0 through 180.

**atanDeg( x )**

Arc tangent. The result is in the range of -90 through 90.

**atan2Deg( y, x )**

Converts rectangular coordinates (x,y) to polar (r,theta). This method computes the phase theta by computing an arc tangent of y/x in the range of -180 to 180.

**B.1.8 Maths**

Standard mathematical and trigonometric functions. Trigonometric functions work with angles in radians.

**sin( theta )**

Sine of an angle.

**cos( theta )**

Cosine of an angle.

**tan( theta )**

Tangent of an angle.

**asin( x )**Arc sine of an angle. The result is in the range of  $-pi/2$  through  $pi/2$ .**acos( x )**Arc cosine of an angle. The result is in the range of 0.0 through  $pi$ .**atan( x )**Arc tangent of an angle. The result is in the range of  $-pi/2$  through  $pi/2$ .**exp( x )**Euler's number  $e$  raised to a power.**log10( x )**

Logarithm to base 10.

**ln( x )**

Natural logarithm.

**sqrt( x )**

Square root. The result is correctly rounded and positive.

**hypot( x, y )**

Returns the square root of the sum of squares of its two arguments. Doing it like this may avoid intermediate overflow or underflow.

**atan2( y, x )**Converts rectangular coordinates (x,y) to polar (r,theta). This method computes the phase theta by computing an arc tangent of y/x in the range of  $-pi$  to  $pi$ .**pow( a, b )**

Exponentiation. The result is the value of the first argument raised to the power of the second argument.

**sinh( x )**

Hyperbolic sine.

**cosh( x )**

Hyperbolic cosine.

**tanh( x )**

Hyperbolic tangent.

**asinh( x )**

Inverse hyperbolic sine.

**acosh( x )**

Inverse hyperbolic cosine.

**atanh( x )**

Inverse hyperbolic tangent.

**E**

Euler's number  $e$ , the base of natural logarithms.

**PI**

$Pi$ , the ratio of the circumference of a circle to its diameter.

**Infinity**

Positive infinite floating point value.

**NaN**

Not-a-Number floating point value. Use with care; arithmetic and logical operations behave in strange ways near NaN (for instance,  $\text{NaN}! = \text{NaN}$ ). For most purposes this is equivalent to the blank value.

**RANDOM**

Evaluates to a random number in the range  $0 \leq x < 1$ . This is different for each cell of the table. The quality of the randomness may not be particularly good.

## B.1.9 Arrays

Functions which operate on array-valued cells. The array parameters of these functions can only be used on values which are already arrays (usually, numeric arrays). In most cases that means on values in table columns which are declared as array-valued. FITS and VOTable tables can have columns which contain array values, but other formats such as CSV cannot.

If you want to calculate aggregating functions like sum, min, max etc on multiple values which are not part of an array, it's easier to use the functions from the `Lists` class.

Note that none of these functions will calculate statistical functions over a whole column of a table.

The functions fall into a number of categories:

- Aggregating operations, which map an array value to a scalar, including `size`, `count`, `countTrue`, `maximum`, `minimum`, `sum`, `mean`, `median`, `quantile`, `stdev`, `variance`, `join`.
- Operations on one or more arrays which produce array results, including `add`, `subtract`, `multiply`, `divide`, `reciprocal`, `condition`.
- The function `array`, which lets you assemble an array value from a list of scalar numbers. This can be used with the aggregating functions here, but it's generally easier to use the corresponding functions from the `Lists` class.

**sum( array )**

Returns the sum of all the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**mean( array )**

Returns the mean of all the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**variance( array )**

Returns the population variance of all the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**stdev( array )**

Returns the population standard deviation of all the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**minimum( array )**

Returns the smallest of the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**maximum( array )**

Returns the largest of the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**median( array )**

Returns the median of the non-blank elements in the array. If `array` is not a numeric array, `null` is returned.

**quantile( array, quant )**

Returns a quantile value of the non-blank elements in the array. Which quantile is determined by the `quant` value; values of 0, 0.5 and 1 give the minimum, median and maximum respectively. A value of 0.99 would give the 99th percentile.

**size( array )**

Returns the number of elements in the array. If `array` is not an array, zero is returned.

**count( array )**

Returns the number of non-blank elements in the array. If `array` is not an array, zero is returned.

**countTrue( array )**

Returns the number of true elements in an array of boolean values.

**join( array, joiner )**

Returns a string composed of concatenating all the elements of an array, separated by a joiner string. If `array` is not an array, `null` is returned.

**add( array1, array2 )**

Returns the result of adding two numeric arrays element by element. Both arrays must be numeric, and the arrays must have the same length. If either of those conditions is not true, `null` is returned. The types of the arrays do not need to be the same, so for example it is permitted to add an integer array to a floating point array.

**add( array, constant )**

Returns the result of adding a constant value to every element of a numeric array. If the supplied `array` argument is not a numeric array, `null` is returned.

**subtract( array1, array2 )**

Returns the result of subtracting one numeric array from the other element by element. Both arrays must be numeric, and the arrays must have the same length. If either of those conditions is not true, `null` is returned. The types of the arrays do not need to be the same, so for example it is permitted to subtract an integer array from a floating point array.

**multiply( array1, array2 )**

Returns the result of multiplying two numeric arrays element by element. Both arrays must be numeric, and the arrays must have the same length. If either of those conditions is not true, `null` is returned. The types of the arrays do not need to be the same, so for example it is permitted to multiply an integer array by a floating point array.

**multiply( array, constant )**

Returns the result of multiplying every element of a numeric array by a constant value. If the supplied `array` argument is not a numeric array, `null` is returned.

**divide( array1, array2 )**

Returns the result of dividing two numeric arrays element by element. Both arrays must be numeric, and the arrays must have the same length. If either of those conditions is not true, `null` is returned. The types of the arrays do not need to be the same, so for example it is permitted to divide an integer array by a floating point array.

**reciprocal( array )**

Returns the result of taking the reciprocal of every element of a numeric array. If the supplied `array` argument is not a numeric array, `null` is returned.

**condition( flagArray, trueValue, falseValue )**

Maps a boolean array to a numeric array by using supplied numeric values to represent true and false values from the input array.

This has the same effect as applying the expression `outArray[i] = flagArray[i] ? trueValue : falseValue`.

**array( values, ... )**

Returns a floating point numeric array built from the given arguments.

**intArray( values, ... )**

Returns an integer numeric array built from the given arguments.

**stringArray( values, ... )**

Returns a String array built from the given arguments.

**B.1.10 Fluxes**

Functions for conversion between flux and magnitude values. Functions are provided for conversion between flux in Janskys and AB magnitudes.

Some constants for approximate conversions between different magnitude scales are also provided:

- Constants `JOHNSON_AB_*`, for Johnson  $\leftrightarrow$  AB magnitude conversions, from Frei and Gunn, *Astronomical Journal* 108, 1476 (1994), Table 2 (1994AJ....108.1476F (<http://adsabs.harvard.edu/abs/1994AJ....108.1476F>)).
- Constants `VEGA_AB_*`, for Vega  $\leftrightarrow$  AB magnitude conversions, from Blanton et al., *Astronomical Journal* 129, 2562 (2005), Eqs. (5) (2005AJ....129.2562B (<http://adsabs.harvard.edu/abs/2005AJ....129.2562B>)).

**abToJansky( magAB )**

Converts AB magnitude to flux in Jansky.

$$F/Jy = 10^{(23 - (AB + 48.6)) / 2.5}$$

**janskyToAb( fluxJansky )**

Converts flux in Jansky to AB magnitude.

$$AB = 2.5 * (23 - \log_{10}(F/Jy)) - 48.6$$

**luminosityToFlux( lumin, dist )**

Converts luminosity to flux given a luminosity distance.

$$F = \text{lumin} / (4 \times \text{Pi} \times \text{dist}^2)$$

**fluxToLuminosity( flux, dist )**

Converts flux to luminosity given a luminosity distance.

$$\text{lumin} = (4 \times \text{Pi} \times \text{dist}^2) \times F$$

**JOHNSON\_AB\_V**

Approximate offset between Johnson and AB magnitudes in V band.

$$V_J \sim V_{AB} + \text{JOHNSON\_AB\_V.}$$

**JOHNSON\_AB\_B**

Approximate offset between Johnson and AB magnitudes in B band.

$$B_J \sim B_{AB} + \text{JOHNSON\_AB\_B.}$$

**JOHNSON\_AB\_Bj**

Approximate offset between Johnson and AB magnitudes in Bj band.

$$B_{jJ} \sim B_{jAB} + \text{JOHNSON\_AB\_Bj.}$$

**JOHNSON\_AB\_R**

Approximate offset between Johnson and AB magnitudes in R band.

$$R_J \sim R_{AB} + \text{JOHNSON\_AB\_R.}$$

**JOHNSON\_AB\_I**

Approximate offset between Johnson and AB magnitudes in I band.  $I_J \sim I_{AB} + \text{JOHNSON\_AB\_I.}$

**JOHNSON\_AB\_g**

Approximate offset between Johnson and AB magnitudes in g band.  $g_J \sim g_{AB} + \text{JOHNSON\_AB\_g.}$

**JOHNSON\_AB\_r**

Approximate offset between Johnson and AB magnitudes in r band.  $r_J \sim r_{AB} + \text{JOHNSON\_AB\_r.}$

**JOHNSON\_AB\_i**

Approximate offset between Johnson and AB magnitudes in i band.  $i_J \sim i_{AB} + \text{JOHNSON\_AB\_i.}$

**JOHNSON\_AB\_Rc**

Approximate offset between Johnson and AB magnitudes in Rc band.

$$R_{cJ} \sim R_{cAB} + \text{JOHNSON\_AB\_Rc.}$$

**JOHNSON\_AB\_Ic**

Approximate offset between Johnson and AB magnitudes in Ic band.

$$I_{cJ} \sim I_{cAB} + \text{JOHNSON\_AB\_Ic.}$$

**JOHNSON\_AB\_uPrime**

Offset between Johnson and AB magnitudes in u' band (zero).

$$u'_J = u'_{AB} + \text{JOHNSON\_AB\_uPrime} = u'_{AB}.$$

**JOHNSON\_AB\_gPrime**

Offset between Johnson and AB magnitudes in g' band (zero).

$$g'_J = g'_{AB} + \text{JOHNSON\_AB\_gPrime} = g'_{AB}.$$

**JOHNSON\_AB\_rPrime**

Offset between Johnson and AB magnitudes in r' band (zero).

$$r'_J = r'_{AB} + \text{JOHNSON\_AB\_rPrime} = r'_{AB}.$$

**JOHNSON\_AB\_iPrime**

Offset between Johnson and AB magnitudes in i' band (zero).

$$i'_J = i'_{AB} + \text{JOHNSON\_AB\_iPrime} = i'_{AB}.$$

**JOHNSON\_AB\_zPrime**

Offset between Johnson and AB magnitudes in z' band (zero).

$$z'_J = z'_{AB} + \text{JOHNSON\_AB\_zPrime} = z'_{AB}.$$



**VEGA\_AB\_J**

Approximate offset between Vega (as in 2MASS) and AB magnitudes in J band.

$$J_{\text{Vega}} \sim J_{\text{AB}} + \text{VEGA\_AB\_J}.$$

**VEGA\_AB\_H**

Approximate offset between Vega (as in 2MASS) and AB magnitudes in H band.

$$H_{\text{Vega}} \sim H_{\text{AB}} + \text{VEGA\_AB\_H}.$$

**VEGA\_AB\_K**

Approximate offset between Vega (as in 2MASS) and AB magnitudes in K band.

$$K_{\text{Vega}} \sim K_{\text{AB}} + \text{VEGA\_AB\_K}.$$

**B.1.11 Strings**

String manipulation and query functions.

**concat( strings, ... )**

Concatenates multiple values into a string. In some cases the same effect can be achieved by writing `s1+s2+...`, but this method makes sure that values are converted to strings, with the blank value invisible.

**join( separator, words, ... )**

Joins multiple values into a string, with a given separator between each pair.

**equals( s1, s2 )**

Determines whether two strings are equal. Note you should use this function instead of `s1==s2`, which can (for technical reasons) return false even if the strings are the same.

**equalsIgnoreCase( s1, s2 )**

Determines whether two strings are equal apart from possible upper/lower case distinctions.

**startsWith( whole, start )**

Determines whether a string starts with a certain substring.

**endsWith( whole, end )**

Determines whether a string ends with a certain substring.

**contains( whole, sub )**

Determines whether a string contains a given substring.

**length( str )**

Returns the length of a string in characters.

**split( words )**

Splits a string into an array of space-separated words. One or more spaces separates each word from the next. Leading and trailing spaces are ignored.

The result is an array of strings, and if you want to use the individual elements you need to use square-bracket indexing, with `[0]` representing the first object

**split( words, regex )**

Splits a string into an array of words separated by a given regular expression.

The result is an array of strings, and if you want to use the individual elements you need to use square-bracket indexing, with `[0]` representing the first object

**matches( str, regex )**

Tests whether a string matches a given regular expression.

**matchGroup( str, regex )**

Returns the first grouped expression matched in a string defined by a regular expression. A grouped expression is one enclosed in parentheses.

**replaceFirst( str, regex, replacement )**

Replaces the first occurrence of a regular expression in a string with a different substring value.

**replaceAll( str, regex, replacement )**

Replaces all occurrences of a regular expression in a string with a different substring value.

**substring( str, startIndex )**

Returns the last part of a given string. The substring begins with the character at the specified index and extends to the end of this string.

**substring( str, startIndex, endIndex )**

Returns a substring of a given string. The substring begins with the character at `startIndex` and continues to the character at index `endIndex-1`. Thus the length of the substring is `endIndex-startIndex`.

**toUpperCase( str )**

Returns an uppercased version of a string.

**toLowerCase( str )**

Returns a lowercased version of a string.

**trim( str )**

Trims whitespace from both ends of a string.

**padWithZeros( value, ndigit )**

Takes an integer argument and returns a string representing the same numeric value but padded with leading zeros to a specified length.

**B.1.12 Formats**

Functions for formatting numeric values.

**formatDecimal( value, dp )**

Turns a floating point value into a string with a given number of decimal places using standard settings.

**formatDecimalLocal( value, dp )**

Turns a floating point value into a string using current locale settings. For instance if language is set to French, decimal points will be represented as a comma "," instead of a full stop ".". Otherwise behaves the same as the corresponding `formatDecimal` function.

**formatDecimal( value, format )**

Turns a floating point value into a formatted string using standard settings. The `format` string is as defined by Java's `java.text.DecimalFormat` (<http://docs.oracle.com/javase/6/docs/api/java/text/DecimalFormat.html>) class.

**formatDecimalLocal( value, format )**

Turns a floating point value into a formatted string using current locale settings. For instance if language is set to French, decimal points will be represented as a comma "," instead of a full stop ".". Otherwise behaves the same as the corresponding `formatDecimal` function.

**B.1.13 CoordsRadians**

Functions for angle transformations and manipulations, based on radians rather than degrees. In particular, methods for translating between radians and HH:MM:SS.S or DDD:MM:SS.S type sexagesimal representations are provided.

**radiansToDms( rad )**

Converts an angle in radians to a formatted degrees:minutes:seconds string. No fractional part of the seconds field is given.

**radiansToDms( rad, secFig )**

Converts an angle in radians to a formatted degrees:minutes:seconds string with a given number of decimal places in the seconds field.

**radiansToHms( rad )**

Converts an angle in radians to a formatted hours:minutes:seconds string. No fractional part of the seconds field is given.

**radiansToHms( rad, secFig )**

Converts an angle in radians to a formatted hours:minutes:seconds string with a given number of decimal places in the seconds field.

**dmsToRadians( dms )**

Converts a formatted degrees:minutes:seconds string to an angle in radians. Delimiters may be colon, space, characters `dm[s]`, or some others. Additional spaces and leading +/- are permitted. The :seconds part is optional.

**hmsToRadians( hms )**

Converts a formatted hours:minutes:seconds string to an angle in radians. Delimiters may be colon, space, characters `hm[s]`, or some others. Additional spaces and leading +/- are permitted. The :seconds part is optional.

**dmsToRadians( deg, min, sec )**

Converts degrees, minutes, seconds to an angle in radians.

In conversions of this type, one has to be careful to get the sign right in converting angles which are between 0 and -1 degrees. This routine uses the sign bit of the `deg` argument, taking care to distinguish between +0 and -0 (their internal representations are different for floating point values). It is illegal for the `min` or `sec` arguments to be negative.

**hmsToRadians( hour, min, sec )**

Converts hours, minutes, seconds to an angle in radians.

In conversions of this type, one has to be careful to get the sign right in converting angles which are between 0 and -1 hours. This routine uses the sign bit of the `hour` argument, taking care to distinguish between +0 and -0 (their internal representations are different for floating point values).

**skyDistanceRadians( ra1, dec1, ra2, dec2 )**

Calculates the separation (distance around a great circle) of two points on the sky in radians.

**hoursToRadians( hours )**

Converts hours to radians.

**degreesToRadians( deg )**

Converts degrees to radians.

**radiansToDegrees( rad )**

Converts radians to degrees.

**raFK4toFK5radians( raFK4, decFK4 )**

Converts a B1950.0 FK4 position to J2000.0 FK5 at an epoch of B1950.0 yielding Right Ascension. This assumes zero proper motion in the FK5 frame.

**decFK4toFK5radians( raFK4, decFK4 )**

Converts a B1950.0 FK4 position to J2000.0 FK5 at an epoch of B1950.0 yielding Declination. This assumes zero proper motion in the FK5 frame.

**raFK5toFK4radians( raFK5, decFK5 )**

Converts a J2000.0 FK5 position to B1950.0 FK4 at an epoch of B1950.0 yielding Declination. This assumes zero proper motion, parallax and radial velocity in the FK5 frame.

**decFK5toFK4radians( raFK5, decFK5 )**

Converts a J2000.0 FK5 position to B1950.0 FK4 at an epoch of B1950.0 yielding Declination. This assumes zero proper motion, parallax and radial velocity in the FK5 frame.

**raFK4toFK5Radians( raFK4, decFK4, bepoch )**

Converts a B1950.0 FK4 position to J2000.0 FK5 yielding Right Ascension. This assumes zero proper motion in the FK5 frame. The `bepoch` parameter is the epoch at which the position in the FK4 frame was determined.

**decFK4toFK5Radians( raFK4, decFK4, bepoch )**

Converts a B1950.0 FK4 position to J2000.0 FK5 yielding Declination. This assumes zero proper motion in the FK5 frame. The `bepoch` parameter is the epoch at which the position in the FK4 frame was determined.

**raFK5toFK4Radians( raFK5, decFK5, bepoch )**

Converts a J2000.0 FK5 position to B1950.0 FK4 yielding Declination. This assumes zero proper motion, parallax and radial velocity in the FK5 frame.

**decFK5toFK4Radians( raFK5, decFK5, bepoch )**

Converts a J2000.0 FK5 position to B1950.0 FK4 yielding Declination. This assumes zero proper motion, parallax and radial velocity in the FK5 frame.

**DEGREE\_RADIAN**

The size of one degree in radians.

**HOURLY\_RADIAN**

The size of one hour of right ascension in radians.

**ARC\_MINUTE\_RADIAN**

The size of one arcminute in radians.

**ARC\_SECOND\_RADIAN**

The size of one arcsecond in radians.

## B.1.14 Coverage

Functions related to coverage and footprints.

One coverage standard is Multi-Order Coverage maps, described at <http://www.ivoa.net/Documents/MOC/> (<http://www.ivoa.net/Documents/MOC/>). MOC positions are always defined in ICRS equatorial coordinates.

MOC locations may be given as either the filename or the URL of a MOC FITS file. Alternatively, they may be the identifier of a VizieR table, for instance "`V/139/sdss9`" (SDSS DR9). A list of all the MOCs available from VizieR can currently be found at <http://alasky.u-strasbg.fr/footprints/tables/vizier/> (<http://alasky.u-strasbg.fr/footprints/tables/vizier/>). You can search for VizieR table identifiers from the VizieR web page (<http://vizier.u-strasbg.fr/> (<http://vizier.u-strasbg.fr/>)); note you must use the *table* identifier (like "`V/139/sdss9`") and not the *catalogue* identifier (like "`V/139`").

**inMoc( mocLocation, ra, dec )**

Indicates whether a given sky position falls strictly within a given MOC (Multi-Order Coverage map). If the given `mocLocation` value does not represent a MOC (for instance no file exists or the file is not in MOC format) a warning will be issued the first time it's referenced, and the result will be false.

**nearMoc( mocLocation, ra, dec, distanceDeg )**

Indicates whether a given sky position either falls within, or is within a certain distance of the edge of, a given MOC (Multi-Order Coverage map). If the given `mocLocation` value does not represent a MOC (for instance no file exists or the file is not in MOC format) a warning will be issued the first time it's referenced, and the result will be false.

**B.1.15 Lists**

Functions which operate on lists of values.

Some of these resemble similar functions in the `Arrays` class, and in some cases are interchangeable, but these are easier to use on non-array values because you don't have to explicitly wrap up lists of arguments as an array. However, for implementation reasons, most of the functions defined here can be used on values which are already `double[]` arrays (for instance array-valued columns) rather than as comma-separated lists of floating point values.

**sum( values, ... )**

Returns the sum of all the non-blank supplied arguments.

**mean( values, ... )**

Returns the mean of all the non-blank supplied arguments.

**variance( values, ... )**

Returns the population variance of the non-blank supplied arguments.

**stdev( values, ... )**

Returns the population standard deviation of the non-blank supplied arguments.

**min( values, ... )**

Returns the minimum of all the non-blank supplied arguments.

**max( values, ... )**

Returns the maximum of all the non-blank supplied arguments.

**median( values, ... )**

Returns the median of all the non-blank supplied arguments.

**countTrue( values, ... )**

Returns the number of true values in a list of boolean arguments. Note if any of the values are blank, the result may be blank as well.

**B.1.16 CoordsDegrees**

Functions for angle transformations and manipulations, with angles generally in degrees. In particular, methods for translating between degrees and HH:MM:SS.S or DDD:MM:SS.S type sexagesimal representations are provided.

**degreesToDms( deg )**

Converts an angle in degrees to a formatted degrees:minutes:seconds string. No fractional part of the seconds field is given.

**degreesToDms( deg, secFig )**

Converts an angle in degrees to a formatted degrees:minutes:seconds string with a given number of decimal places in the seconds field.

**degreesToHms( deg )**

Converts an angle in degrees to a formatted hours:minutes:seconds string. No fractional part of

the seconds field is given.

**degreesToHms( deg, secFig )**

Converts an angle in degrees to a formatted hours:minutes:seconds string with a given number of decimal places in the seconds field.

**dmsToDegrees( dms )**

Converts a formatted degrees:minutes:seconds string to an angle in degrees. Delimiters may be colon, space, characters `dm[s]`, or some others. Additional spaces and leading +/- are permitted. The :seconds part is optional.

**hmsToDegrees( hms )**

Converts a formatted hours:minutes:seconds string to an angle in degrees. Delimiters may be colon, space, characters `hm[s]`, or some others. Additional spaces and leading +/- are permitted. The :seconds part is optional.

**dmsToDegrees( deg, min, sec )**

Converts degrees, minutes, seconds to an angle in degrees.

In conversions of this type, one has to be careful to get the sign right in converting angles which are between 0 and -1 degrees. This routine uses the sign bit of the `deg` argument, taking care to distinguish between +0 and -0 (their internal representations are different for floating point values). It is illegal for the `min` or `sec` arguments to be negative.

**hmsToDegrees( hour, min, sec )**

Converts hours, minutes, seconds to an angle in degrees.

In conversions of this type, one has to be careful to get the sign right in converting angles which are between 0 and -1 hours. This routine uses the sign bit of the `hour` argument, taking care to distinguish between +0 and -0 (their internal representations are different for floating point values).

**skyDistanceDegrees( ra1, dec1, ra2, dec2 )**

Calculates the separation (distance around a great circle) of two points on the sky in degrees.

## B.2 Activation Functions

The following functions can be used only for defining custom Activation Actions (Section 8) - they mostly deal with causing something to happen, such as popping up an image display window. They generally return a short string, which will be logged to the user to give an indication of what happened (or didn't happen, or should have happened). More complete documentation of them is available from within TOPCAT in the Available Functions Window.

### B.2.1 Browsers

Displays URLs in web browsers.

**basicBrowser( url )**

Displays a URL in a basic HTML viewer. This is only likely to work for HTML, text or RTF data. The browser can follow hyperlinks and has simple forward/back buttons, but lacks the sophistication of a proper WWW browser application.

**systemBrowser( url )**

Attempts to display a URL in the system's default web browser. Exactly what counts as the default web browser is system dependent, as is whether this function will work properly.

**mozilla( url )**

Displays a URL in a Mozilla web browser. Probably only works on Unix-like operating systems, and only if Mozilla is already running.

**firefox( url )**

Displays a URL in a Firefox web browser. Probably only works on Unix-like operating systems, and only if Firefox is already running.

**netscape( url )**

Displays a URL in a Netscape web browser. Probably only works on Unix-like operating systems, and only if Netscape is already running.

**mozalike( cmd, url )**

Displays a URL in a web browser from the Mozilla family; it must support flags of the type `"-remote openURL( url )"`. Probably only works on Unix-like operating systems, and only if the browser is already running.

**B.2.2 Mgc**

Specialist functions for use with data from the the Millennium Galaxy Survey.

**imageMgc( mgc\_id )**

Displays the postage stamp FITS image for an MGC object in an image viewer.

**MGC\_IMAGE\_BASE**

String prepended to MGC\_ID for the FITS image URL.

**MGC\_IMAGE\_TAIL**

String appended to MGC\_ID for the FITS image URL.

**B.2.3 System**

Executes commands on the local operating system. These are executed as if typed in from the shell, or command line.

**exec( words, ... )**

Executes an operating system command composed of a command and one or more arguments.

Each of the `words` values is treated as a single (possibly quoted) word in a shell command. The first argument is the filename (either a full pathname, or the name of a file on the current path) of an executable file. These values can be numeric, or strings, or something else, and are converted automatically to string values.

**exec( line )**

Executes a string as an operating system command. Any spaces in the string are taken to delimit words (the first word is the name of the command).

**B.2.4 Sdss**

Specialist display functions for use with the Sloane Digital Sky Server.

**sdssShowCutout( label, ra, dec, pixels )**

Displays a colour cutout image of a specified size from the SDSS around a given sky position. The displayed image is square, a given number of (0.4arcsec) pixels on each side.

**sdssShowCutout( ra, dec, pixels, scale )**

Displays a colour cutout image of a specified size from the SDSS around a given sky position with pixels of a given size. Pixels are square, and their size on the sky is specified by the `scale`

argument. The displayed image has `pixels` pixels along each side.

### **sdssShowCutout( ra, dec )**

Displays a colour cutout image of a default size from the SDSS around a given sky position. The displayed image is 128 pixels square - a pixel is 0.4arcsec.

### **SDSS\_DR2\_BASE\_URL**

Base URL for SkyServer JPEG retrieval service, DR2.

### **SDSS\_BASE\_URL**

Base URL for SkyServer JPEG retrieval service.

## **B.2.5 SuperCosmos**

Specialist display functions for use with the SuperCOSMOS survey. These functions display cutout images from the various archives hosted at the SuperCOSMOS Sky Surveys (<http://www-wfau.roe.ac.uk/sss/> (<http://www-wfau.roe.ac.uk/sss/>)). In most cases these cover the whole of the southern sky.

### **sssShowCutout( ra, dec, pixels )**

Displays a cutout image in one of the available bands from the SuperCOSMOS Sky Surveys. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square. Sky coverage is complete.

### **sssShowCutout( ra, dec )**

Displays a cutout image of default size in one of the available bands from the SuperCOSMOS Sky Surveys. Sky coverage is complete.

### **sssShowBlue( ra, dec, pixels )**

Displays a cutout image of default size from one of the blue-band surveys from SuperCOSMOS. Sky coverage is complete.

### **sssShowRed( ra, dec, pixels )**

Displays a cutout image of default size from one of the red-band surveys from SuperCOSMOS. Sky coverage is complete.

### **sssShowUkstB( ra, dec, pixels )**

Displays a cutout image taken from the SuperCOSMOS Sky Surveys UK Schmidt Telescope Bj-band survey. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square.

Sky coverage is  $-90 < \text{Dec} < +2.5$  (degrees).

### **sssShowUkstR( ra, dec, pixels )**

Displays a cutout image taken from the SuperCOSMOS Sky Surveys UK Schmidt Telescope R-band survey. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square.

Sky coverage is  $-90 < \text{Dec} < +2.5$  (degrees).

### **sssShowUkstI( ra, dec, pixels )**

Displays a cutout image taken from the SuperCOSMOS Sky Surveys UK Schmidt Telescope I-band survey. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square.

Sky coverage is  $-90 < \text{Dec} < +2.5$  (degrees).

### **sssShowEsoR( ra, dec, pixels )**

Displays a cutout image taken from the SuperCOSMOS Sky Surveys ESO R-band survey. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square.



Sky coverage is  $-90 < \text{Dec} < +2.5$  (degrees).

#### **sssShowPossE( ra, dec, pixels )**

Displays a cutout image taken from the SuperCOSMOS Sky Surveys Palomar E-band survey. The displayed image is square, and `pixels` pixels in the X and Y dimensions. Pixels are approximately 0.67 arcsec square.

Sky coverage is  $-20.5 < \text{Dec} < +2.5$  (degrees).

#### **SSS\_BASE\_URL**

Base URL for SuperCOSMOS image cutout service.

### **B.2.6 BasicImageDisplay**

Functions for display of graphics-format images in a no-frills viewing window (an `ImageWindow`). Supported image formats include GIF, JPEG, PNG and FITS, which may be compressed.

#### **displayBasicImage( label, loc )**

Displays the file at a given location as an image in a graphical viewer. `label` may be any string which identifies the window for display, so that multiple images may be displayed in different windows without getting in each others' way. `loc` should be a filename or URL, pointing to an image in a format that this viewer understands.

### **B.2.7 TwoQZ**

Specialist functions for use with data from the the 2QZ survey. Spectral data are taken directly from the 2QZ web site at <http://www.2dfquasar.org/> (<http://www.2dfquasar.org/>).

#### **image2QZ( name )**

Displays the postage stamp FITS image for a 2QZ object in an image viewer.

#### **jpeg2QZ( name )**

Displays the postage stamp JPEG image for a 2QZ object in an external viewer.

#### **get2qzSubdir( name )**

Returns the name of the subdirectory (such as "ra03\_04") for a given 2QZ object name (ID).

#### **TWOQZ\_SPEC\_BASE**

String prepended to the object NAME for the FITS spectra file URL.

#### **TWOQZ\_SPEC\_TAIL**

String appended to the object NAME for the FITS spectra file URL.

#### **TWOQZ\_FITS\_IMAGE\_BASE**

String prepended to the object NAME for the FITS postage stamp URL.

#### **TWOQZ\_FITS\_IMAGE\_TAIL**

String appended to the object NAME for the FITS postage stamp URL.

#### **TWOQZ\_JPEG\_IMAGE\_BASE**

String prepended to the object NAME for the JPEG postage stamp URL.

#### **TWOQZ\_JPEG\_IMAGE\_TAIL**

String appended to the object NAME for the JPEG postage stamp URL.

### **B.2.8 Image**

Functions for display of images in a window. Supported image formats include GIF, JPEG, PNG and FITS, which may be compressed. The SoG program (<http://www.starlink.ac.uk/sog/> (<http://www.starlink.ac.uk/sog/>)) will be used if it is available, otherwise a no-frills image viewer will be used instead.

**displayImage( label, location )**

Displays the file at the given location in an image viewer.

**B.2.9 Output**

Functions for simple logging output.

**print( str )**

Outputs a string value to the user log.

**print( num )**

Outputs a numeric value to the user log.

**B.2.10 Sog**

Functions for display of images in external viewer SOG (<http://www.starlink.ac.uk/sog/> (<http://www.starlink.ac.uk/sog/>)).

**sog( label, loc )**

Displays the file at a given location as an image in a graphical (SoG) viewer. `label` may be any string which identifies the window for display, so that multiple images may be displayed in different windows without getting in each others' way. `loc` should be a filename or URL, pointing to an image in a format that SOG understands (this includes FITS, compressed FITS, and NDFs).

## C Release Notes

This is TOPCAT, Tool for OPERations on Catalogues And Tables. It is a general purpose viewer and editor for astronomical tabular data.

**Author**

Mark Taylor (Bristol University)

**Email**

m.b.taylor@bristol.ac.uk

**WWW**

<http://www.starlink.ac.uk/topcat/>

User comments, suggestions, requests and bug reports to the above address are welcomed.

Related software products are

**STIL**

The Starlink Tables Infrastructure Library, which provides the table handling classes on which TOPCAT is based.

**STILTS**

The STIL Tool Set, which provides some command-line table manipulation and analysis tools based on STIL. This is a non-GUI counterpart to TOPCAT, providing many of the same facilities (matching, row selection, format conversion, plotting etc) but in a form which can be incorporated into scripts, web services, etc.

See the TOPCAT web page, <http://www.starlink.ac.uk/topcat/> for the latest news and releases.

## C.1 Acknowledgements

TOPCAT was initially (2003-2005) developed under the UK Starlink project (1980-2005, R.I.P.). Since then it has been supported by grant PP/D002486/1 from the UK's Particle Physics and Astronomy Research Council, the VOTech project (from EU FP6), the AstroGrid project (from PPARC/STFC), the AIDA project (from EU FP7), grants ST/H008470/1, ST/I00176X/1, ST/J001414/1 and ST/L002388/1 from the UK's Science and Technology Facilities Council (STFC), the GAVO project (BMBF Bewilligungsnummer 05A08VHA), the European Space Agency, and the FP7 project GENIUS. All of this support is gratefully acknowledged.

Nearly all of this work has been done in the Astrophysics Group of the School of Physics, in the University of Bristol, UK, although the University of Cambridge and the University of Heidelberg have also been involved.

Inspiration for some of TOPCAT's features has been taken from the following pre-existing tools:

- Mirage (Bell labs)
- VOPlot (VO-India)

Apart from the excellent Java 2 Standard Edition itself, the following external libraries provide parts of TOPCAT's functionality:

- JEL (Konstantin Metlov, GNU) for algebraic expression evaluation
- PixTools (Nikolay Kuropatkin, Fermilab EAG) for HEALPix-based celestial sphere row matching
- IVOARegistry (NVO) for ADQL/X->S translations in registry searches
- HTM (Sloan Digital Sky Survey) for HTM-based celestial sphere row matching (now

deprecated within TOPCAT)

- Ptpplot (Ptolemy) for some 2D axis plotting
- iText for PDF output
- EPSGraphics2D (Jibble) for encapsulated postscript output
- GifEncoder (Acme) for GIF output
- IVOA FITS Package (Sloan Digital Sky Survey) for simple (non-SoG) display of FITS images
- ADQL (Grégory Mantelet, CDS) for ADQL parsing
- MOC (CDS) Multi-Order Coverage HEALPix maps for footprint manipulation
- nom.tam.fits (NASA) for parts of FITS I/O
- Skyview in a Jar (NASA) for sky axis drawing
- JLaTeXMath (Scilab) for LaTeX typesetting in plots
- JIDE Common Layer (Jidesoft) for double-ended slider controls
- BrowserLauncher2 for launching default web browser

Contributed code from the following people is gratefully acknowledged:

- Thomas Boch (CDS), for the Vizier load dialogue.
- Gerard Lemson (MPE, Garching) for the GAVO load dialogue.
- Marco Molinaro (INAF-OATS) for the BaSTI load dialogue.

The TOPCAT logo in use since version 4 was kindly drawn by Phil Hall.

Colour maps available for the plotting have been taken from various places, including SkyCat/GAIA, MatPlotLib 1.5, Gnuplot, Daniel Michalik, Paul Tol, Color Brewer, HCL Wizard, Dave Green, and probably some others I forgot.

TOPCAT has benefitted from, and is largely shaped by, advice, comments and feedback from very many users, testers and developers. Some of the notable contributors are these:

- Alasdair Allan (Starlink, Exeter)
- Mark Allen (CDS)
- Heinz Andernach (Guanajuato)
- Thomas Boch (CDS)
- Laurent Bourges (JMMC)
- Danny Boxhoorn (Kapteyn, Groningen)
- Baptiste Cecconi (LESIA)
- Igor Chilingarian (CfA)
- Malcolm Currie (Starlink, RAL)
- Clive Davenhall (Royal Observatory Edinburgh)
- Luke Davies (Bristol)
- Christoph Deil (MPI Heidelberg)
- Markus Demleitner (ARI Heidelberg)
- Sebastien Derriere (CDS)
- Emilio Donoso (ICATE)
- Peter Draper (Starlink, Durham)
- Pierre Fernique (CDS)
- Shashikiran Ganesh (PRL, India)
- David Giaretta (Starlink, RAL)
- Eduardo Gonzalez-Solares (IoA)
- Norman Gray (Starlink, Glasgow)
- Phil Hall (Son of Ken)
- Haggis Harris (Bath)
- Evanthia Hatziminaoglou (ESO)
- Avon Huxor (Bristol/ARI)
- Jonathan Irwin (Cambridge)
- Tim Jenness (JACH)

- JJ Kavelaars (NRC)
- Grant Kennedy (IoA)
- Gerard Lemson (MPE Garching)
- Mark Mahabir (Leicester)
- Bob Mann (Royal Observatory Edinburgh)
- Grégory Mantelet (ARI)
- Ben Maughan (Bristol)
- Richard McMahon (IoA)
- Guillaume Mella (JMMC)
- Konstantin Metlov (Donetsk)
- Daniel Michalik (Lund)
- David Mills (Bristol)
- Douglas Morgan (Bristol)
- August Muench (CfA)
- Simon Murphy (ANU)
- John Murrell
- Ada Nebot (CDS)
- François Ochsenbein (CDS)
- Clive Page (Leicester)
- Menelaus Perdikeas (ESAC)
- François-Xavier Pineau (CDS)
- Roy Platon (RAL)
- James Price (Bristol)
- Paul Price (Princeton)
- Anita Richards (Jodrell Bank)
- Kristin Riebe (AIP)
- Aaron Robotham (Bristol)
- Yannick Roehlly (OAMP)
- Mike Read (ROE)
- Henrik Rydberg (Gothenburg drinking water laboratory)
- Juan Carlos Segovia (ESAC)
- Petr Škoda (AIAS, CZ)
- Mattia Vaccari (Western Cape)
- Juande Sandander Vela (IAA-CSIC)
- Mike Watson (Leicester)
- Ivan Zolotukhin (SAI MSU)
- Jonathan Zwart (Columbia)

If you use this software in published work, the following citation would be appreciated:

2005ASPC...347...29T: M. B. Taylor, "TOPCAT & STIL: Starlink Table/VOTable Processing Software", in *Astronomical Data Analysis Software and Systems XIV*, eds. P. Shopbell et al., ASP Conf. Ser. 347, p. 29 (2005)

## C.2 Version History

Releases to date have been as follows:

### **Version 0.3b (4 June 2003)**

First public release

### **Version 0.4b (8 July 2003)**

- Row subset count column in subsets window.
- Fixed and improved broken Parameter window.
- Fixed output of table name and parameters.

- Output to LaTeX `tabular` environment now available.
- SQL access buttons now greyed out when no JDBC drivers are present.
- UCD selection now available from New Synthetic Column dialogue.
- Column metadata display made more logical and flexible.
- Column cardinality now calculated in Stats window.
- Synthetic column expressions and most other column metadata now editable.
- Null value support in evaluated expressions.
- Integral example data provided.
- Hierarchical browser for tables available in load dialogue.

#### **Version 0.4-1b (10 July 2003)**

- Fixed a VOTable output bug (not escaping XML special characters).
- Fixed a serious FITS output bug.
- Improved bad value handling for FITS tables.

#### **Version 0.5b (20 October 2003)**

- Can now read plain text format tables.
- FITS files of arbitrary size can now be read (they are not loaded into memory).
- TOPCAT can now run without errors in a secure environment (e.g. as an unsigned jar file under WebStart). Of course some actions such as Save Table are unavailable in this context.
- Files compressed with Unix `compress` now work (as well as `gzip` and `bzip2`).
- Added hierarchy browser to load dialog.
- Added integral demo data (accessible from load dialog).
- Can now drag tables from Treeview into TOPCAT load dialog (or load button).
- Some bugfixes.

#### **Version 0.5-1 (18 November 2003)**

- Rewrite of FITS binary table access for big efficiency improvements.
- It's now possible to plug in user-defined algebraic methods at runtime.
- Improved unit testing leads to some bugfixes.

#### **Version 1.1-0 (21 April 2004)**

- User interface redesigned - now based around Control Window not table browser (much easier to work with multiple tables).
- Extensive facilities for table joining by matching rows between multiple tables or within a single one.
- Documentation much improved and available on- or off-line as SUN/253.
- Self-contained table access library STIL now provided as a separate product.
- Tables can be concatenated.
- Better top-level control over window proliferation.
- Columns can now be hidden/revealed not just deleted.
- Additional output formats/variants supported:
  - VOTable using BINARY or FITS encoding, inline or by reference
  - Machine-readable plain ASCII stream
  - HTML `<TABLE>` element or document
- Hybrid DOM/SAX parsing of VOTables for improved efficiency/memory usage.
- New flag `-demo` starts up with demo data.
- Miscellaneous efficiency improvements and bug fixes.

#### **Version 1.1-3 (5 May 2004)**

- Functions provided for radians $\leftrightarrow$ sexagesimal conversion

**Version 1.3 (20 October 2004)**

This version has introduced many improvements in scalability, efficiency and functionality. TOPCAT is now quite happy with tables of a million rows or more (and hundreds of columns) even on systems with quite modest memory/CPU resources. The main improvements are as follows:

**Plotting**

- Plotting is much faster and can handle many more points
- Subsets can be selected from plot window by tracing out a non-rectangular region
- You have more choice over plotting symbols (including semi-transparent ones)
- Finally X or Y axes can be flipped!
- Export to encapsulated PostScript is of improved quality (though for many points file sizes can get large)
- Export to GIF format is available
- Regression lines can be plotted and coefficients displayed (experimental capability - could be improved)

**Table Formats**

- "-disk" flag allows use of disk backing storage for large tables
- New 'FITS-plus' format stores rich table/column metadata in a FITS file
- VOTable handler now fully VOTable 1.1 and 1.0 compatible
- VOTable parsing now works with Java 5.0 platform
- Comma-Separated Value format now supported for input and output
- ASCII input handler rewritten to cope with much larger tables
- ASCII handler now understands d/D as exponent letter as well as e/E
- ASCII handler now uses Short/Float not Integer/Double where appropriate to save memory
- ASCII format fixed bug for -0 degrees/hours in sexagesimal angles
- Null handling improved for FITS & VOTable formats
- FITS files store column descriptions in TCOMMx headers
- Better error messages for unparsable tables

**Table Joins**

- Various efficiency improvements and reductions in memory requirements
- In cases of multiple possible matches, the closest is now chosen rather than picking one at random
- Pair match now adds column containing score for each match (distance between points)
- Units can be selected RA/Dec columns and match errors (so it doesn't need to be all in radians)
- New match types suitable for multivariate matching (anisotropic Cartesian, Sky+X, Sky+XY)

**Data/Metadata Manipulation**

- Can add/remove table parameters
- One-step column replacement dialogue from data or column windows
- Synthetic column expressions now written out to column descriptions

**GUI Navigation and Display**

- Improved rendering of numbers in tables (esp. Floats)
- Better detection of displayed table column widths
- New Control Window option on File menus
- Better window resizing for some dialogue boxes

- Less confusing error messages in many places

### **Algebraic Expressions**

- All available functions are now fully documented in help document and interactive Method Window
- Many new trig, coordinate, type conversion, string manipulation functions
- Big performance improvements for null values

### **Activation Actions**

- Clicking a point in the plot highlights the corresponding row in the data window and vice versa
- Row selection can trigger display sky cutout region display
- Row selection can trigger user-defined actions on activation

In addition, the following incompatibilities and changes have been introduced since the last version:

- The input format for tables can now be specified in the load window (via a selection box) or on the command line (using the "-f" flag). FITS files and VOTables can still be identified automatically (i.e. it's not necessary to specify format in this case) but ASCII tables cannot: you *must* now specify the format when loading ASCII tables. This change allows better error messages and support for more text-like formats.
- Algebraic expressions referencing row subsets now use the underscore character ("\_") rather than the hash character("#") to indicate a subset ID.
- Classes containing user-supplied functions for algebraic expressions are now specified using the properties "jel.classes" and "jel.classes.activation", not "gnu.jel.static.classes".
- The default output format for FITS tables is now the so-called "FITS-plus" format, which has a BINTABLE in its first extension as before, but the text of a VOTable stored in its primary HDU. This can store more metadata for TOPCAT, but should behave just the same for other FITS-compatible applications. The old behaviour can be restored if desired by specifying "FITS-basic" format.

### **Version 1.3-1 (10 November 2004)**

Minor changes:

- 2MASS cutout servers now available from Activation Window
- Added Starlink logo to all windows

### **Version 1.3-2 (6 Dec 2004)**

Bug fix:

- Error in parsing empty VOTable TD elements fixed.

### **Version 1.4 (4 Feb 2005)**

#### **Load Dialogues**

The graphical table load dialogue has been overhauled, and now has two main new features. First, it has been rewritten so that the GUI does not freeze during a long load; it is still currently not possible to interact with other TOPCAT windows while a load is taking place, but you can now cancel a load that is in progress.

Secondly, the provision of load dialogues has been modularised, and a number of new dialogues provided. The new ones are:

- Cone Search
- MySpace Browser



- Registry Query
- SIAP Query

If the required classes are present, you can acquire tables from these external sources as well as the traditional methods of loading from disk etc. New command line flags corresponding to each of these have been added to ensure that they are present and make them prominent in the load dialogue. Furthermore it is possible to plug in additional load dialogues at runtime using the `startable.load.dialogs` system property.

The appearance of the **Load Window** has changed; now only the **File Browser** button is visible along with the **Location** field in the body of the window, but the **DataSources** menu can be used to display other available table import dialogues.

### Packaging

The program can now be obtained in two standalone forms: `topcat-full.jar` and `topcat-lite.jar`. The former is much larger than before (11 Mbyte), since it contains a number of classes to support custom load dialogues such as the MySpace browser and web service interaction, as well as the SoG classes. The latter contains only the classes for the core functionality, and is much smaller (3 Mbyte).

### Explode Array Column action

There is now a new button in the Columns Window which replaces an array-valued column with a scalar column for each of its elements.

### Paste'n'Load

You can now load a table by pasting its filename or URL as text into the table list in the Control Window (using the X selection on X-windows - not sure if or how this works on other platforms).

### Help message

The result of `topcat -help` is now more comprehensive, describing briefly what each option does and listing system properties as well as arguments/flags proper.

## Version 1.4-1 (8 February 2005)

- Added Search Column item to Data Window column popup menu

## Version 1.5 (17 March 2005)

### File Access

Load dialogues have changed again somewhat, and save dialogues as well. The default file browser in both cases is now a *Filestore Browser*, which is very much like the standard file browser, but can browse files in remote filesystems as well; currently supported are files in AstroGrid's MySpace or on an SRB (Storage Resource Broker) server. You can now save files to these remote locations as well as load from them.

In addition, the save dialogue now displays the current row subset and sort order - this makes it easier to see and/or change the details of the table you're about to save.

### BugFixes

A few more minor changes have been made.

- Error display dialogue boxes have been improved in some places
- Various bugs relating to JDBC database access have been fixed
- Some minor issues relating to VOTables with single-character columns have been addressed

## Version 1.6 (30 June 2005)

### Activation Actions

Some more activation functionality has been added:

- New **View URL as Web Page** option introduced in Activation Window
- New **System** class of activation functions containing `exec` functions which execute commands on the local operating system
- New **Browsers** class of activation functions for displaying URLs in web browsers (external or basic fallback one)

### Algebraic Functions

New **Times** class added containing functions for converting between Modified Julian Day and ISO 8601 format epochs.

### Sky Matching

The default sky matching algorithm now uses HEALPix rather than HTM for assigning sky pixels to RA,Dec positions. This gives much faster sky matches in most cases, and uses somewhat less memory so can be used on larger tables. It has also fixed a bug which missed out some possible matches. HTM-based matching is currently still provided as an option, but this is mainly for debugging purposes and may be withdrawn in the future.

### Logging

The message logging has been tidied up. The main observable consequence of this is that fewer untidy messages are written to the console when TOPCAT is run from a standalone jar file rather than a full starjava installation. By specifying the new `-verbose` (or `-v`) flag one or more times you can get those messages back. The messages (in fact all logging messages at any level) can also be viewed from the GUI by using the new **File|Show Log** menu option from the Control Window.

### SOAP Services

TOPCAT now acts as a SOAP server; SOAP requests can now be made to a running instance of TOPCAT to get it to display tables by location or by sending XML for a VOTable direct. Because of limitations in Axis, this latter method won't work for arbitrarily large tables.

### Documentation changes

The `tablecopy` tool is no longer covered in this document; it is replaced by the `tcopy` tool in the separate STILTS (<http://www.starlink.ac.uk/stilts/>) package. There has also been some reorganisation of this document, mainly in the appendices.

### Minor changes

- Added `-version` flag
- Added (dummy) **Print** option to Data Window. This just presents a message to the effect that you should save to a printable format.
- Fixed a bug which gave errors when expressions contained a `NULL_` test on the first column of a table.
- Modified one of the demo tables to contain a column with URLs in it.

### Version 1.6-1 (7 July 2005)

Bugfixes:

- Work around AstroGrid/Sun bug which prevented loading short VOTables from MySpace.
- Ensure that filestore browsers are kept up to date when dialogues are displayed.

### Version 1.7 (30 September 2005)

#### Crossmatching

There have been major improvements in the flexibility, and minor improvements to performance, of two-table crossmatching.

- New match algorithm **Sky with Errors** introduced. This allows you to specify a

column giving the maximum permissible match error (so it can vary per row) rather than a fixed value for the whole table.

- In the case of multiple possible matches between the two tables, instead of automatically giving you only the closest match, you can now select whether you'd like only the closest one or all those which fit your criteria.
- You can now specify which rows you want to see in the output: 1 and 2, 1 or 2, All from 1, All from 2, 1 not 2, 2 not 1, 1 xor 2. This is pretty much all the possibilities which make sense, and in particular allows you to do 'left outer joins' (1 not 2).
- The match score column which results from most matches now comes (a) in sensible units where possible (e.g. arcseconds not radians) and (b) with metadata which tells you what its meaning and units are.
- More information is available in added columns after the match; as well as the match score, information about matched groups is inserted where appropriate.
- The "Spherical Polar" match algorithm is now rebadged as the hopefully less confusing "Sky 3d".

Similar changes for 1-table and multi-table matches should follow in future versions.

### MySpace Access

MySpace I/O has been re-implemented to use the ACR rather than the (now deprecated) CDK classes it was using before. As well as probably being more reliable and less likely to break with future changes in MySpace server protocols, this gives the benefit of single sign on. The effect of this is that you will need to have the AstroGrid desktop running on your machine before you can access MySpace from TOPCAT.

### Algebraic functions

- Added Julian Epoch and Besselian Epoch conversion functions to `Times` class.
- Added `RANDOM` special function.

### Miscellaneous

- When you select a column in the Columns window, it now scrolls the table in the Data Window so that the selected column is visible. This is a boon when dealing with tables that have very many columns.
- String "null" interpreted as a blank value in ASCII tables.
- Added new activation action to launch system default browser.

### Bugfixes

- Fixed some relatively harmless bugs to do with actions available when you select the dummy "Index" column. You can now unsort from a popup menu in the table viewer window.
- Believed to work fine with Java 1.5 now (there were previously some issues with MySpace at Java 1.5).
- Fixed bug in ASCII input handler which misidentified blank lines, or DOS-format line ends, as end of file.

### Version 1.7-1 (4 October 2005)

Bugfixes:

- Fixed broken MySpace access on MS Windows.

### Version 1.8 (13 October 2005)

- Added **Sky Coordinates Window**; it's now easy to add new sky coordinate columns based on old ones in different coordinate systems.
- `roundDecimal` and `formatDecimal` functions introduced for more control over visual appearance of numeric values.

- Now copes with 'K'-format FITS binary table columns (64-bit integers).
- Modifications to JNLP files.

### **Version 2.0 (3 February 2006)**

A major upgrade of TOPCAT's visualisation capabilities has taken place in this release. There are considerable improvements in functionality, flexibility and efficiency over previous versions:

#### **New graphics windows**

In addition to the 2-d scatter plot from previous versions, the following visualisation windows are now available:

- Histogram (1-d)
- 3-d Cartesian scatter plot
- 3-d Spherical scatter plot (with optional radial dimension)
- 2-d Density map (2-d histogram)

The new 3-d functionality does *not* require you to install Java3D or any other third-party 3D toolkit to work (nor does it take advantage of any such toolkit which may be present).

#### **Multi-dataset/multi-table plotting**

The plot windows are no longer associated with a single table. All of them allow you to display data from different tables, or from different tuples of columns of the same table, on the same plot. You can layer as many plots as you like on the same axes, using different plotting styles for the different datasets. As before, you can still display data from different subsets of the same table and same columns using different styles.

#### **Plotting Styles**

All the graphics windows allow you to set the plotting style for each data set individually, using a wide range of options including colour, line width, marker size, (histogram) bar style, etc. This allows you considerable control over the visual details of the plots.

#### **Transparent markers**

All the 2-d and 3-d scatter plots allow you to render points using markers of variable transparency. In a crowded plot, this allows you to see much more information than using opaque points, since you can get some idea of how many points (of different data sets) have been drawn at a given point on the plotting surface.

#### **Line drawing & linear correlation**

The 2-d scatter plot can now optionally plot lines associated with data sets. It can either draw straight line segments joining all the plotted points in a set, or draw per-dataset linear correlation lines. In the latter case it will report line gradient, intercept and correlation coefficient.

#### **Improved column selection**

The selector boxes for selecting which columns to plot are now 'editable' - that is, instead of selecting the column from a drop down list it is now also possible to type an expression into them instead. This may be more convenient if there is a very long list of columns. It also means that you can use an algebraic expression based on the names of one or more columns instead of a simple column name. The selectors also have small arrow boxes next to them which makes it easy to cycle through the list of known columns. These features are also available at some other places in the program where a column value is required.

#### **Manual axis configuration**

As well as zooming in and out using the mouse, you can now set the axis limits by typing them into an axis dialogue box. You can also set the text which will form the axis annotation on the plot.

#### **Status line**

Most plots now feature a panel at the bottom of the window indicating how many points

have been plotted and the current position of the mouse pointer (if any) on the plotting surface.

### **Performance**

Scatter plots of large datasets now use considerably less memory and around an order of magnitude less CPU time than previously (a 2-d million point replot now takes about 1 second - plotting it the first time may be rather longer since it needs to acquire the data which may be I/O intensive).

Some non graphics-related improvements have also been made as follows:

- Selection of a subset in the Control Window now triggers its selection in other windows (plot, statistics, subsets). The same thing doesn't happen the other way around, since that might lead to confusing consequences.
- Boolean columns now display null values distinctly from false ones. Additionally, null/false distinctions are handled more carefully in FITS and VOTable files.
- The Sky Coordinates Window now suggests names for new columns.
- The Filestore Browser now allows you to enter the position in a file of the table to load (e.g. HDU index for FITS or TABLE index for VOTable).
- Added Hide All & Reveal All actions to the Columns window.
- When joining tables, column name comparison to determine whether deduplication is required is now case-insensitive.
- Fixed a problem which was causing TOPCAT to crash when attempting to save an altered copy of a FITS file under the same name.
- The manual has been reorganised somewhat, and a new Quick Start (Section 2) section added.
- There is an experimental implementation of the Aladin interoperability interface. This hasn't really been tested however, so may not work. Improved Aladin interoperability is expected in future releases.
- Fixed a bug in Cartesian crossmatching algorithms which failed to match if the required error in any dimension was zero.

### **Version 2.1 (7 April 2006)**

A number of graphical and other improvements have been made at this release.

#### **Stacked Line Plot**

A new Stacked Line Plot (Appendix A.5.4) visualisation window has been added. This is especially suitable for use with time series data.

#### **Asynchronous data reading in graphics windows**

All the graphics windows now read data for plotting asynchronously. What this means is that when you change the plot in a way which requires the data to be read or re-read then the GUI will not lock up and you can do other things, including changing the plot in other ways before it has completed drawing. A progress bar at the bottom of the window indicates progress. This is only noticeable for large (slow to read) files.

#### **Axis Zooming**

The existing Histogram, 2-D Scatter Plot and Density Map windows, as well as the new Stacked Line Plot, now allow you to do a 1-d zoom by dragging the mouse near the axis, as well as the 2-d zoom by dragging on the plot surface.

#### **PLASTIC Tool Interoperability**

TOPCAT now sends and services messages using the PPlatform for AStronomical Tool InterConnection protocol. See Section 9.

#### **IPAC Data Format**

Data files in the IPAC format defined by the CalTech's Infrared Processing and Analysis Center can now be read. This is how data from the Spitzer satellite amongst others is normally provided.

### String-typed coordinate columns

If suitably identified (e.g. by UCD or units), string-valued columns which represent data in certain well-known forms, currently sexagesimal angles and ISO-8601-format dates, can now be treated directly as numeric values in column selector boxes such as the ones in the plotting windows, rather than having to define a new column using a string->numeric converter.

### Memory Error Reporting

When the JVM runs out of memory, it now attempts to post a popup window explaining this rather than just writing a terse message to the console. This popup also gives you the opportunity to view a new section added to the documentation: Tips for Large Tables (Section 10.4).

### SOAP Service deprecated

TOPCAT no longer by default runs a SOAP server for accepting tables. You can choose to run it by specifying the `-soap` flag on the command line. This facility may be withdrawn in future versions, in view of the fact that the PLASTIC service can provide similar functionality.

### Logo fade

The Starlink logo at the top of every window is set to fade out gradually over the period 1-APR-2006 to 1-SEP-2006.

### Bug fixes and minor alterations

- Export of GIF and Encapsulated PostScript images from the graphics windows has been improved - unwanted grey backgrounds round the edges are now rendered transparent in GIFs and white in EPS. A section on Exporting Graphics (Appendix A.5.1.7) has also been added to the manual.
- A number of graphical anomalies such as the plot bounds being reset when an axis Flip button was toggled have been ironed out.
- The JDBC behaviour has been improved: some bugs have been fixed and it now no longer asks for an SQL username/password twice.
- Work around intermittent bug when loading small files from MySpace.
- Density maps are now exported into FITS format with better WCS headers where appropriate.
- The SDSS JPEG cutout activation action URL has been updated from DR2 to DR4.
- A bug which caused the first row in a table to be included erroneously in graphical subset selections has been fixed.
- Added an offset selector toggle button to the Histogram window.

## Version 2.1-1 (13 April 2006)

### Sphere Zoom

You can now zoom in the Spherical Plot.

### PLASTIC

There are some improvements to the PLASTIC functionality and documentation.

- In addition to **Broadcast** actions for the various PLASTIC messages transmission types, there are now corresponding **Send** actions which allow you to direct the message to only a single selected listener application. These actions are now enabled/disabled according to whether any suitable PLASTIC listeners are registered.
- You can now pop up a little window from the **Control Window's Interop** menu which displays the currently registered PLASTIC listeners. The functionality of this window may be extended in future releases.
- When TOPCAT receives a `showObjects` message, it now checks if a matching subset exists rather than always creating and adding a new one. If it does, it just sets

current the existing one. This can cut down (a bit) on proliferation of Row Subsets.

- The PLASTIC section of the manual has been improved.
- Some of the screenshots have been updated to include PLASTIC-related features.
- URLs using the `file:` scheme sent by TOPCAT in PLASTIC messages now correctly conform to RFC 1738.

### MySpace

A workaround has been introduced for MySpace browsing performance problems. Run with `-Dmyspace.cache=true` to speed it up at the expense of accuracy.

### Version 2.1-2 (21 April 2006)

This is mainly a bug fix release.

- Added explode columns option to Data Window popup menu.
- Fix bug present since v2.0 which was sometimes drawing wrong numeric labels near zero in 2D plots. A side effect is that exponential multipliers will more often be used on axis labels (as in versions prior to v2.0). Hopefully somewhat improved axis labelling will be present in future versions.
- Fixed bug which was writing unclosed LINK elements in VOTables.
- Fixed some NullPointerException bugs in column selection and 2D plot.
- Fixed bug in scrolling to correct column in Data Window when column is selected in Columns Window.

### Version 2.1-3 (11 May 2006)

Bug fixes and some minor enhancements.

- Implements send and receive of PLASTIC `ivo://votable.org/votable/highlightObject` message.
- Added new `csv-noheader` output format.
- Added skew and kurtosis calculations to Statistics Window.
- Improved, though still imperfect, retention of table-wide metadata in VOTables.
- Fixed bug in writing `votable-fits-href` and `votable-binary-href` format tables from the file browser.
- `mark.workaround` system property, see Section 10.2.3.

### Version 2.2 (7 July 2006)

#### Column-oriented storage

New features for permitting column-oriented storage (`colfits` format, `startable.storage` policy `"sideways"`) have been introduced. These can provide considerable efficiency improvements for certain tasks when working with very large (and especially wide) tables.

#### Bug fixes

- Fixed bug which caused some points to be missed out in spherical plots.
- ISO-8601 strings now permit times of 24:00:00 as they should.
- Quoted unit values are examined to determine probable column data types.

#### Other items

- Added flux conversion functions (`Jansky<->magnitude`).

### Version 2.2-1 (3 August 2006)

- Added new coordinate system ICRS for sky coordinate conversions.
- Eliminated some unsightly but harmless stack dumps that could occur during plotting.
- TUCDnn header cards now used in FITS files to transmit UCDs (non-standard

mechanism).

- Added TST (Tab-Separated Table) format input and output handlers.
- Efficiency improvements for column-oriented access.
- Improved responsiveness when viewing table data for extremely large tables.

### Version 2.3 (5 October 2006)

Various modifications and improvements:

- Removed Starlink logo from the top right corner of all TOPCAT windows, including screenshots in documentation. New logos (Starlink, Astrogrid, VOTech, Bristol) are added in **About** window.
- Added new **Add Sample Subset**, **Add Head Subset** and **Add Tail Subset** utility tool buttons to the Subset Window.
- Direct MySpace access using `ivo:` or `myspace:` URLs is now provided - see new Section 4.2.
- Added time conversion functions between MJD and Decimal Year.
- ISO-8601 column numerical values (e.g. in plots) are now by default Decimal Year rather than Julian Epoch. The numerical values are quite close, but Decimal Year is probably more what you'd intuitively expect.
- Added `toHex` and `fromHex` numeric conversion functions.
- Added `-J` flag to `topcat` startup script for passing flags directly to Java.
- Added Sample (as opposed to Population) Standard Deviation/Variance calculation options to the Statistics Window.
- Improved listing for functions in Appendix B of user document (subsection for each function class).
- Modified presentation of HTML version of user document using CSS.

Bug fixes:

- Fixed bug (introduced at v2.2-1) which made table display and adding columns slow, and column widths the wrong size.
- Fixed bug (introduced at v2.0) which prevented Exact matches using non-numeric columns. Exact match now also deals with array-valued columns.
- Fixed PLASTIC bug which was ignoring ID argument of `ivo://votech.org/votable/loadFromURL` message.
- Fixed bug in coordinate conversion that caused bad behaviour in the presence of malformed input coords.
- Fixed bug with plotting very large or very small axis labels in 3-d plots.
- Fixed a couple of minor crossmatching bugs (which would not have affected results).

### Version 2.3-1 (Starlink Hokulei release)

- Output to MySpace can now be streamed, if running under J2SE1.5 or later.
- Intercept `OutOfMemoryErrors` during 3D plotting to give an explanatory popup.
- Fixed bug in axis range checking.
- Fixed bug in handling of single quotes in FITS file metadata.
- Concatenation column assignments now smarter for fully compatible input tables.
- Some optimisations for very large files.
- Added Vega $\leftrightarrow$ AB magnitude conversion constants to Fluxes functions.
- Encapsulated PostScript exported from graphics windows is improved. These files should now print when sent directly to a printer, and the image size is now guaranteed sensible for a portrait A4/letter page.
- Added hyperbolic trig functions (`sinh`, `cosh`, `tanh` and inverses) Maths class (`sinh`, `cosh`, `tanh` and inverses).
- Added cosmology distance calculations in class Distances.

### Graphics upgrades



- Column selector panels in graphics windows are now in scrollers.

### Version 3.0 (5 July 2007)

This release includes major visualisation improvements and some changes to the plotting user interface:

#### Error bars

Comprehensive plotting of symmetric and asymmetric error bars (and crosshairs, and boxes, and ellipses, ...) in 1, 2 and 3 dimensional plots is now provided -- see Appendix A.5.1.3.

#### Auxiliary (colour) axes

It is now possible to modify the colour of plotted points according to values in one or more axes additional to the spatial ones -- see Appendix A.5.1.5.


#### Histogram Weighting

The Histogram and Density Map windows now have the option of weighting the quantity represented, so that the binned values are sums rather than just counts.

#### Density Map Colour Scale

When the density map is in indexed (non-RGB) mode you can now select from various colour maps to represent bin occupancy. This makes level differences easier to see.

#### Plot window rearrangement

Buttons which affect data selection are now arranged in their own toolbar below the plot (Appendix A.5.1.1). The size of the plot control area can be manually resized to give more room for the plot itself by using the **Split Window** () button, which is useful on

small screens.

#### Expression language enhancements

Table parameters can now be referenced in expressions using `param$` notation (Section 7.3), and both columns and parameters can be referenced by UCD using `ucd$` notation (Section 7.1).

#### Statistics & Histogram result re-use

The Statistics and Histogram windows now both provide **Save** and **Import** actions, which take the data presented in the window and allow it to be saved to a file or imported directly into TOPCAT as a new table.

#### Cartesian 3D zoom

It is now possible to zoom the 3D Plot towards the centre in the same way as the Spherical Plot.

#### GAVO Millennium simulation database load dialogue


You can now directly query the GAVO Millennium database service (Appendix A.6.10). Thanks to Gerard Lemson of GAVO for contributing code for this.

#### Row subset improvements


Row Subset names can now be re-used, and when you are asked to select a new subset name you are allowed to choose one from the list of existing ones. When a row selection is received from another application via PLASTIC the subset name (the name of the sending application) is re-used in this way. This helps to keep the number of subsets under control.

Receiving a row subset from PLASTIC in this way, and certain other actions, now cause the subset to be shown straight away (and updated if necessary) on any existing plots, which makes this kind of PLASTIC interaction more responsive.

The size of each subset, and also the corresponding percentage of the table it represents, is now calculated automatically and displayed in the Subset Window. The old behaviour of

 menu item if required.

### Fix Vertical in 3D plots

A **Stay Upright** () button is provided in the Cartesian and spherical 3D plots which allow you to keep the Z axis/north pole vertical during rotations.

### Quantile calculation

The Statistics Window can now calculate quantiles (median, quartiles, .001, .01, .99 and .999) on request.




### Minor changes and bugfixes

- JPEG and PNG now plot export available from all graphics windows.
- Dramatic performance improvement for transparent pixel plotting in 3D plots on some platforms.
- Added locale-specific `formatDecimalLocal()` functions in class `Formats`.
- Added `fluxToLuminosity` and `luminosityToFlux` functions in class `Fluxes`.
- Improved deduplication of column names when joining tables.
- 3D plots now correctly report number of visible points.
- Improved error message for broken CSV files.
- Improved error reporting in the presence of a deficient JVM (such as GNU `gcj`).
- Worked around a Mac OSX Java bug which caused table display to go funny when columns were deleted.
- Fixed error in output of FITS table `TNULL n` header cards - write them as numeric not string values.
- Fixed a number of bugs in responding to PLASTIC messages (threading, failure to respond).
- 3D plots now plot symbols for all subsets not just one, as for 2D plots.
- Fixed bug which prevented non-string parameter values from being reassigned.

### Version 3.0-1 (Starlink Puana release)

- Added error bars capped by arrows.
- Fixed bug which caused error bars in legend to draw in the wrong colour in EPS output.

### Version 3.1 (29 August 2007)

- You can now annotate plotted points using text labels (Appendix A.5.1.4) in 2d and 3d scatter plots - use the **Draw Label** () button to assign a column containing label text.
- Radial axis control in sphere plot is now not visible by default - you need to click the **Radial Coordinates** () button in the dataset toolbar to enable it.
- New **Show Legend** () button in plot windows determines whether a legend will be displayed next to the plot.
- New **Hide Legend** checkbox in style editor windows which allows subset legend entries to be excluded from the displayed legend individually.

### Version 3.2 (28 January 2008)

#### Functionality enhancements

- Considerably improved display of per-table metadata in the Parameter Window. It is now possible to view long or multi-line parameter values.
- FITS header cards in table HDU are now read (but not written) as table parameters.
- Add mouse wheel zoom for 3D plots.


- Add **Write Mode** selector to SQL Output Dialogue, to allow appending and optional drop prior to create of new table in database.
- Now TOPCAT will automatically register with a PLASTIC hub any time one becomes available - the hub does not have to be running when TOPCAT starts. As a corollary, if you start an internal or external hub from the **Interop** menu, you no longer need to explicitly invoke the **Register** item afterwards.
- Add `-exthub` flag which starts a new external PLASTIC hub.
- Activation window permits selection of different image viewers for **View URL as Image** option. Note this is not currently provided for the **Display Cutout Image** option - this may be addressed in the future.

### Bugfixes and minor improvements

- Added `-stilts` convenience flag so you can easily run STILTS from a TOPCAT installation.
- Fix calculation error in `fluxToLuminosity` function.
- Fix some bugs associated with text label plotting.
- Error bars for zero/negative values are now correctly extended to the axis (rather than being omitted) in 2D plots with logarithmic axes.
- Modified the size of some windows - now posts somewhat smaller windows if a small screen is in use.
- System properties specified explicitly on command line now take precedence over those in `.starjava.properties` file.
- Embedded spaces in output ASCII format table column names are now substituted with underscores, which prevents writing ASCII tables that cannot be read.
- Downgraded from WARNING to INFO log messages about the (extremely common) VOTable syntax error of omitting a FIELD/PARAM element's `datatype` attribute.
- Starting an external hub now works more reliably.

### Version 3.3 (19 August 2008)

#### Functionality enhancements:

- Added normalisation option () to Histogram Window.
- In overplots using different tables, an attempt is now made to use the same columns/expressions for axis values as for the main table. This may be convenient when overplotting data from several similar tables.
- FITS files with columns using variable-length arrays ('P' or 'Q' data type descriptors) can now be read (in random access mode, i.e. from an uncompressed FITS file on disk, only).
- The **lut.files** system property can now be used to configure custom colour maps for auxiliary axes and density maps - see Section 10.2.3.
- New class Arrays added to algebraic functions.
- Histogram **Offset** control now works for logarithmic X scale as well as linear.

#### Bug fixes and other changes:

- Efficiency improvements (~25%? in both CPU time and memory usage) for HEALPix-based sky crossmatching. Thanks to Nikolay Kouropatkine at Fermilab for a new version of the PixTools library which provided this improvement (this version also fixed a - minor? - HEALPix bug).
- Sexagesimal field identification for ASCII input files has become more forgiving; it now accepts minutes or seconds equal to 60 without a syntax error.
- Fixed a pair matching performance bug (slower if tables were not given in the right order) introduced at v1.4.

- The `-disk` flag is now honoured when loading tables from JDBC, which makes it possible to input larger datasets from RDBMS.
- Fixed problem which caused the graphics system to hang or fail when drawing Dot-to-dot lines in 2D plot for points a very long way off the screen.
- Fixed bug which caused incorrect plotting of cumulative histograms with small weights.
- Slight modification to spike-style bar drawing in Histogram, now has tidier appearance when using dashed lines.

### Internal

There has been extensive internal reorganisation of the plotting classes. There should be no user-visible effects of this, but please report anything which seems to be misbehaving.

## Version 3.4 (23 December 2008)

### SAMP

TOPCAT now uses SAMP as a (preferred) alternative to PLASTIC for inter-tool interoperability. SAMP is now usually the default; the old behaviour can be recovered using the `-plastic` command-line flag. Although much of the messaging behaviour remains the same, various improvements to the GUI accompany this change:

- New SAMP Panel at the bottom of the Control Window summarising status
- New **SAMP Status** ( ) toolbar button in Control Window, which pops up new

SAMP Window

- Other slight rearrangements of toolbar buttons and menu items in Control Window
- **Broadcast Row** checkbox added near Activation Action in Control Window

For more discussion see Section 9.

### Registry Access

TOPCAT's registry access has finally been updated to use VOResource 1.03 and the Registry Interface 1.01 (it was previously using very out of date protocols). This means that the Cone Search dialogue (as well as the semi-supported SIAP and Registry dialogues) will now pick up a much more up to date set of services. Additionally, you can choose your own registry service, and the registry searches in those dialogues now feature keyword searches rather than picking up all known cone/SIAP services. See Appendix A.9.2.

### Other enhancements

- Table columns and parameters can now be referenced by Utype using `utype$` syntax in a similar way to using UCDs - see Section 7.1.
- When a new subset is created by whatever method, the user can now elect to transmit it to other applications using SAMP/PLASTIC as an alternative to adding a new subset to the table's subset list.
- Add items to **Help** menu for viewing help in the default web browser rather than TOPCAT's help browser.
- Add a plot output format option to write gzipped PostScript.
- `-version` command line flag now reports subversion revision as well as symbolic version number.

### Bugfixes

- Fixed histogram bug (blank plot) for small values when using logarithmic Y axis.
- Fix stacktrace error in 3D plots when no data is available.
- The SIMBAD name resolver is fixed in the cone search window (it was using a no-longer-supported service format).

**Version 3.4-1 (27 March 2009)**

- Note Sybase ASE works (documentation edit only)
- Fixed bug which caused SIAP queries to fail every time.
- SIAP load dialogue now has name resolution capability.
- Added SSAP load dialogue. Basic functionality only.
- Added documentation for previously-unadvertised SIAP/SSAP load dialogues.
- Fixed bug which caused registry queries (Cone, SIA, SSA) to fail for Java 1.6.
- Fixed bug which caused blank Contact field in registry query results.
- Can now query registry for more registry services.
- Fits BINTABLE TZERO/TSCAL value reading improvements:
  - Columns with integer TZERO values now read as integers rather than floating point values where possible. This includes unsigned longs ('K'), which were previously represented as doubles with lost precision. Unsigned longs which are too large however ( $>2^{63}$ ) are read as nulls.
  - Byte-valued columns can now be written out by fits-basic output handler as signed byte values (TFORM=B,TZERO=-128) rather than signed shorts (TFORM=I).
  - More comprehensive testing.
  - Fixed bug in calculating value scaled double ('D') values.
  - Fixed bug in typing value for scaled float ('E') arrays.
- Added column selector documentation Appendix A.1.4.
- Fix bug when decoding file:-type URLs containing %xx escapes.
- Fix error reporting bug in registry search.
- Report application name and version in User-Agent header of outgoing HTTP requests.
- Provide aliases -sia and -ssa for -siap and -ssap command line flags.
- Fix SIA and SSA queries to avoid exponential notation in parameters; there is currently no standard concerning numeric representations, but this avoids problems with some servers.
- The fixed length Substring Array Convention for string arrays (TFORMmn=rAw) is now understood for FITS binary tables.
- Fixed -noserv flag.
- Fixed so that attempting to run services under unfavourable firewall conditions does not cause startup to hang.
- Minor SAMP bugs fixed (JSAMP upgraded to 0.3-1).

**Version 3.4-2 (17 July 2009)**

- Plot windows now allow a plot title to be set from the **Configure Axes and Title** window (Appendix A.5.1.2).
- Loading message now displayed in control window table list when table is being loaded from PLASTIC or SAMP, or during startup.
- Incoming and outgoing SAMP load.table messages now use the name parameter, which gives a better chance of a useful human-readable name being attached to transmitted tables.
- Sexagesimal delimiters can now be spaces as alternative to colons in Cone/SSA/SIA load dialogues.
- GAVO load dialogue updated in line with modified database service.
- Work around J2SE mark/reset bug when loading table direct from URL.
- Produce null rather than nonsense results from sky coordinate conversions with unphysical latitudes.
- Produce null rather than questionable results from sexagesimal conversions with mins/secs out of range.
- Startup script improved to provide some cosmetic improvements for Mac users: the

TOPCAT icon is shown in the dock rather than the generic Java icon, and the menu application name is set to "TOPCAT" rather than the classname. Note these don't take effect if running directly using `java -jar`.

- STILTS run using `topcat -stilts` from standalone jar can now correctly report starjava revision.

### **Version 3.4-3 (31 July 2009)**

- Fixed two bugs related to drag and drop: intermittent failure of drop (Mac only?), and application crash when making a drop that fails.

### **Version 3.5 (6 November 2009)**

There are several improvements in this version, many of them to do with improved functionality and usability for loading or acquiring data both from local files and from various Virtual Observatory services.

#### **Multiple-table load**

TOPCAT will now load multiple tables from a single file if multiple tables are present, rather than just the first one. This applies to multi-extension FITS files which contain more than one table HDU, and to VOTable documents which contain more than one TABLE element. It affects the Filestore dialogue, naming files on the command line, and the new Vizier load dialogue.

In previous versions, when you selected a multi-extension FITS file or VOTable from one of the load dialogues or on the command line, only the first table contained in it would be loaded, and any others were silently ignored. Such files often only contain a single table, but not always. Now, if multiple tables are present in the container file, each one will appear as a new table in the tables list. Other table file formats are not currently affected, since there is no mechanism for storing multiple tables in, say, a CSV file.

#### **Multiple Cone, SIA and SSA searches**

In previous versions, it was possible to use the Virtual Observatory Cone Search, Simple Image Access and Simple Spectral Access services to locate catalogue records, images and spectra respectively in a single (cone-shaped) region of the sky, with position and radius entered manually. In this version it is possible to execute such a query for each row of an input table, with position and radius obtained from table data in appropriate columns. This effectively gives a join between a local table and a remote one exposed by a cone-like searchable VO interface.

#### **Vizier load dialogue**

A new Vizier load dialogue has been introduced to make queries directly to CDS's Vizier service. Thanks to Thomas Boch for contributing code for this.

#### **VO menu**

A new menu labelled **VO** (for Virtual Observatory) has been added to the Control Window. This currently contains single and multiple Cone, SIA and SSA queries and the service specific load dialogues, namely the Vizier and Millennium ones. All these features are also available elsewhere in the GUI, but the new menu groups together VO-related functionality for convenience.

#### **Positional search dialogue enhancements**

A number of improvements have been made to the Cone Search, SIA and SSA dialogues; these applies both to the old single-region search dialogues and the new multi-region searches.

- The service URL can now be entered (e.g. cut'n'pasted) directly as an alternative to obtaining it from a registry search. This allows use of unregistered services, as well as making it easy to select the service URL for pasting elsewhere.
- The resource table returned from the registry search can now be sorted by column

contents, by clicking on the selected column header.

- While searching the registry, some indication of the number of services found so far is displayed.
- Searching the registry is less likely to result in an out of memory error.
- The (not very useful, and potentially confusing) registry **Update** button has been withdrawn; its function is now available only from the **Registry** menu instead.
- The metadata contained in tables returned from a successful search is improved.
- The SIA and SSA dialogues now have a selector for choosing the desired image/spectrum format.
- The SIA and SSA dialogues now longer require explicit entry of a size parameter (these protocols define default behaviour for when no non-zero size has been specified).
- The SIA and SSA now automatically query the registry for all appropriate services when first displayed. This is sensible behaviour at the moment, since there is at time of writing a reasonable number of them. A directed search may be given instead at the user's option.
- The Cone Search window provides better instructions on use when first displayed.

A new section Section 6 has been added to the manual providing an introduction to the concepts of VO data access, and a new appendix section Appendix A.9 describes the various single- and multiple-query windows.

### Storage Policy change

The way that TOPCAT stores large tables has changed. In previous versions, table data was held in memory (the default *Storage Policy* was "memory")) unless the `-disk` flag was specified (to use the "disk" storage policy), in which case it was stored in temporary disk files. If you didn't put `-disk` on the command line, it was common to run out of memory when working with large tables.

Now, the default storage policy is "adaptive". This means that relatively small tables are held in memory, and larger ones are stored in temporary files. The effect should be that most of the time, you don't need to specify any special options, and you can work with small or large tables without running out of memory. Using `-disk` is still permissible, and may help in unusual circumstances. Running out of memory is still a possibility, but should be much less common.

The old behaviour can be restored by using the new `-memory` command-line flag.

The only differences you notice from this change should be good ones, however it is slightly experimental. Please contact the author if you suspect it has introduced any problems.

### SAMP-related enhancements

- Spectrum Display activation action now uses SAMP or PLASTIC to display in a suitably compliant external spectrum viewer. The old behaviour was to display the spectrum in an internal SPLAT window, but it didn't work very well.
- The JSAMP library has been upgraded to v1.0, giving some SAMP behaviour changes, enhancements and bugfixes. Control of things like TOPCAT server port and server endpoint hostname are now configurable using `jsamp.*` system properties.
- The `-jsamp` command-line flag has been added for convenience so that the TOPCAT jar file can be used easily as a JSAMP toolkit.

### VOTable-related enhancements

- Namespacing of VOTable documents made more intelligent, and configurable using the `votable.namespacing` system property.
- VOTable 1.2 supported.

- The VOTable 1.2 `xtype` attribute is now used to try to identify columns containing ISO-8601 format dates, which allows them to be treated as numeric values for plotting etc.
- As described above, when loading a VOTable document which contains multiple TABLE elements, all the contained TABLEs are now separately loaded into TOPCAT rather than just the first one, which is what used to happen.

### Loading changes

- The Load Window now has toolbar buttons as well as items in the **DataSources** menu for more convenient access to the various load dialogues.
- The command-line flags corresponding to load dialogues (`-tree`, `-file`, `-sql`, `-cone`, `-gavo`, `-registry`, `-sia`, `-ssa`) have been withdrawn from use; they were probably not widely used, and are of minimal usefulness with the new load window toolbar buttons.
- While a table is loading an entry is now visible in the Control Window table list. Previously it was not always clear whether a slow load was in progress or not.

### Bugfixes and minor enhancements

- There are improvements (and possibly new problems) in the loading dialogue framework. One item is that a failed load less often closes the dialogue.
- HTML table output is now HTML 4.01 by default (includes THEAD and TBODY tags).
- There is a workaround for illegally truncated type declarations in IPAC tables.
- XML, including VOTable, output is handled more carefully; fix VOTable output encoding to be UTF-8, and ensure no illegal XML characters are written.
- Fixed plotting bugs - caused intermittent GUI corruption (e.g. missing toolbar buttons), especially on MacOS.
- Bug fixed in crossmatching output: entries which should have been null were sometimes written as non-null (typically large negative numbers) in FITS and in non-TABLEDATA VOTable output. This affected cells in otherwise non-nullable columns where the entire row was absent. The previous behaviour is not likely to have been mistaken for genuine results.
- `parse*` string->numeric conversion functions now cope with leading or trailing whitespace.
- The list of example queries in the GAVO load dialogue has been extended to match those in its web interface.
- Fixed an obscure bug which could under rare circumstances cause truncation of strings with leading/trailing whitespace read from text-format files.

### Version 3.5-1 (21 December 2009)

This version has one significant new feature:

- Plot windows can now export to PDF as well as other formats.

Other than that, is is mostly performance and usability enhancements and bugfixes:

- Fixed a bug which caused an error when attempting to load multiple files from the command line.
- Improve documentation of 3D windows, including zoom options.
- Fixed a bug which caused an error when attempting to do cone/SIA/SSA search from directly supplied service URL.
- Improve error reporting from multi-cone (etc) window.
- You can now cut and paste from the linear correlation coefficient display in the 2-d plot style editor.
- Faster display of load window and VO menu options.



- Fixed internationalisation bug which could cause cone searches etc to fail in locales that use "," for a decimal point.
- Registry searches are now faster and less likely to run out of memory. Registry-search-based operations (cone search dialogues etc) now run in topcat-lite configuration as well as topcat-full.
- Now warns about non-active resources returned from registry searches (which shouldn't happen).
- Improved helpfulness of error messages in multi-cone etc windows.
- Performance improvements when loading from VOTable documents which contain many small tables, both in general and in the special case of SIA/SSA results. This makes a big difference in some cases.

### Version 3.5-2 (24 March 2010)

- Add memory monitor indicator in control window.
- Considerable performance and scalability improvements to the crossmatching algorithms. For several common regimes, using default settings, memory use has been decreased by a factor of about 5, and CPU time reduced by a factor of about 3.
- Add optional tuning controls (**Tuning Parameters** and **Full Profiling**) to the crossmatching windows. Experimentation with these can lead to significant performance improvements for given matches.
- Fixed a crossmatch bug; it was giving a possibility of suboptimal "Best Match Only" match assignments when pair matching in crowded fields. Crossmatch results thus may differ between earlier versions and this one. Both are reasonable, but the newer behaviour is more correct. In non-crowded fields, there should be no change.
- Add new functions to Arrays: `array` functions for constructing arrays, and new aggregating functions `median` and `quantile`.
- Added **Verbosity** selector in multi-cone window.
- Improved logging of registry searches.
- Adjusted the way that data types are read from JDBC databases. Date, Time and Timestamp type columns will now be converted to Strings which means they can be written to most output formats (previously they were omitted from output tables).
- Match Window documentation extended and somewhat reorganised (they now have their own appendix subsection).
- MacOS icons improved: thanks to Simon Murphy (ANU) for this contribution.
- Memory management adjusted further - default (Adaptive) storage policy now uses direct allocation (`=malloc()`) for intermediate-sized buffers to avoid running out of java heap space.
- Further performance improvement for VOTable documents with very many TABLES.
- Fix bug in code for handling very large mapped FITS files. This was causing fatal read errors in some cases.
- Fix issue when sending some SAMP messages (unhelpful reuse of message tags).
- Fixed minor error when a directory in the file browser got deleted while visible.
- Fix minor scrolling bug in table and column windows.
- Fix histogram save/export bug: the exported table now has the same X bounds as the currently displayed histogram.
- Withdraw embedded SPLAT application. SPLAT could previously be used in some circumstances for in-JVM display of spectra, but this facility didn't work well and was presumably little used. You can still view spectra by configuring an Activation Action to talk to an external spectrum viewer (SPLAT or some other choice) via SAMP or PLASTIC.
- Withdraw SOAP server functionality. This has been deprecated since v2.1. External control is better handled using SAMP (or PLASTIC). If anybody misses this functionality, get in touch and I will consider a similar but less SOAPy alternative.

### Version 3.6 (6 August 2010)

This version contains some new features and usability enhancements, and a number of bugfixes.

#### Load/Save Dialogue improvements:

- There is a new **System Browser** option when loading and saving. This does essentially same job as the old **Filestore Browser** (still available), selecting a file to load or save, but it uses the system-default GUI to do it. For Mac and Windows users, this may present a more familiar and functional interface to the file system (for Linux users, it is probably not an improvement). Both options are available, so you can choose whichever you prefer.
- You can now save multiple tables at once. The Save Window now offers three ways of saving: **Current Table**, **Multiple Tables** and **Session**. Current Table saves the current Apparent Table as before. Multiple Table saves some or all of the loaded Apparent Tables to a container file (typically Multi-Extension FITS or multi-TABLE VOTable). Session saves some or all tables as well, but additionally saves many aspects of the TOPCAT state, such as table subsets, sort order, hidden columns etc.

#### Other significant enhancements:

- Finally - you can delete and rename row subsets in the Subset Window! Apologies to those who've been asking that this has taken so long to implement.
- There are new actions in the Control Window File menu to change the order of loaded tables in the table list (move them up and down). You can use ALT-up/down keys to do the same thing. This may be useful in conjunction with the new session/multi-table save facility.
- Utype metadata items are now visible in the Parameter and Column windows.
- The JSAMP library has been upgraded to version 1.1. This means that the hub will appear in the "System Tray" where possible (when using java 1.6+, and when using a suitable display manager). It also recognises the SAMP\_HUB environment variable for non-standard hub locations.
- There are a couple of improvements in the Vizier load dialogue: it has options to display sub-tables within catalogues and to include or exclude obsolete catalogues.

#### Minor enhancements:

- The current table can now be discarded by hitting the **Delete** key in the Control Window table list.
- SDSS image cutout downloads updated to DR7 (was DR4).
- The unofficial column type "long" is recognised in IPAC format tables.
- An efficiency warning is now issued for large compressed FITS files.
- Format row counts are displayed with group separators (e.g. thousands separated by commas) in some places for better readability.

#### Fixes for bugs and misfeatures:

- Table parameter values in algebraic expressions now evaluate to their current value, rather than the value when the expression was entered.
- Tables in TOPCAT now have private copies of column and table metadata. This means that changing a column name or parameter in one table will not affect other tables. It should also prevent a related bug that allowed the column headings in the table browser to get out of step with those in the columns window and elsewhere.
- Receipt of a SAMP `table.highlight.row` message will no longer cause a similar message to be sent back (and similarly for the PLASTIC `ivo://votable.org/votable/highlightObject` message). The previous behaviour was an unwanted implementation side effect which could sometimes cause problems in interaction with other tools.
- Fixed bug in FITS-plus metadata output (table parameters were getting lost).

- Better behaviour (warn + failover) when attempting to read large files on 32-bit OS or JVM.
- Corrected literature references in Fluxes conversion class documentation (thanks to Mattia Vaccari).
- Fix memory usage issues (unnecessarily large output graphics files, possible out of memory errors) for density plots with large pixel sizes.
- Fixed bug in CSV file parsing that could ignore header row in absence of non-numeric columns.
- Fix some minor bugs with Cancel operation during table save.
- Reinstate help buttons from Save Window toolbar (erroneously removed since version 2.0).
- Make sure that failed load of table with no rows reports as such rather than failing silently.
- Loading tables by typing their name directly in the Load Window now loads multiple tables if present not just the first one, in common with most of the other load dialogue types.
- Upgraded PixTools HEALPix library to 2010/02/09 version. This fixes a bug that could theoretically cause deficient crossmatch results, though I haven't managed to produce such errors.
- Fixed (I hope) "Table no longer loaded" error when sending tables via SAMP.

Finally, from this release TOPCAT requires version 1.5 (a.k.a. 5.0) of the Java J2SE Runtime Environment; it will no longer run on version 1.4, which is now very old. I don't expect this to cause compatibility issues for anyone, but I'm interested to hear if that's not the case.

### Version 3.7 (23 December 2010)

This version contains some significant enhancements and a number of minor improvements and bugfixes.

#### Load Dialogues:

Table loading has had a major overhaul. Load dialogues are no longer *modal*, that is you can now interact with the rest of the application while they are open. This has a number of benefits:

- A slow load doesn't prevent you from doing other things.
- You can be loading several different tables at once, either from the same or from different places.
- You can use the Help system while a load dialogue is open (load dialogues now have the usual Help button like other TOPCAT windows).

It has also enabled some related changes:

- Progress is reported for each table load - you can see how many rows have loaded and, if known, how many there are in total
- Cancelling a load in progress is more straightforward and works more reliably than in earlier versions

A few related bugs and idiosyncracies have been corrected at the same time.

#### Registry Queries:

The registry search interface used in single and multiple VO access windows has been improved:

- When specifying a keyword search for Cone, SIA or SSA services, you can select which registry record fields to match (ShortName, ID, Title, Publisher, Subjects etc). This makes it much easier to get the records you're interested in.
- The Subjects field is now displayed for retrieved records.
- Lists of registry records (i.e. Cone, SIA or SSA services) can now be sent to/received

- from other tools via SAMP (`voresource.loadlist{,.cone,.siap,.ssap}` MTypes).
- Sky coordinates can be received from other tools via SAMP and used to set Cone, SIA or SSA position (`coord.pointAt.sky` MType).

### **BaSTI Load Dialogue:**

A new load dialogue is available for accessing the INAF-OATS BaSTI (Bag of Stellar Tracks and Isochrones) service -- code contributed by Marco Molinaro.

### **Searchable Help:**

The Help browser now has a Search tab as well as the Table of Contents tab, so you can do text searching on the extensive help document.

### **Minor Enhancements:**

- Storage management improvements; removed restriction on large (>2Gb) non-FITS datasets in some circumstances.
- Fix so FITS tables >2Gb can be used in 32-bit mode (though slower than 64-bit).
- All windows have a new Scrollable option in their File menu. It is not generally recommended to use this option, since in general the windows are arranged so that resizing them will resize sensible parts of them, but it may be useful if using some of the larger windows on an unusually small screen.
- FITS files now store table names in EXTNAME (and possibly EXTVAR) header cards.
- Window placement should now behave a bit more like platform norms, rather than sitting in the top left corner.
- When a table is discarded, it is now deselected from any table selector (for instance in a plot, match or concatenation window).
- HTML table output now writes cell contents which look like URLs in HTML <A> tags.
- Basic authorization (`http://user:pass@host/path`) on table URLs handled.
- Logs current version and whether it is up to date with latest release on startup. This behaviour can be controlled by the `[-no]checkversion` flag.
- Belatedly added STFC logo to About window.
- Add recommendation to use 64-bit java in large table tips section.

### **Bug Fixes:**

- Fix bug in registry result table which displayed resource identifier instead of publisher in the Publisher column.
- Fix SAMP table load bug which tripped over "+" characters in URLs.
- Fix regression bug from v3.6 - loading jdbc tables from command line or text entry now works again.
- Fixed file pointer `int` overflow bug in FITS MultiMappedFiles.
- Fixed a couple of bugs relating to VOTable strict mode.
- Possibly fixed a not-well-characterised bug in registry search.

## **Version 3.8 (9 May 2011)**

### **TAP**

A new TAP load dialogue (Appendix A.9.8) supporting data retrieval using the Table Access Protocol (TAP) has been introduced. This provides SQL-like free-form access to any remote databases exposed with the TAP protocol.

### **SAMP changes**

An internal SAMP hub is now by default started when TOPCAT starts up if no existing hub is apparently running, so under normal circumstances the SAMP panel at the bottom right of the Control Window should be active all the time. This means that you don't need to do any explicit preparation to communicate with other SAMP-aware applications. If

you want to return to the previous behaviour (hub only started by explicit user action), use the `-nohub` flag on the command line. You can connect and disconnect with the hub by clicking the connection logo at the right of the SAMP panel.

The JSAMP library is upgraded to JSAMP v1.2, which fixes one or two bugs and supports the experimental Web Profile, though Web Profile support is not switched on by default.

### Other enhancements

- Random Groups HDUs are now tolerated, though not interpreted, within FITS files.
- Add quintuple match option to multi-table match window (someone requested it!).
- You can now re-activate a row in the table viewer window by deselecting and then re-selecting it.
- Expressions in column selectors are now left (and coloured grey) rather than deleted in the case of a syntax error.
- Improve text rendering in Available Functions window.
- GAVO load dialogue now uses fixed-width font for SQL entry.
- Registry search now includes Publisher field by default.

### Bug fixes and workarounds

- JDBC table input handler now effectively downcasts BigInteger/BigDecimal types to Long/Double. The PostgreSQL JDBC driver seems to use the Big\* types routinely for numeric values (which I don't think it used to do).
- Add workaround for J2SE bug #4795134, which could cause errors when reading compressed FITS files.
- Fix FITS character handling bug which could cause corrupted FITS files on output in presence of non-ASCII characters.
- Fix (some) JDBC connection leaks.
- Attempt case-sensitive matching before case-insensitive for column names.

## Version 3.9 (27 October 2011)

### Crossmatching:

- Additional asymmetric match options have been added to the Pair Match window - as well as the old symmetric Best Match option, you can now choose to do a match which identifies the best match for each row of one or the other input table (see Output Rows Selector Box). They correspond to finding the best match in table B for each row in table A, and in crowded fields often provide more intuitive semantics than the previous symmetric `best` option (in non-crowded fields there is generally no difference). This replicates the matching performed by some other tools, including Aladin.
- New Match Algorithms, **2-d Cartesian Ellipses** and **Sky Ellipses**, have been introduced to permit matching of general elliptical, rather than just circular, regions in both planar and sky coordinates. Another, **N-d Cartesian with Errors**, has been added for dealing with per-object errors in Cartesian coordinates (previously per-object errors could only be handled in sky coords).
- Fixed a significant bug in sky crossmatching. If all points in a table were on one side of the RA=0 line, but the error radius extended across that line, matches on the other side could be missed. Matches could also be missed if different tables used different conventional ranges for RA (e.g. -180..180 in one case and 0..360 in another). This fix may in some, but not most, cases result in slower matching than previously.
- Semantics of the **Sky With Errors** match algorithm have changed slightly.

### Notable usability improvements:

- You can now change the column order of a table by dragging the rows up and down in the Columns Window.
- Algebraic functions involving angles are now mostly available using degrees as well as radians. The `Coords` class has been replaced by `CoordsDegrees` and `CoordsRadians` classes providing sky coordinate functions, and a new class `TrigDegrees` provides normal degree-based trigonometric functions alongside the radian-based versions in `Maths`. The functions in the various activation action classes now take degrees and not radians. Some of the old function names have changed to make clear that they use radians and not degrees.
- The **File|Discard Table** action in the Control Window can now remove multiple tables at once if more than one is selected.
- ADQL parsing and syntax highlighting has been added to the TAP load dialogue, thanks to Grégory Mantelet's ADQL library.
- TAP queries are now synchronous by default (the **Synchronous** checkbox is now checked by default).
- The SQL Query Dialogue (finally) has a multi-line entry field for SQL query text.
- Add **To Browser** button to Help Window - this displays the currently displayed help page in your normal web browser.
- New convenience button added to Available Functions window which is a shortcut to display of the expression syntax in the Help browser.

#### Other upgrades and enhancements:

- JSAMP has been upgraded to version 1.3-1. The main change is that the Web Profile is now enabled by default in the SAMP hub that TOPCAT launches. There are also more options in the Hub menus, and it will overwrite any moribund `.samp` file at startup, which *should* eliminate annoying persistent "404 No handler for URL" warnings.
- Add **Delete On Exit** checkbox to Running Jobs tab of TAP load dialogue. Jobs are now created by default with Delete On Exit set true (you can change the default with the Deletion menu).
- You can now run JyStilts from the `topcat-*.jar` jar files in the same way as from `stilts.jar`.
- Add `experimental` `system` `properties` `star.basicauth.user` and `star.basicauth.password`.
- Added the experimental `topcat.exttools` system property to allow custom tools to be added to the main toolbar at runtime.
- Improve resilience of multi-cone operation in the presence of unreliable or inconsistent DAL services.
- Added new constants to expression language `POSITIVE_INFINITY`, `NEGATIVE_INFINITY` and `NaN`.
- Add `-running` flag which loads tables specified on the command line into a existing instance of TOPCAT if one is already running.
- Add new `join` function to Arrays class to combine array elements into a string.
- TAP load dialogue uses modified upload ID as per most recent TAPRegExt draft. This may cause some TAP services incorrectly to appear not to support uploads if they have not made a similar update.

#### Bug fixes:

- Improved some issues (reporting wrong row subset counts and membership) related to changing subset definitions.
- Fix cone search verbosity parameter so that `VERB=3` is not erroneously ignored.
- Fix bug introduced at v3.7 related to labelling loaded files (improper handling and propagation of `LOAD_SOURCE` table parameter).
- A `PARAMref` element with no referent in a `VOTable` no longer causes an uncaught

NullPointerException.

- Fix a small bug related to enabledness of buttons in control window and hub connection.
- Work round an obscure java misfeature which could cause the wrong cell to be edited if a sort interrupts an edit.

### Version 4.0b (28 March 2013)

TOPCAT version 4 includes a complete rewrite of the plotting (see below for details). The new plotting functionality is however experimental and changes may be made to the GUI following user feedback. For this reason, the classic plot windows remain available from their usual toolbar buttons and the new plot windows are currently hidden away in the **Graphics** menu. The "b" designation of this version acknowledges the experimental nature of the visualisation changes, but other new features are considered stable.

#### Logo:

I finally bid an affectionate farewell to the Top Cat Hannah-Barbera cartoon graphic. Never having had any legal right whatsoever to use this logo my conscience has eventually got the better of me and TOPCAT now has an excellent new logo kindly drawn by my friend Phil Hall. It's still a yellow cat.

#### Plotting:

Four new "Layer plot" windows are available from the **Graphics** menu: Plane, Sky, Cube and Sphere layer plots. These feature a new more powerful (and, yes, more complicated) user interface, and are fully documented in Appendix A.4. New functionality includes:

##### New Sky Coordinate Plot

- Choice of projection: Sin (rotatable), Aitoff, Plate Carrée
- Data and view sky coordinate systems selected separately: options are equatorial, galactic, supergalactic, ecliptic
- Sky coordinate grid labelled and visible at all zooms

##### New data plot options

- vectors
- ellipses (with position angle)
- pair, triple, ... data point lines/polygons
- contours
- variable size markers

##### Improved interactive response

- In 2d and sky plots mouse wheel zooms around cursor position
- In 2d and sky plots you can drag the plot around
- In 3d plots right mouse button recentres cube on selected point
- In 3d plots zooming zooms data in the cube rather than enlarging the cube wireframe itself
- Many controls are sliders which update the plot as you slide

##### New plot colouring modes

Density colour coding for all plot types, with density colour map based either on dataset colour or chosen from a fixed set. Flat, transparent and aux colour coding still available as before.

##### Better support for large datasets

Several features have been introduced to provide more meaningful visualisation of large datasets. Improved density-like plots and contours give you better ways to understand plots containing many more points than there are pixels to plot them on.

There is separately some improvement in scalability: up to roughly 10 million points is currently feasible depending on available memory etc, though it depends what you're doing. However, I hope to improve this limit in future.

### Improved axis labelling

- Choice of font size and style
- Option of LaTeX input for non-ASCII characters etc
- Log axes labelled better
- Minor tick option

### Legend options

External or manually positioned internal placement.

### Analytic function plotting in 2D

Plot functions of X or Y coordinate using TOPCAT expression language.

*(Note this plot change list was missing from the v4.0 release and retrospectively added at v4.0-1).*

### Other new capabilities:

- Use MOC footprints to speed up multi-cone searches, including footprint icon display in multicone window.
- Add IPAC output format.
- Add new class KCorrections to the expression language, containing a method for calculating K-corrections following the method of Chilingarian and Zolotukhin.
- VOTable input and output are now supported for version 1.3 of the VOTable standard.
- The version of the VOTable format used for VOTable output can now be selected, by using the system property `votable.version`. Output version is VOTable 1.2 by default. VOTable output no longer includes `schemaLocation` attribute.
- You can now reference tables in multi-extension FITS files by name (EXTNAME or EXTNAME-EXTVER) as an alternative to by HDU index.

### Other enhancements:

- ADQL parsing in the TAP window can now highlight all unknown symbols not just one of them. This is possible by upgrading to the official v1.1 of Grégory Mantelet's ADQL library, which improves some other aspects of ADQL parsing as well.
- Add new function `hypot` ( $=\sqrt{x^2+y^2}$ ) to the Maths class in expression language.
- Add new `split` functions for string splitting to the Strings class in expression language.
- Some changes to the `toString` function: it now works on non-numeric values, gives the right answer for Long integers and character values, and returns a blank value rather than the string "null" or "NaN" for blank inputs.
- Sexagesimal to numeric angle conversion functions now permit the seconds part of the sexagesimal string to be missing.
- Changes to the IPAC format definition are accommodated: the "long"/"l" type, which is apparently now official, no longer generates a warning, and headers may now use minus signs instead of whitespace.
- Fixed SAMP table loads to honour "name" parameter.
- PNG graphics output no longer has transparent background.
- Upgrade JSAMP library to version 1.3-3.
- Work around change in Vizier output so that Vizier catalogue searching works again. This may be a temporary change.
- Update URLs for GAVO Millennium database service at request of Gerard Lemson.
- Add new class Coverage to the expression language containing MOC-related functions (currently, just `inMoc`).



- Add explanatory section to manual on Multi-Object Matches.
- Add `-debug` flag to manage logging messages better.

#### Bug fixes:

- Fix serious and long-standing bug (bad TZERO header, causes subsequent reads to fail) for FITS output of boolean array columns.
- Fix small but genuine sky matching bug. The effect was that near the poles matches near the specified threshold could be missed. The bug was in the PixTools library, fixed at the 2012-07-28 release.
- Fix bug which failed when attempting to read FITS files with complex array columns (`TFORMn=rC/rM`).
- Fix failure when loading very large sequential tables.
- Fixed table concatenation so that column datatype array size etc is consistent with both input tables not just the first one.
- Adjust SQL writer to avoid type error for MySQL.
- Fix plotting bug that might have caused mysterious failures to update the plot. Or it might not.
- Fix ADQL parsing in TAP window so that TAP\_UPLOAD tables are treated correctly.
- Fix bug which could cause truncation of strings in FITS and possibly VOTable output when tables were hand-edited to add strings longer than previously-declared length.
- Fixed bug in multi-table matches (>2 tables) which could result in output rows with columns from only a single table, i.e. not representing an inter-table match.
- Fix bug in HMS sexagesimal formatting: minus sign was omitted from negative angles. Now the output is forced positive.
- Fix minor bug associated with deleting the current row subset (this now causes All to become current).
- Cope with 1-column CSV files.
- Fix (some, though probably not all) possible bug(s) related to running on Java 1.7 on Mac.
- Use the correct form "rows"/"bytes" rather than "row"/"byte" for TAP capability unit values.
- Infinite floating point values are now correctly encoded in VOTable output ("`+Inf`"/"`-Inf`", not "`Infinity`"/"`-Infinity`" as in previous versions).
- Fixed some layout problems which could lead to zero-width text entry fields for RA/Dec.
- Fixed bug when attempting to explode a hidden array column.
- Fix error bar rendering bug which could result in diagonal lines being offset near the edge of plots.
- Fix problem with GUI locking up when plotting outsize histograms.
- Fixed bug in stacked line plot which caused scribbly drawing in the presence of null X axis values.
- Fixed problem with row highlight messages bouncing for ever between applications when the Transmit Row activation action is in use.
- Improve behaviour when deleting a parameter in the Parameter Window.

#### Version 4.0-1 (1 July 2013)

##### New Functionality

- Add read-only support for CDF (NASA Common Data Format) files. Currently, no attempt is made to present time-like values in a human-friendly way, but this may improve in a future release. The SAMP `table.load.cdf` MType is also supported for receiving CDF files.

- Improve handling of HTTP basic authorization. HTTP 401s now pop up a user/password dialogue window, unless the `star.basicauth.*` system properties have been set up.
- Add options to calculate the (scaled or unscaled) Median Absolute Deviation in the Statistics window.

### Minor enhancements

- Now subscribes to `SAMP voresource.loadlist.tap` MType, accepting resource list in TAP load dialogue "Select Service" tab.
- The Columns window now makes a better attempt at displaying non-standard per-column metadata of unusual types, in particular array values.
- Line plotting improved in layer plots (including better dashing, new antialias option).
- Implemented fixes to reduce the chance of users inadvertently overloading external Cone/SIA/SSA services with multicone-like queries. First, fix it so that abandoned queries are properly terminated, rather than continuing to hit the server until completion or JVM shutdown. Second, implement a sensible default maximum value for the **Parallelism** field in the multi-cone (etc) panel (though this may be adjusted with a system property).
- Source code is now managed by git and not subversion. The format of the "Starjava revision" string reported by the `-version` flag and the Help|About menu item has changed accordingly.

### Bug fixes

- Fix CSV regression bug introduced at v4.0b - CSV files now work again with MSDOS-style line breaks.
- TAP example queries now quote table and column names where necessary. This fixes a bug that was particularly evident with the Vizier TAP service for which nearly all table names are not legal ADQL identifiers.
- Quoting behaviour has changed when generating SQL to write to RDBMS tables. This ought to reduce problems related to mixed-case identifiers. However, it is possible that it could lead to unforeseen new anomalies.
- Fix bug with reading session files containing 1-column tables.
- Fixed FITS output bug which could result in badly-formed string-valued header cards (no closing quote).
- Turn off layer plot optimisation that could result in lost precision for double values.
- Fix layer plot bug that broke the plot if both handles of a range slider were dragged to the same position.
- Fixed it so that layer plot axis range settings are reset when the Rescale button is used.
- Fixed a bug that caused an exception when a table in a layer plot was deleted from the application.
- Fixed a bug when zooming way out from a sky plot.

### Version 4.1 (7 March 2014)

In this version, the new plotting windows appear for the first time in the main Control Window toolbar, and are no longer considered experimental. There have been many changes to their user interface and functionality since the last release. The old plotting windows are still available from the Control Window **Graphics** menu, but are now considered somewhat deprecated, and will not be developed or fixed further.

### Navigation changes to the new plotting windows:

- In 2d plots, right-button drag (or ctrl-drag) does a 2-d anisotropic zoom - you can stretch or squash the plot in both directions.

- In 3d plots, right-button drag (or ctrl-drag) does a 2-d anisotropic zoom along the 2 axes most face-on.
- In 3d plots, middle-button drag (or shift-drag) slides the plot along the 2 axes most face-on.
- In 2d plots, the pan/zoom actions (wheel, pan-drag, zoom-drag) can be done outside the plot to affect just one axis.
- Most navigation mouse actions now show visual feedback on the screen.
- A new tab **Navigation** is added to the Axis controls. This lets you configure zoom/pan actions, for instance selecting zoom factor and X/Y/XY zoom options for when mouse wheel is used.
- A hints panel is added below the plot showing current mouse navigation actions.
- Removed zoom in/out buttons in the toolbar of the layer plot windows; there are better ways to zoom.

#### Other changes to the new plotting windows:

- Histogram plot type added. You can also add histograms to a normal plane plot if you want to.
- Layer plot windows now have a progress bar at the bottom for loading data, and optionally for other potentially slow operations like replotting, identifying a point, turning blob-selected points into a subset etc.
- For very large/slow plots, intermediate subsampled plots are optionally displayed while panning/zooming to improve responsiveness.
- Toolbar buttons rearranged; the buttons for adding and removing layer controls are now in the layer control panel not the main plot window toolbar.
- When a pair crossmatch completes successfully, you get the option to plot the results, which shows you exactly how the located matches relate to the input positions (pairs of points are plotted).
- Text label plotting now has crowding configuration options: according to configuration, text labels are only drawn if there aren't too many close together.
- New grid and label colour configuration options for plane and sky plots.
- More colour maps added.
- Antialiasing option added for grid line and label drawing.
- New auto transparency shading mode.
- Add some padding around auto-ranged plot limits.
- Changing table in a plot now tries to retain the coordinate values if they still make sense.
- Axis tick labelling improved, including avoiding overlapping tick labels.
- Sky plot re-ranging improved: new data set now re-ranges unless lock button is set.
- Clip selector added for density and aux colour maps, which means you can now use just part of a colour map, giving you much more control over colouring.
- Sky plot position formatting now honours sexagesimal setting.
- Some minor threading/performance improvements.
- Experimental time plot type added. This plot type is considered experimental and appears in the Graphics menu but not currently on the main toolbar. It can plot time series and spectrograms. May be improved in future versions.
- The control for plotting pairs of points works differently (now has its own layer control not just a Form). The triple-point form (which didn't work properly) is withdrawn for now.

#### Other enhancements:

- A new experimental SAMP message `table.get.stil` is supported, allowing SAMP clients to retrieve tables from TOPCAT.
- URL selector fields in Activation Action window are now editable.
- Add cuboid match algorithm to match windows.

- The Exact matcher now considers scalar numeric values equal if they have the same numeric value; they are no longer required to have the same type.
- Array-valued per-column metadata items are now displayed properly.
- The `toString` function now works for byte and boolean values as well as other data types.
- Replace the (slightly dangerous, also undocumented) Apply Subset option in the table viewer window with a new Highlight Subset option. Also add a new Highlight Subset option in the Subsets window.
- In the old-style histogram plot, exporting it now provides cumulative counts if the visible histogram is currently in cumulative mode.
- Keyboard focus is more helpfully positioned by default in some dialogue windows (it starts off in the first text field).

#### Bugfixes etc:

- Fix a registry access bug related to namespaces. It is now finally possible to query the STSci/NVO/VAO registry, as well as the AstroGrid and Euro-VO registries, from the Cone/SIA/SSA/TAP VO windows. Queries to the STSci registry have not been working since mid-2010 as a consequence of bugs in (mostly) the TOPCAT client and (partially) the STSci registry service.
- Work around bug in MacOS text rendering in plots that painted text in the wrong direction when not horizontal (e.g. letters going backwards when running vertically). The workaround involves defaulting to antialiased text when drawing axis grids for MacOS. As well as getting the text the right way round this does look a bit nicer, but it is also perceptibly slower. If speed is more important than correctness you can turn it off using the **Antialias** checkbox in the **Axes-Grid** tabs of the various plot windows.
- Withdraw **Labels** tab in the Sphere plot, since it didn't do anything.
- Rename the **File** menu in most windows to **Window**, which is much more appropriate.
- Fix bug when concatenating tables whose data cells may have been edited by hand.
- Fix service selection bug (NullPointerException) in multi-SIA/SSA windows.
- Fix bug which prevented access to long integer array elements from expression language.
- Fix bug in Statistics window so that Boolean columns correctly report mean values (true/true+false proportion).
- Fix TST input handler so TST files with fewer than 3 columns can be read.
- Upgrade JSAMP to version 1.3-4 (hub now supports web clients with https URLs).
- Moved text documenting `table.load.cdf` and `table.load.stil` SAMP MTypes from the transmit section (wrong) to the receive section (right) of this document.
- Minor improvements to CDF support.
- Fix some plot threading issues.
- The reported point count is improved in plots.
- Better handling of OutOfMemoryErrors during plotting.
- Better handling of LaTeX errors in axis labelling.

#### Version 4.2 (4 July 2014)

##### Significant new functionality and changes:

- Add the new CDS X-Match Upload Window. In most cases (matching with a table that can be found in VizieR) this window can and should be used instead of the old Multi-Cone window - it is *much* faster. The button for this window replaces the MultiCone button in the main Control Window toolbar. MultiCone is still available from the VO and Joins menus, but is now for most purposes deprecated.
- Windows with a Registry Query Panel can now talk to the registry using either RI1.0

(the only option in previous versions) or RegTAP. RegTAP is generally much better (faster, less bandwidth, more reliable), and is the default choice, but a selector allows you to revert to RI1.0 if preferred.

- The "frame"-style zooming that was used in old-style plots is now available in the new-style plots - drag the mouse using the middle button or shift key.
- The TAP Window button replaces the Concatenate window button on the toolbar. The Concatenate window is still available from the Joins menu. This is not new functionality, but the UI change reflects the growing usefulness of TAP services.

#### Minor improvements:

- Form configuration options are now preserved, rather than reset, when a new table is selected in one of the plot windows.
- Registry search panels now display resource count at the bottom of the resource list.
- Update the endpoint for the NVO/VAO RI1.0 registry (now deprecated in favour of RegTAP), on advice from Theresa Dower.
- Fixed TAP load dialogue so that by default it stays open after a table is loaded (the **Stay Open** option is set true by default).
- Improve TAP window Running Jobs display. The Job URL is now always shown even when the other job details scroll off screen, and the details scroll to the top not the bottom when a job is selected or refreshed.
- Upgraded Gregory Mantelet's ADQL library to version 1.2. Better ADQL parsing.
- Modify the way antialiasing is controlled in plot windows. The Grid antialiasing option, where present, now just covers the actual grid lines. Text antialiasing is now selected with a new "Antialias" option in the (now possibly misnamed) "Text Syntax" item in Font configuration. This now affects the Aux axis and Legend labelling as well as the grid labelling.
- Added new graphics output format `png-transp` to generate PNG files with transparent backgrounds.

#### Bug fixes:

- Set antialiasing on by default for OSX - get it right this time.
- Fixed some bugs related to TAP table uploads.
- Fix out-by-one error in reporting of Row of min/max values in Statistics window.
- Add missing geometric reserved words to ADQL reserved word list. Has a slight effect on ADQL validation in TAP window.
- No longer disable "View URL As ..." options in Activation window when no suitable columns are present.
- Fixed bug in which cube plot navigation help button gave you sphere plot help and vice versa.

### Version 4.2-1 (13 November 2014)

#### New functionality:

- Add (experimental) read-only support for Gaia/DPAC GBIN format.
- Add special variables `$ncol` and `$nrow` to the expression language to refer to the column and row counts in the underlying table. The special variable `index` is also deprecated in favour of `$index` or `$0`.
- Support viewing tables up to  $2^{31}$  (2 billion) rows in the table viewer window. The previous limit was  $2^{27}$  (134 million) rows. This does not necessarily mean that using TOPCAT on GigaRow tables is a good idea!

#### Minor improvements and bugfixes:

- Change the default RegTAP registry endpoint to `http://reg.g-v.o.org/tap`, which

should have good reliability, since it can point to different RegTAP services as required.

- Add some more colour maps for aux/density shading.
- Fix regression bug (introduced at v4.2) which caused authentication to fail when using the SQL Query load dialogue.
- Attempting to write FITS tables with >999 columns now fails with a more helpful error message.
- Somewhat improved Unicode handling in VOTables. If you load a VOTable with columns marked `datatype="unicodeChar"` and save it again, the columns now remain `unicodeChar` instead of getting squashed to type `char`. Some lurking Unicode-related issues remain.
- Fix time plot to work for ISO-8601 times. This change also breaks use of ISO-8601 times in non-time plots, but that didn't work particularly well before. Please complain if this change of functionality causes a problem.
- Fix problem with activation action image viewer selection box.
- Update JSAMP to v1.3.5.

## Version 4.2-2 (6 February 2015)

### Plotting enhancements:

- New Linear Fitting form for Plane plot. This reinstates behaviour available from old plot windows, but it provides some more options, like point weighting.
- New SizeXY form, which allows plotting (optionally autoscaled) markers with horizontal and vertical extents independently determined by input data.
- New Frame fixed control in plot windows with two tabs, **Size** and **Title**. Size allows explicit setting of output plot dimensions and alignment in pixels - can be useful when exporting plots. Title lets you add a text title at the top of the plot.
- New **Save/Import** options in Histogram window allow you to export histogram bin data as a new table (this reinstates behaviour available from old plot windows).
- More flexibility when assigning colour maps, in the Aux Axis control, Density shading mode, and Spectrogram layer control. These now provide a **Scaling** option allowing use of Square and Sqrt mappings as well as Linear and Logarithmic, and a **Quantise** slider allowing quantisation of the colour map to discrete levels.
- **Rescale Y** button added to Histogram window.
- Histogram normalisation option adjusted so that total area under bars, rather than total height of bars, is fixed.
- Tick Crowding configuration option added to Aux Axis control Ramp tab. Also changed default for Aux axis to use fewer ticks.
- Some "dart" options (fixed-base open or filled triangles) added for plotting vectors.
- Some "triangle" options (variable-base open or filled triangles) added for plotting ellipses.
- Max Marker Size option replaced with with Autoscaling option in plot Size form. You can now optionally turn off autoscaling and specify marker size in pixels instead.
- Every slider control in plot windows now has a little button on the right that resets it to its default value.
- The text labels for configuration options and coordinate entry fields in the plot windows now provide a tooltip of the form "`name=value`". These strings give the STILTS parameter assignment corresponding to the current setting.

### Other enhancements

- New Column Classification Window lets you add mutually exclusive subsets based on column contents.

**FITS I/O:**

- Reworked part of the FITS input implementation, in particular adjusting the way memory mapping is done to reduce resource requirements on some platforms. If you notice any difference, it should be reduced virtual and perhaps resident memory usage, and some (~10%?) performance improvements, when reading large FITS/colfits files. However, please report anything that appears to be working worse than before.
- Colfits files can now be accessed from streams, not just uncompressed disk files (though that's not necessarily a good idea).

**MacOS packaging**

Some changes have been made to the MacOS DMG file/application bundle. I believe these are at worst harmless, but pre-release testing of platform-specific features is patchy, so problems are possible - please report if you find any, or if the following items don't seem to be working.

- OSX application bundle adjusted so that the DMG file *should* work on any MacOS with either legacy Apple Java 6 or Oracle/OpenJDK Java 7+ installed.
- OSX application bundle adjusted (I think) so that application working directory is user's home directory.

**Bugfixes and workarounds:**

- Fixed a query bug (missing REQUEST=queryData parameter) in the Multiple SSA Query window. This long-standing bug would have stopped this window working at all with well-behaved SSA services.
- Fixed error in fits-var output (PCOUNT header card did not include block alignment gap).
- Graphics coordinates are now calculated in floating point rather than as integers. This fixes problems that could cause scaled vectors, ellipses etc to be drawn with shapes or orientations badly wrong due to rounding errors. It also improves plotting of analytic functions, especially to vector contexts (PDF/EPS).
- Fix some problems to do with zooming to very large/small plot axis ranges.
- Hide error bars (etc) that would extend to negative values on logarithmic axes; previously they were being drawn in anomalous places.
- Fix Aux axis positioning for 3D plots so that the numeric labels don't get snipped off at top and bottom.
- Add a hack that allows LDAC FITS tables to be treated sensibly in auto-format-detection mode.
- Make VOTable handling more robust against unknown (illegal) datatypes.

**Version 4.2-3 (14 April 2015)****Histogram enhancements:**

- New Kernel Density Estimate plotting option variants added in the Histogram and Plane plot windows: KDE, KNN and Densogram.
- More flexible normalisation options for histogram/KDE plots: normalise by area, maximum bar height, or total bar height.
- The Bars fixed control is now relabelled Bins, and has separate tabs for Histogram, KDE and General controls.
- The existing Bin Size slider, and new kernel Smoothing sliders, now display current width in data units as the slider is moved.
- More bar styles are available for Histogram plots; **Semi Filled** (the new default) and **Semi Steps** options make it easier to understand multiple overplotted data sets. Note Semi Steps does not currently export very nicely to PDF/EPS.

- Sketch mode is switched off by default.

#### Minor improvements:

- More slider controls have options allowing explicit entry of values from the keyboard (colour map quantisation, marker/vector/ellipse scaling values).
- More slider controls have text value readout displays alongside the slider (colour map quantisation, marker size/errorbar scaling values).
- Column data read in as unsigned bytes will now be written out as unsigned bytes where the output format permits; previously they were forced to 16-bit signed integers. This affects FITS, VOTable and CDF I/O handlers.
- The Subset Styles panel in table data layer controls now has a **Visible** checkbox indicating whether the subset undergoing subset-specific configuration is currently visible in the plot.
- Be less strict about recognising colfits files (tolerate implicit TDIMn headers).
- Tables specified on the command line may now use syntax that reads data from a unix pipeline.
- OSX application bundle fixed to run application in user's home directory (broken in previous version).

#### Bug fixes:

- Fix minor ADQL parse failure when using TAP tables named without explicit schema prefix by service.
- Fix regression bug (from v4.2-2) that caused problems in Histogram and Spectrogram layers: *New subset from visible* action wasn't working, and the displayed point count was always zero.
- Work round FITS read bug that could cause problems for VOTables using inline FITS serialization, and possibly elsewhere.
- Fix bug that caused trouble when auto-ranging a plot with a single sky position.
- Catch and report the error properly if the user enters a bad explicit axis range (max value  $\leq$  min value).

### Version 4.3 (17 August 2015)

#### TAP client enhancements:

The TAP client GUI and implementation in the TAP Load Window have been much improved. As well as being generally more capable and easy to use, the GUI is now quite usable for services providing very large datasets, for instance TAPVizieR (~30,000 tables). In detail:

- TAP service discovery is now done by default by requesting search terms relating to the *tables*, rather than the *services*, that you're interested in. The old (by service) GUI is still available in a different tab.
- The TAP metadata display has been reorganised; schemas and tables are displayed in a searchable tree on the left of the panel, with details of the currently highlighted schema/table in tabs on the right. Organisation of tables at the schema level is now visible. Much more information about the schemas and tables is available, and it's displayed better.
- The Columns and Foreign Keys metadata display tables can now be sorted by clicking on column headers.
- Much more information about the TAP service is displayed, including a full service description and listing of available geometry functions, user-defined functions and data models (if the service/registry provides this information).
- For large services, metadata acquisition is now done as required on a table-by-table basis, rather than attempting to download all the metadata up front. This works much better for services with very many tables.



- The query entry panel has many new features, including multiple query entry tabs, full query text Undo/Redo, and limited support for pasting table and column names that have been selected in the GUI.
- A new **Quick Look** mode has been introduced alongside the Synchronous and Asynchronous modes, to execute a query whose result is displayed but not loaded into TOPCAT.
- The **Examples** menu has been extended to include various sub-menus with more ADQL query examples appropriate to the current service, including examples provided by the service itself for services supporting the (somewhat) standard `/examples` endpoint.
- The Examples menu GUI has been modified to facilitate browsing, and a button is provided to link to web pages describing the selected example in more detail.
- The **TAP** menu has new options for configuring **Metadata Acquisition**, **Response Format**, **Upload Format**, **Service Discovery**, and **HTTP-level compression**. The old Deletion menu has been moved under this TAP menu as **Job Deletion**.
- Where supported by the service, the BINARY2 VOTable serialization is used by default to retrieve TAP results, which should reduce bandwidth and (marginally) improve handling of null values.
- For upload queries, tables are now by default uploaded to the service using the TABLEDATA rather than BINARY VOTable serialization. BINARY ought to work, but at least some services (e.g. CADC) currently fail with BINARY uploads. The default can be changed using the **TAP|Upload Format** menu item.
- There is a change in the way that TAP service metadata is interpreted: table/column names are now expected to be delimited if required. An attempt is made to fix up names that not delimited when they obviously should be. May cause some ADQL examples to work incorrectly in the presence of non-compliant service table metadata reports.
- The ADQL syntax highlighting is now sensitive to User-Defined Functions declared by the service (if any are declared), so that misuse of UDFs or use of non-existent UDFs can be signalled. This is thanks to an upgrade to Grégory Mantelet's ADQL parser library.
- Improve error reporting in the face of non-VOTable TAP error responses. In many cases this makes it much easier to see what's going wrong with a TAP query.
- As a diagnostic tool, when making TAP queries, a log message giving a roughly equivalent `curl(1)` command is now issued at the CONFIG level (visible using flags `-verbose -verbose`).

### Bugfixes:

- Fix a serious bug in processing of FITS bit vector (`TFORMn='rX'`) columns. Values read from these columns are presented as a `boolean[]` array. In all previous versions of STIL the bits have appeared in that array in the wrong sequence (LSB..MSB per byte rather than the other way round). Apologies to anyone who may have got incorrect science results from this error in the past, and thanks to Paul Price for helping to diagnose it.
- Fix a less serious bug with `TFORMn='rX'` processing; attempting to read a single-element bit vector column (`TFORMn=1X` or `X`) previously resulted in an error making the file unreadable. Values read from such columns are now presented as Boolean scalars.
- Fix a VOTable reading bug relating to bit vector columns (`datatype="bit"`) appearing in BINARY/BINARY2 serializations. This one was more obvious, it would usually generate an error when attempting to read the file.
- Fix serious bug in time conversion for CDF TIME\_TT2000 data types. This was giving completely wrong (deep future) labels when using TIME\_TT2000 columns in time plots.
- Fix sky plot vector bug: delta latitude and delta longitude values were being used the

wrong way round.

- Upgrade JEL to v2.0.2. Fixes problem with evaluating void-typed expressions (especially user-supplied activation actions), and possibly some other obscure bugs.
- Prevent submission of multi-cone/SIA queries for which the Search Radius parameter is blank.

#### Minor enhancements and behaviour changes:

- Added some functions to the Arrays class that return array-valued results from array-valued parameters: add, subtract, multiply, divide, reciprocal, condition.
- HTTP-level compression is now requested (Accept-Encoding gzip request headers included) in most places where it makes sense. This should substantially reduce required bandwidth in communications with HTTP services that support this.
- When clicking on the headers of columns to sort the rows in displayed JTables that show registry services and Vizier catalogues, clicks now cycle between three options (up, down, none) rather than just two (up, down).
- Minor improvement to version reporting (reports java specification version, no longer issues warning for absent revision string).
- Update JCDF library to v1.1 (minor changes to do with leap seconds).
- Now only runs in PLASTIC mode if the `-plastic` flag is set explicitly (otherwise runs in the default SAMP mode). Previously it ran in PLASTIC mode if a `.plastic` lockfile was present.
- Activation action summary messages are now written via an INFO log message rather than just printing to standard output. To see them you now need to run the application with the `-verbose` flag. The various `exec` functions however do now write their output direct to stdout/stderr rather than including them in a return value.
- Custom Code activation action invocation behaviour has subtly changed when one of the values in the expression is from an integer or boolean column, and that column has a blank value. In that case, it's not possible to represent the column value properly to the function (since integer/boolean values can't be null). Previously, the function was invoked with a zero/false value where the null would have been. Now, the function is not invoked at all. Note this only applies to Activation Actions (which are expected to have side-effects); expressions evaluated as column values etc are, as previously, executed but the result is ignored.
- Modify the heuristics that determine whether the first row of a CSV file is a header.

#### Version 4.3-1 (22 October 2015)

##### New features:

- A new **Hints** tab is added to the TAP load dialogue, with useful tips to jog your memory about ADQL syntax.
- The expression language has a new way of referring to a column; if you use the form `"Object$ <column-id>"` you get the value as an Object not a primitive. This is a special-interest measure for user-defined activation functions that need to see null numeric values.
- Adjustments to the GBIN input handler: avoid descending into Class-typed members of gbin list objects, and add logging for object->column translations.

##### Bugfixes

- Fix error reporting bug when a non-VOTable response is received from a TAP service.
- Fix TAP service location bug that caused an error if searching in By Tables tab with only Service field selected.
- Fix bug in ADQL parsing that failed to recognise column names that are delimited identifiers when qualified by table names.

- Bugfix updates to ADQL library: disallow use of spurious schema names for tables with no schema name; disallow table-qualified names in ORDER/GROUP BY clauses.
- Upgrade to JCDF v1.2 - fixes a bug when reading large (multi-Gb) CDF files.

### Version 4.3-2 (27 October 2015)

#### Crossmatching bug fix

Fix a long-standing crossmatch bug relating to range restriction during pre-processing. This could have caused missed associations (but not false positives) near the edge of coverage regions when using per-row errors, if the scale of the errors differed (especially differed significantly) between the matched tables. It affected match algorithms **Sky with Errors**, **<n>-d Cartesian with Errors**, **2d\_ellipse** and **Sky Ellipses** only. Thanks to Grant Kennedy (IoA) for reporting this bug.

#### Density plots

Some more options for making weighted density plots have been added. Since v4.0 the Density shading mode has let you see the density of plotted points, but this lacked some features. Three new ways to do density plots are added; these all give you the option of weighting by an additional coordinate (like the Aux mode), choosing the combination method (mean, median, sum, max, ...), and displaying the quantitative value-colour mapping on the shared colour ramp (previously aux axis) beside the plot. The new density plots are:

- Weighted shading mode, using shaped marker kernels on the screen pixel grid, available for all plot types
- SkyDensity form, using HEALPix bins on the celestial sphere, for the Sky plot
- Density form, using square N\*N screen pixel bins, for the Plane plot

The details are somewhat experimental and may undergo some adjustments in future releases (feedback welcome).

#### Colour maps

There are various changes affecting selection and display of colour maps used for density and aux axis shading:

- The default colour map for Aux mode, and other layers using the shared colour map, is no longer Rainbow! It's Inferno. Rainbow colour maps are much hated by visualisation experts. Of course you can still choose Rainbow if you like.
- Add some new colour maps: *Viridis*, *Inferno*, *Magma* and *Plasma* from Matplotlib 1.5, the *SRON* rainbow variant developed by Paul Tol, some diverging maps (*HotCold*, *RdBu*, *PiYG*, *BrBG*) and a qualitative constant chroma/luminance map *HueCL*.
- The options for Density and Aux shading are now mostly the same as each other except where there's good reason to differ. Previously they were different in haphazard ways. The order of the options is now a bit more rational too (similar maps are grouped together, more common ones near the start).
- An attempt is made to give the default form of each colour map a sensible name, without leading minus signs.
- Fix it so that the whole range of each map is distinguishable from white. This is a good idea when you're plotting symbols on a white background, which is common in topcat. Perhaps there are cases it's not such a good idea; if you think so, complain and I may change it back.
- Try to fix it so that all the colour maps go in the same direction (light->dark) where applicable.
- Throw out a couple of particularly useless colour maps.
- Colour map ramp display is now different for non-absolute maps; their effect is shown on a selection of base colours, not just for one base colour.

### Minor items

- Try harder to identify epoch columns (suitable for time plot), in particular look for VOTable `xtype` of JD or MJD, and `units` of year.
- Add some functions to the Tilings class to do with solid angles (`healpixSqdeg`, `healpixSteradians`, `steradiansToSqdeg`, `sqdegToSteradians`, `SQDEG`).
- Tweaked `topcat` startup script to work better if it's invoked from a symbolic link.
- Fix the description in SUN/253 document of Sky ellipse and vector plot forms, previously they were not distinguished from the Plane variants of those forms.
- Fix plot bug; titles were painted in white for pixel export formats.
- Rationalise plot report logging. Some more diagnostic information about plots is now logged at the INFO level (visible if `topcat` is run with the `-verbose` flag).

### Version 4.3-3 (10 June 2016)

#### Expression Language

The JEL library underlying the expression language parser has been upgraded to v2.1.1 (thanks to Konstantin Metlov), now supporting variable-length argument lists among other things. This allows the following improvements:

- New functions that support any number of arguments are provided: `array`, `intArray`, `stringArray` in class `Arrays`; `concat`, `join` in class `Strings`; `sum`, `mean`, `variance`, `stdev`, `min`, `max`, `median`, `countTrue` in new class `Lists`; and `exec` in activation class `System`.
- Some old lists of similarly-named functions with fixed numbers of arguments have been replaced by single functions that take an arbitrary number of arguments (e.g. `array(x1)`, `array(x1,x2)`, ... `array(x1,x2,x3,x4,x5,x6,x7,x8)` replaced by `array(values...)`).
- The 2-argument `min/max` functions in class `Arithmetic` have been renamed `minNaN/maxNaN` to avoid confusion, but in most cases existing expressions involving `min/max` will work as before.
- Some functions that used to require string arguments will now auto-convert numeric types (e.g. `exec("do_stuff.py",toString(RA),toString(DEC))` can now be written `exec("do_stuff.py",RA,DEC)`).
- You can now implement user-defined functions with variable numbers of arguments.
- Writing large ( $\geq 2^{31}$ ) literal integers used to fail with an inscrutable error message. Now the message tells you to append the "L" character.

#### Time plot

The Time Plot window (available from the Control window **Graphics** menu) has been enhanced so that it can make multi-zone plots - multiple plots stacked vertically that share the same horizontal (time) axis but have independent vertical axes. The time plot itself and this multi-zone feature are currently experimental; in future versions they may be improved or changed, and the multi-zone feature may be extended to other plot types. Some other changes have gone along with this:

- Plotted points (Mark form) in the time plot now supports shading modes as for other plot types.
- The spectrogram layer now uses the (per plot, or more precisely per-zone) Aux colour map rather than a colour map specific to the layer. This means that the colour ramp is now displayed alongside the plot, so you can see quantitatively what colour corresponds to what value. It also means that to adjust the colour map, you now need to use the Aux fixed control rather than controls in the spectrogram's **Style** tab.
- The plot window tries to guess the time coordinate and fill it in automatically.
- The Function layer control now works in the experimental Time plot, rather than throwing an error. However, it doesn't work very well, since the time coordinate is in

unix seconds rather than something more user-friendly.

- Fixed a serious bug in ISO-8601 axis labelling in the time plot. In some cases axis labels were being drawn at positions badly different from the correct position.

### Miscellaneous enhancements and changes

- This and subsequent releases target **Java SE 6**, so will no longer run under the (now very ancient) Java 5 runtime.
- Add Fill plotting mode for Plane and Time plots.
- The TAP window now uses blocking HTTP requests rather than repeated polls to wait for asynchronous TAP job completion from services that declare themselves UWS 1.1 compliant.
- Improve sky plot border painting.
- Clean up noisy Cubehelix colour map.
- New function `countTrue` in class `Arrays`.
- VERbose selector added to Cone Search dialogue.
- UType matching is now case-insensitive when identifying likely columns in Activation window.
- Replace Opaque Limit with Transparency config option for plane and sky density plots.
- Update list of VizieR mirrors.
- Changed implementation of GIF exporter for plots, from Acme to ImageIO. Shouldn't be any noticeable difference.

### Bug fixes

- Fix the sky join syntax example in the TAP window Hints panel, which was incorrect (comma and parenthesis misplaced).
- Fix bug in cumulative histogram calculation.
- Fix read failure for FITS files with non-blank TDIM for zero-length columns.
- Fix bugs that led to timezone-dependent results when reading ISO-8601 or decimal year time columns.
- Fix numeric field truncation bug in LaTeX table output.
- Fix bug in which expression metadata was sometimes not revealed in Columns Window display following addition of a new synthetic column (regression bug since v4.0b).
- Don't use white as one of the auto-selected plotting colours, since it's likely to be invisible (this was a regression bug introduced at v4.3-2).

## Version 4.3-4 (13 September 2016)

### Visualisation and HEALPix

- Add options (from layer control Report panel or plot window Export menu) to export the HEALPix density map calculated by a SkyDensity plot; either save it to disk or import it as a new table.
- New Healpix layer control in the sky plot lets you plot tables that represent healpix maps (e.g. as exported from `SkyDensity` plot).
- Add some HEALPix-related functions to the Tilings class: conversions from pixel index to sky position and conversions between ring and nested schemes.
- Colour ramp **Clip** control (e.g. for the Aux Axis) now has a new **Default** checkbox. When checked, a default clip is applied, which avoids very light colours, and when unchecked you can use the double-slider over the whole range. Previously, the double-slider range was pre-clipped, so very light colours were not available.
- Improve performance of some plot types. In weighted shading mode you can now adjust the colour map without appreciable delay. SkyDensity plots are somewhat improved.

### TAP/VO enhancements

- The TAP window now honours the (TAP 1.1) `column_index` column if it is present in the `TAP_SCHEMA.columns` table. This means that columns can be displayed in a predefined, rather than random, order in the column metadata display tab if the service supports that.
- The TAP window also now acquires and displays non-standard, as well as standard, table and column metadata items if provided by the TAP service. These come from non-standard columns in the `TAP_SCHEMA.tables` and `.columns` tables, or from attributes in custom namespaces in the TableSet document, depending on metadata acquisition policy. Services can use this if they wish to report things like table row counts or column data limits.
- Add VOSI 1.1-style `detail-sensitive` `/tables` endpoint query options to the **TAP|Metadata Acquisition** menu in the TAP window.

### Bug fixes and minor enhancements

- The GBIN input handler can now pick up more metadata from the classpath. For suitable tables, metadata included in datamodel classes if present can be interrogated to provide table and column descriptions and UCDs. There are still some deficiencies of this functionality (no column order, utypes and units missing, large file "temp.xml" written to current directory) dependent on issues in the upstream Gaia libraries and ICD.
- Fix bug that caused read failures for large (>0.5Gb) FITS files outside of the current directory on 32-bit JVMs. This was a regression bug since v4.2-2.
- Fix long-standing bug that failed to release file descriptors when opening FITS tables (could cause an error if very many FITS files were opened).
- Improve the Concatenation Window. You can now use algebraic expressions for the appended table columns, and changes to the appended table no longer reset the window state.
- Fix minor border-painting bug in sky plot window (sky outline painted on GUI components).
- Combiner option `stdev` replaces `variance` in density, sky density and weighted plot layers.

### Version 4.3-5 (23 September 2016)

- Add **Lock Aux Range** action to plot window toolbars and Aux Axis Control **Range** tab. This lets you prevent the display from updating the Aux colour ramp range during navigation. It also affects some dynamic Auto-Scale range calculations in a similar way.
- Fix aux ranging bug in some plot types (SkyDensity, Healpix, Density, Weighted) that used an aux data range too small to show colour variations when Aux scaling was logarithmic and negative data values were present.
- Add Registry menu to TAP load window, allowing list of RegTAP registries to be updated (as for other VO windows).
- Improved performance (better memory use, faster) of some plot types (SkyDensity, Healpix, Density, Weighted).