



xmmselect

January 12, 2017

Abstract

GUI meta-task controlling the selection-extraction-viewing cycle

1 Instruments/Modes

not applicable

2 Use

pipeline processing	no
interactive analysis	yes

3 Description

3.1 Introduction

xmmselect is a GUI (Graphical User Interface) meta-task which allows the SAS user to control the generation of final products, i.e., filtered event lists, images, histograms, OGIP spectra/rate curves, Auxiliary Response Files (ARFs), and Redistributuion Matrix Files (RMFs) from an initial event list data set in an interactive manner. It is intended to be used quite extensively in a typical analysis session in which the user will want to start from a calibrated event list as created by the automatic pipeline processing and generate OGIP-compliant products eligible for the final stages of interactive analysis, such as, spectral fitting with Xspec, and timing analysis with Xronos.

In this process one wants to enter a loop of stepwise refinements of the products in question:

1. Filter the event list under the control of a boolean selection expression
2. Accumulate a product from the events which passed the filters
3. Visualize the product
4. Correlate features in the products with made selections and refine the selection



5. Go back to step 1.

xmmselect provides functionality to perform this type of processing in a convenient manner.

3.2 General overview

3.2.1 Used tasks

xmmselect makes use of the following other SAS tasks to perform the above listed functions:

functions	SAS task used
select events	evselect
extract images, rate curves	evselect
extract spectra	evselect + backscale
generate ARF/RMF	eregionanalyse + especget
display histograms, spectra and rate curves	dsplot
display images	imgdisplay/dsplot
add new columns	tabcalc, phasecalc
delete columns	dsrcm

Users unfamiliar with these packages are advised to consult the relevant documentation prior to reading any further.

3.2.2 Task invocation and GUI initialization

xmmselect must be invoked with the name of a table in an existing data set via the task parameter **table**. This specified table is first inspected for names, types and numerical limits of existing columns and a list of corresponding pushable buttons with associated check buttons is displayed within the main window (see Sect. 3.3.4).

During the initial scan it is also checked whether or not the table comprises DSS (Data SubSpace) information, i.e., if it has been filtered in a previous **xmmselect** or **evselect** run. The DSS formalism is a means to store the selection history of a table in a systematic and easily accessible form. For more details on this please consult the documentation of package **dsslib**.

If DSS information has been found the following dialog box is presented right after the display of the main window:

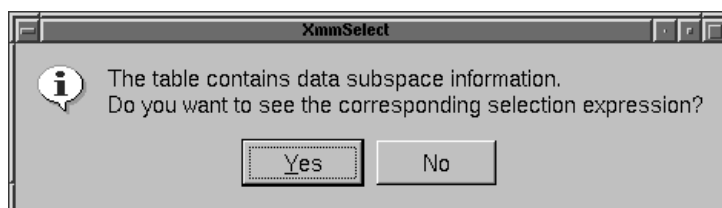


Figure: Dialog signaling presence of DSS information in input table



Ending the dialog with **Yes** will convert the DSS specification into a selection expression and present this in an editable field in a new dedicated dialog:

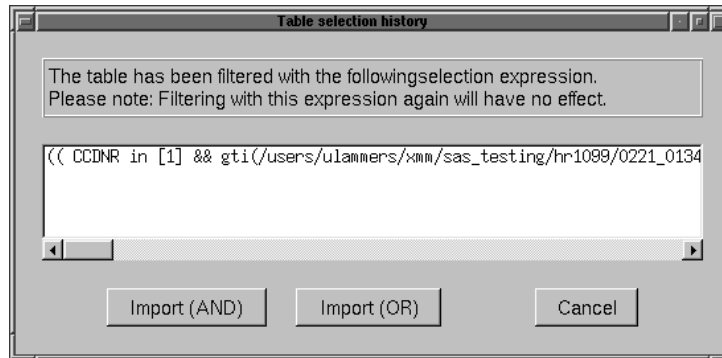


Figure: Dialog to import expression

From there it can be imported into the selection expression widget (see Sect. 3.3.2) of the main window with the buttons **Import (AND)** or **Import (OR)**. The difference between the two import methods is not relevant in this context and will be described later (see Sect. 3.3.2). Pushing the **Cancel** button will end the dialog disregarding the DSS specification. *Please note:* If DSS information is present it is very likely that the table has been filtered with the corresponding expression *already*. Importing the expression and re-doing the filtering will, thus, have no effect. In this regard the DSS import is to be regarded as an informative aid for inquiring the table's selection history.

3.2.3 Column selections

Four types of scalar columns are distinguished: *integer*, *real*, *string*, and *boolean*. Vector columns of any kind (fixed or variable length) are ignored. The extraction of images, histograms, spectra, and rate curves requires the specification of one or two numerical columns as detailed in the following table:

type of product	number of columns (axes in product)	type of columns
image	2	integer or real
histogram	1	integer or real
OGIP spectrum	1	integer
OGIP rate curve	1	real

Columns can be selected for usage in a product extraction if their type is of appropriate type according to the above table. The two columns with check buttons to the left of the push-buttons serve this purpose. The left column refers to the image and the right one to the histogram, spectrum and rate curve generation. Note: For the latter only one table column can be selected at a time while for the image extraction two columns need to be specified.

3.2.4 Selection expression

The selection of events prior to product extraction is carried out by the first stage of **evselect**. This processing step is driven by a boolean expression, i.e., events for which the expression evaluates to “true” are included in the subsequent product generation. The selection expression to be applied to the data is



displayed in the **xmmselect** main window and is editable at all times. In addition, several GUI elements facilitate the construction of complex selection expression in a convenient manner (see below for more details).

3.2.5 Product generation and visualization

Accumulation of a specific product is triggered by pushing one of seven buttons (see Sect. 3.3.6) at the bottom of the main window. This transparently generates a specific product data set with a default name which is subsequently displayed with either **imgdisplay** for images or **dsplot** for histograms, spectra, and rate curves. No display tool is by default invoked on filtered tables and when you click on **Fix Expression** button.

3.2.6 Selections from displayed products

Once a product is displayed the user can interactively define one (spectra/histograms and rate curves) or two-dimensional (images) regions with the aid of the mouse and make **xmmselect** transform these into a boolean selection expression for a subsequent new product extraction.

3.2.6.1 Spatial - 2D regions from imgdisplay or dsplot

imgdisplay and **dsplot** support the definition of one or multiple regions on displayed images (please consult the respective documentation if you are not familiar with this feature). Region specifications can be propagated to **xmmselect** by pushing a button (see Sect. 3.3.5) in the main window. **xmmselect** shall read the regions from **imgdisplay/dsplot** and translate this into a corresponding selection expression which will be appended to the current expression visible in the entry field (see Sect. 3.3.2) at the top of the main window. The selection stage of a subsequent product extraction will then merely consider events which lie inside/outside any of the specified “include”/“exclude” regions.

Please note: Region markers with the color *red* will be ignored in this process. This is useful in cases where one merely wants to import a selected subset of all defined region markers. Please consult the Ds9 documentation about how modify the properties of region markers.

3.2.6.2 Others - 1D regions from dsplot

dsplot supports the definition of one or multiple regions on displayed data plots (please consult the respective documentation if you are not familiar with this feature). Region specifications can be propagated to **xmmselect** by pushing a button (see Sect. 3.3.5) in the main window. **xmmselect** shall read the regions from **dsplot** and translate this into a corresponding boolean selection expression which will then be handled in the same manner as described above for 2D regions. Please note: The Grace region marker types “Above”/“Below” line do *not* translate into expressions which select events that constitute the areas above/below the marker line. The section between the two little arrows on the marker rather define an interval on the abscissa. Also, the region types “Inside Polygon”/“Outside Polygon” can not be converted into a one-dimensional selection expression. Attempts to do so will result in a corresponding error condition.



3.3 Detailed layout of GUI

If `xmmselect` is invoked it comes up with a main window as illustrated in the following figure 3.3:

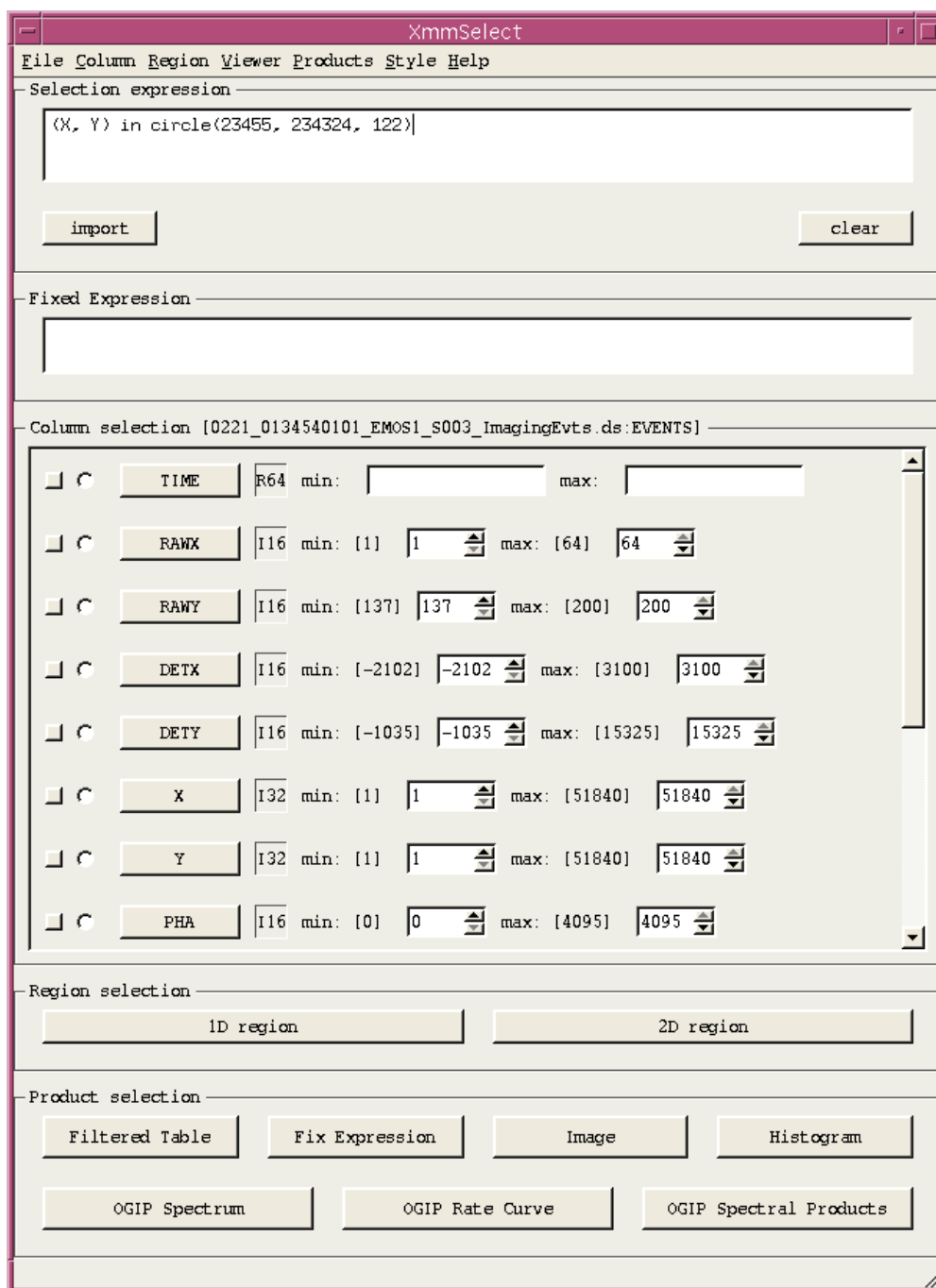


Figure: Layout of elements in main `xmmselect` window



The window consists of the following five main areas:

3.3.1 Menus

At the top of the main window are five pull-down menus labeled **File**, **Column**, **Region**, **Viewer**, **Style**, and **Help** respectively.

The **File** menu contains the following selectable entries:

field	key shortcut	purpose
New table	Control+N	Bring up a data set browser for selecting a new input table. If a new table has been selected the GUI will be re-initialized.
Reload table	Control+R	Rescan the input table and re-initialize the GUI
Exit	Control+X	Terminate xmmselect

The **Column** menu contains the following selectable entries:

field	key shortcut	purpose
Add new column	Control+D	Bring up the parameter dialog of tabcalc . If the task has run successfully the table will contain a new column. In this case the table will be rescanned and the GUI re-initialized.
Add PHASE column	Control+P	Bring up the parameter dialog of phasecalc . If the task has run successfully the table will contain a new column PHASE . In this case the table will be rescanned and the GUI re-initialized.
Delete column	Control+L	Delete the column from the table that is currently marked to be used in a 1-D product extraction.

The **Region** menu contains the following selectable entries:

field	key shortcut	purpose
Union	Control+U	If selected multiple 2-D image regions are combined with logical <i>or (union)</i>
Intersection	Control+I	If selected multiple 2-D image regions are combined with logical <i>and (intersection)</i>

The **Viewer** menu contains the following selectable entries:

field	key shortcut	purpose
Binned Intensity Image	Control+B	If selected images are displayed as binned intensity images with the task imgdisplay
Scatter plot image	Control+S	If selected images are displayed as scatter plots with the task dsplot

The **Products** menu contains the following selectable entries:



field	key shortcut	purpose
backscale on spectra	Control+C	If selected the task backscale is run automatically on all extracted spectra prior to display
eregionanalyse on source image	Control+R	The generation of OGIP spectral products (see Sect. 3.3.6.2) requires the selection of a source extraction region. If the checkbox is selected the task eregionanalyse will be invoked on the selected region to optimize its centroid and radius before the generation of spectral products is performed.

The **Style** and **Help** menus comprise entries which are common to all interactive SAS tasks, e.g. **Font** to select the font of text label, etc. As a bare minimum the **Help** menu will always contain at least the following entries:

field	key shortcut	purpose
Help	Control+H	Bring up a HTML browser displaying the xmmselect on-line documentation
About	Control+A	Display a dialog box with general information about xmmselect

3.3.2 Selection expression

A wide, editable entry field which contains the boolean selection expression with which **evselect** shall be invoked when the generation of a product is requested. Below that are two or three push buttons:

button label	purpose
PrefSelections	If a task group preference with the name PrefSelections exists it will be read and presented in the form of a cascaded submenu structure. Selecting an entry will present the value of the corresponding selection preference in the form of the below illustrated dialog. From there the expression can be accepted or dismissed with the Ok or Cancel button. See Sect. 5 for more details on the PrefSelection preference.
import	This brings up a data set browser dialog which allows to select a block (table or array) or an attribute. A selected block is inspected for the presence of DSS specifications. If none are found a corresponding warning dialog shall be presented, otherwise, the DSS specification is read, converted into a selection expression and presented in a dialog as illustrated below. Accepting this with either Import (AND) or Import (OR) will combine the expression with the current total selection expression in the indicated manner through logical AND (&&) or OR () respectively. Pushing Cancel will disregard the expression. Selecting an attribute in the data set browser behaves in the same manner, i.e., the value will be read and presented in the dialog from where it can be accepted or rejected.
clear	delete any currently defined selection expression and re-initializes the entry field

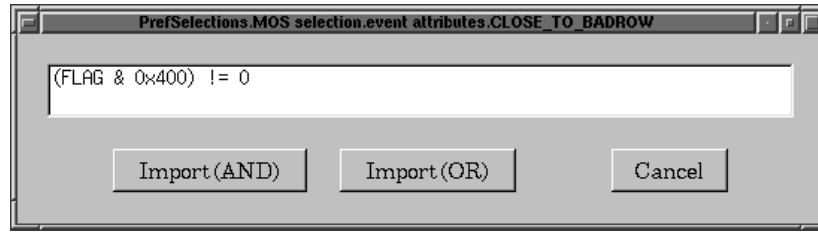


Figure: Dialog to accept/rejected imported or preference selection expressions.

3.3.3 Fixed Expression

This is a non editable entry field that contains the boolean selection expression with which **evselect** was invoked when the fix expression (see Sect. 3.3.6.1) product was requested.

3.3.4 Column selection

This constitutes the central area of the **xmmselect** GUI. It consists of a vertically arranged list of pushable buttons, each corresponding to one numeric, string, or boolean column found in the specified input table. Each button is labeled with the name of the column it represents. To its left are up to two associated check buttons which, when being selected, mark the column as primary axis in the extraction of a an image or spectrum/histogram/rates curve, respectively (for more information on the function of the check buttons see Sect. 3.3.6).

To the right of the button with the name of the column is a short framed label which specifies the types of the column:

label	column type	canonical value range
I8	8-bit integer	$[-128, 127]$
U16	16-bit unsigned integer	$[0, 65535]$
I16	16-bit signed integer	$[-32768, 32767]$
U32	32-bit unsigned integer	$[0, 4294967295]$
I32	32-bit signed integer	$[-2147483648, 2147483647]$
R32	32-bit real	$[\approx -10^{38}, \approx +10^{34}]$
R64	32-bit real	$[\approx -10^{308}, \approx +10^{308}]$
S	string	n/a
B	boolean	true/false

To the right of this label are more GUI elements whose type depend on the type of the column.

If the number of data columns in the input is large the list of push buttons might become too large to fit on the screen. In this case the entire middle area will become vertically scrollable as shown in figure 3.3.

3.3.4.1 Integer columns

Integer columns are suitable to serve as primary axis in image and spectrum/histogram extractions and, thus, have two associated check buttons left to it. To the right are two numerical, editable entry fields which determine the minimum and maximum allowed value of the quantity in that column. If the focus is set in a field, the user can simply type in new values or use the up and down arrows next to the field to increase or decrease the displayed value by 1. Pushing the button with the column name will then



determine whether any of the fields has been changed, construct a valid selection expression of either form

```
colname in [min, max]
colname >= min
colname <= max
```

and append this to the global expression in the top expression entry field (see Sect. 3.3.5). If the column possesses associated hard upper and/or lower limits these will be indicated in the attached text label and non compliant user inputs will not be accepted.

3.3.4.2 Real columns

Real column behave identical to integer columns except for

1. In addition to images and histograms they can be used as primary axis in a rate curve extraction (rightmost check button).
2. The minimum and maximum fields do not have attached arrows to increase/decrease the value by one.

3.3.4.3 String columns

String columns cannot be used in the extraction of images, spectra/histograms, and rate curves and, thus, will not be shown with associated check buttons. To the right of the pushable button with the name of the column is an editable entry field in which the user can input arbitrary text. Pushing on the column button will then append the sub-expression

```
colname==value
```

to the global expression in the top expression entry field (see Sect. 3.3.5).

3.3.4.4 Boolean columns

Boolean columns cannot be used in the extraction of images, spectra/histograms, and rate curves and, thus, will not be shown with associated check buttons. To the right of the pushable button with the name of the column are two linked check buttons labeled “true” and “false” corresponding to the two possible values. Only one value can be selected at a time and pushing the column button shall append either of the sub-expression

```
colname==TRUE
colname==FALSE
```

to the global expression in the top expression entry field (see Sect. 3.3.5).

3.3.5 Region selection

This area below the “Column selection” area contains two buttons labeled “1D region” and “2D region”. Pushing either of them will trigger **xmmselect** to read spatial region definitions from a running **dsplot** or



imgdisplay respectively. If this was successful the region specifications are converted into a corresponding boolean selection expression which then gets appended to the global expression in the top expression entry field (see Sect. 3.3.5). If no region specification could be read an error dialog box will be shown indicating the most likely cause of the failure.

3.3.6 Product extraction

This is the area at the bottom of the main window and consists of seven horizontally arranged buttons labeled **Filtered Table**, **Fix Expression**, **Image**, **Histogram**, **OGIP spectrum**, **OGIP Rate Curve**, and **OGIP Spectral Products** respectively. Pushing either of these will trigger the following actions:

1. Execution of **evselect** with the selection expression in the top expression entry field (see Sect. 3.3.5). Before **evselect** is actually executed, its parameter dialog is presented which gives the user full control over the product extraction process, i.e., all of **evselect**'s parameters (e.g. bin size for rates curve extraction, name of output data set, etc.) can be altered.
2. Accumulation of the respective product with a default name
3. Except for event lists, execution of a product displayer (either **dsplot** or **imgdisplay**) with the name of the generated product

A prerequisite to the extraction of images, histograms, spectra, and rate curves is the selection of data set columns which define the primary product axes (see Sect. 3.2.3)). This is done with the check buttons left to each push button in the middle. The leftmost check button column corresponds to the image extraction. Two and only two data set columns can be selected which defines these columns as image axes. When a column is selected in this way the associated check button will be labeled with either of the letter "X" or "Y" when being pushed. "X" is the axis which is displayed horizontally by **imgdisplay**. The attempt to select a third column will result in the last selected one getting unselected again. Trying to perform an image extraction without a prior column selection will result in an error condition. For histogram, spectrum, and rates curve extractions the user has to select one and only one data column with the rightmost check button column. These act as "radio" buttons, i.e. pushing a button will undo the previously made choice (if any has been made yet). As for images, trying to extract a histogram, spectrum or rates curve without a prior axis selection will lead to an error condition.

If an image, histogram, spectrum or rate curve has been successfully generated and displayed, the user can define one and two dimensional regions which **xmmselect** will translate into a proper boolean selection expression for a subsequent refined product generation.

3.3.6.1 Fix expression

The aim of this **Fix expression** button is to work with a smaller (in size) event list file, and therefore, speed up all the subsequent steps (product extraction, new filter expressions). If you click on **Fix expression** button, you will filter your event list file with the expression you have in the selection expression entry. And automatically, the new event list file will be loaded as the current event list file.

Once you have filtered your event list file, and the new filtered event file has been loaded automatically, the expression used for filtering is written in the fixed expression window (see Sect. 3.3.3).

3.3.6.2 OGIP spectral products

Unlike the other product categories the generation of spectral products yields more than one output data set, viz.:



- OGIP spectrum for selected source region
- OGIP spectrum for selected background region
- Redistribution Matrix File (RMF) corresponding to source spectrum
- Auxiliary Response File (ARF) corresponding to source spectrum

These are created by the task **especget** which in turn is a wrapper around **rmfgen** and **arfgen**. Please consult the description of those package to learn more about functionality, flow of operations, and limitations. The primary input to **especget** is an image data set and two spatial selection expressions which need to mark an extraction region around a point source and a source-free background region. These can be generated with **xmmselect** in a convenient manner:

1. create an image in detector or sky coordinates as described above
2. when the image is displayed in Ds9 mark a normal circular region around the source of interest
3. mark another (not necessarily circular) region that shall comprise only background events. In order that **xmmselect** can distinguish the source from the background region, the latter must be assigned the Ds9 attribute *Background*. To do that, double click on the region marker and in the upcoming dialog select “Property” and “Background”. When done the border line style should change from solid to dashed. Be sure that no “exclude” regions have been selected.
4. click on **OGIP Spectral Products**. At this point **xmmselect** shall read the two regions from Ds9 and perform some basic checks for correctness (e.g. there is exactly one source and one background region). In case this fails informative error messages with suggested corrective actions shall be presented. If the two regions could be successfully read **especget** is directly called with the necessary inputs unless the option to invoke **eregionanalyse** on source image (see Sect. 3.3.1) is active. In this case there is an additional step which aims at optimizing the position and extent of the source extraction region before it is passed to **especget**. This is done by running **eregionanalyse** on the source region. Its primary output is an optimized source region which shall be displayed as a white circle marker on the image. In addition a dialog box with some numerical results from the **eregionanalyse** run will be shown:

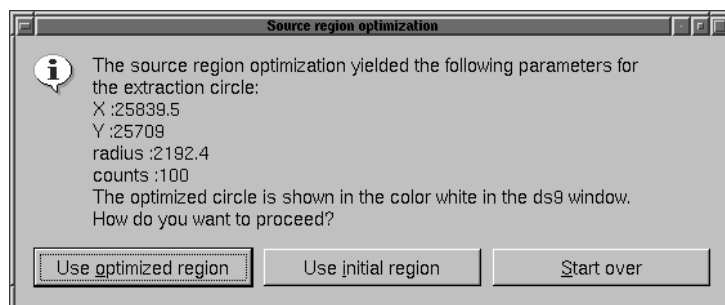


Figure: Dialog with results of source region optimization process

The dialog gives the option to proceed in three possible ways:

- (a) Use optimized region
The initially user-defined source region will be replaced by the new optimized region



- (b) Use initial region
The optimized region will be erased from the image and execution proceeds with the initially defined region
- (c) Start over
All regions will be deleted from the image - new initial selections for source and background regions must be made

In the first and second case **especget** is eventually invoked with the appropriate input.

Please note that the production of an RMF with **especget** is a computationally intensive task and can take several tens of minutes on a moderately loaded workstation. There is however the option to skip the RMF-generation stage (please consult **especget** documentation for a description) in which case the task should run to completion within minutes. If it finishes successfully the generated source spectrum shall be shown via **dsplot**.

3.4 Typical usage scenarios

It is assumed **xmmselect** has been brought up to process a table in a data set, e.g.

```
xmmselect table=event.fits:RAW
```

The following is a collection of common scientific processing problems together with instructions on how to perform them within **xmmselect**:

- Generate a spectrum image or rate file without doing any selections and visualize it.
 - 1. Select two integer/real (image), or one integer (spectrum), or one real column (rate curve)
 - 2. Push the corresponding product extraction button at the bottom to extract and subsequently visualize the respective product.
- Generate a spectrum/rate curve from events stemming from a particular region in the FOV
 - 1. Select two integer/real columns
 - 2. Push the **Image** button at the bottom and wait until the corresponding image is displayed within an **imgdisplay** window
 - 3. Mark the region(s) in question within the **imgdisplay** window
 - 4. Push the **2D region** button of **xmmselect** - a string will appear in the top expression entry field (see Sect. 3.3.5).
 - 5. Select an integer (spectrum) or a real (rate curve) column
 - 6. Push the **Spectrum** or **Rate curve** button at the bottom - the respective product shall be generated and visualized
- Identify the spatial distribution of events which account for a particular feature in a spectrum/rate curve
 - 1. Select an integer (spectrum) or a real (rate curve) column
 - 2. Push the **Spectrum** or **Rate curve** button at the bottom - the respective product shall be generated and visualized with **dsplot**
 - 3. Select one or multiple intervals in the displayed product



4. Push the **1D region** in the “Region selection” area
5. Select two integer/real columns for the image extraction
6. Push the **Image** button at the bottom and wait until the corresponding image is displayed within an **imgdisplay** window

3.5 Miscellaneous

Please note: **xmmselect** is a generic table processing tool, i.e., the input table is not required to be an event list.

4 Parameters

This section documents the parameters recognized by this task (if any).

Parameter	Mand	Type	Default	Constraints
-----------	------	------	---------	-------------

table	yes	filename		table specifier
--------------	-----	----------	--	-----------------

a specifier which must point to a table in a data set. It must be in either of the three forms **setname**, **setname+tableid** or **setname[tableid]** where **setname** must be the name of an existing data set and **tableid** an identifier of a table in that data set. If the first form, **setname**, is used, the table data are sought in the *first* block of the named data set. **tableid** can either be an identifier starting with a letter which must then be the name of a table in the specified data set or a simple number which is directly interpreted as a block sequence number (starting from 1).

centerpaneheight	no	integer	8	1 ≤ centerpaneheight ≤ 10
-------------------------	----	---------	---	---------------------------

initial height of central pane (see Sect. 3.3.4) in units of initially visible column buttons; if the input table contains more than the specified number of columns a scrollbar will be provided with which the invisible portions of the pane can be brought into view.

5 Preferences

For general information about the preference formalism, please consult the documentation of package **preferences**. The following preferences are recognized by the task:

quicklook

default value = **false**

type = **bool**

This determines the default values of the **evselect** parameters

- **writedss** (see documentation of **evselect**),
- **updateexposure** (see documentation of **evselect**)



when the generation of a product is requested. For *quicklook*=**true** those will be set to **false** which, unless altered by the user in the parameter dialog, will result in **evselect** generating the requested product in the most run-time efficient way. Setting any of the above parameters to **true** triggers additional calculations which can represent a significant run-time overhead. With *quicklook*=**true** also by default the task **backscale** will not be invoked on any extracted spectrum. This is changeable after startup via the **Products** menu entry.

If the product is not to be used by downstream analysis packages it is acceptable to skip those calculations. Please consult the corresponding **evselect** parameter descriptions for more details on the nature and the necessity of the involved computations.

withbackscaleddiag

default value = **true**

type = **bool**

When the task **backscale** is to be invoked automatically after each extraction of a spectrum (see 3.3.1 for details) *withbackscaleddiag*=**true** will bring up the parameter dialog of this task prior to running it. This gives the user the option to alter its parameter values.

PrefSelections

default value = **none**

type = **group**

This preference groups a list of frequently used selection expressions that will be available in every **xmmselect** session. The set of expressions does not have to be flat but can be hierarchically structured, i.e., the *PrefSelections* group can contain other preference groups, etc. The following is an example of a possible contents of the *PrefSelection* group:

```
"PrefSelections" {
  "MOS selection" {
    "badpix"    "FLAG & 0x111234 == 0"
    "source"    "(RAWX, RAWY) in circle(300,300,30)"
  }
  "PN selection" {
    "badpix"    "FLAG & 0x111233 == 0"
    "source"    "(X, Y) in circle(20000,20000,1000)"
    "Crab" {
      dum       "(CCDID==4&&(RAWX, RAWY) in circle(32,10,10)"
    }
  }
  "energy"     "PHA in [100:200]"
}
```

6 Errors

This section documents warnings and errors generated by this task (if any). Note that warnings and errors can also be generated in the SAS infrastructure libraries, in which case they would not be documented here. Refer to the index of all errors and warnings available in the HTML version of the SAS documentation.

NoSuchTable (*error*)

table parameter does not point to an existing table

all errors of all tasks listed above (see Sect. 3.2.1)



7 Input Files

1. a data set with a table to be processed

8 Output Files

1. image (optional)
2. histogram (optional)
3. OGIP spectrum (optional)
4. OGIP rate curve (optional)
5. filtered table (optional)
6. OGIP source + background spectrum (optional)
7. OGIP ARF + RMF file (optional)

9 Comments

None.

10 Future developments

Future versions

- might be more streamlined towards the handling of event lists, i.e., the current general purpose nature of the task might become partially compromised.
- might be more generic in terms of creating general 1- or 2-dimensional histograms from selected data in addition to the specific products spectra/rate curves, images.



References