Base of version 1.4.12, Last update April 2010

Data Subset management

On most observation sites, it is frequent that different types of geological structures can be observed and measured. The structural data are consigned in a note book with the precise indication of the type of structure measured (e.g. bedding plane, foliation plane, mineral lineation, fold axes, fault with slip line, fracture, …). When typed in most software's for processing structural data, the different types of structures have to be recorded in different files, one for each different type of structure.

Data Set and Data Subsets

With Tensor and Win-Tensor, the different kind of data can be recorded on the same data file, where they are differentiated using what is called a Subset Index. As explained in Delvaux and Sperner (2003), a *Data Set* represents the total amount of structural data measured on an, observation site independently from their type and a *Data Subset* is a portion of the total *Data Set*, generally corresponding to a particular type of structure. A Subset Index is attributed to each data set during the data input procedure. Within the old Tensor for Dos programme (and the related .dat file format), the Subset Index was an Integer ranging from 0 to 9, allowing to separate the total data set into 10 different subsets. This proved to be insufficient in some cases and was not very convenient for large data sets. With the Windows version (and the .wtd, .csv, .txt and .xls formats), the Subset Index has been expanded to decimal value ranging from 0 to 99.9 (1000 different subsets). The integer part is considered as a Major (or Primary) Index) and the full index with the decimal part, as a Minor (or Secondary) Index.

General principle

The data subset management consists as the precise definition of the data used for processing, by the definition of subsets extracted from the total data set and selecting the desired subset. Different tools are provided for this:
1. The Subset Index is defined for each data during its input in the Data Worksheet This is called the Input Index (Fig. 1)
2. When calling the Processing Sheet, the compiled data from the Data Worksheet are transferred to the Processing Worksheet, and the Working Subset Index (Subsets/Wor in the data table) receives the value of the Input Index (Fig. 2).
3. Within the Processing Worksheet, only the Working Index can be modified and this modification will not affect the Input Index of the Data Worksheet. This can be done as detailed later, either manually, using Menu / Data / Transfer or Modify Subset Indexes or during the data separation procedures.
4. The Subset Manager presents the active data subsets in the form of a TreeView with each node (or item) representing a subset and associated to a check box.
5. The check box associated to each node of the tree view is used for selecting the related Primary or Secondary subset.
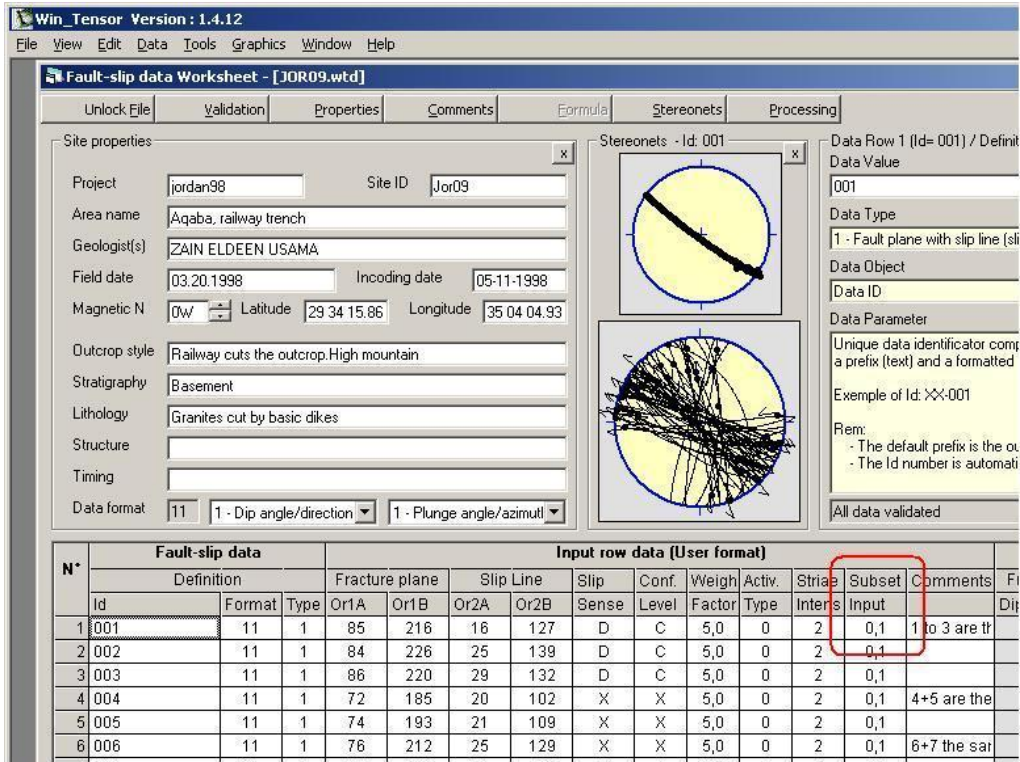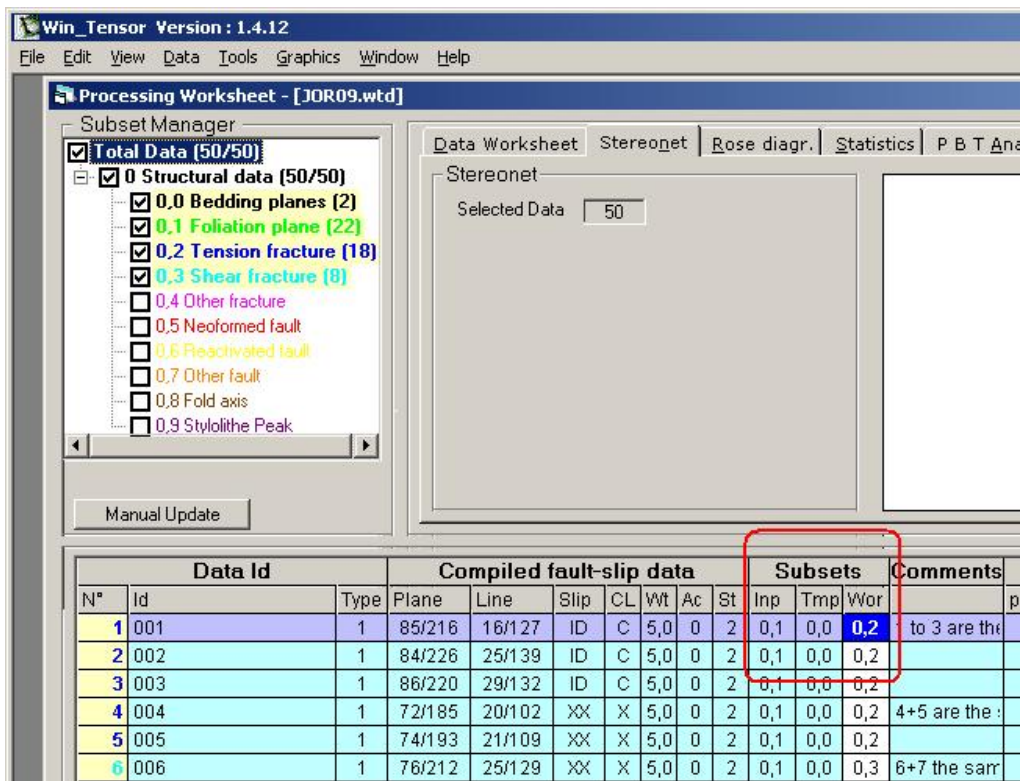
Figure 1: Subset Input Index in the Data Worksheet.

Figure 2: Subsets in the Data Worksheet (Inp: Input Index, Tmp: Temporary Index, Wor: Working Index).
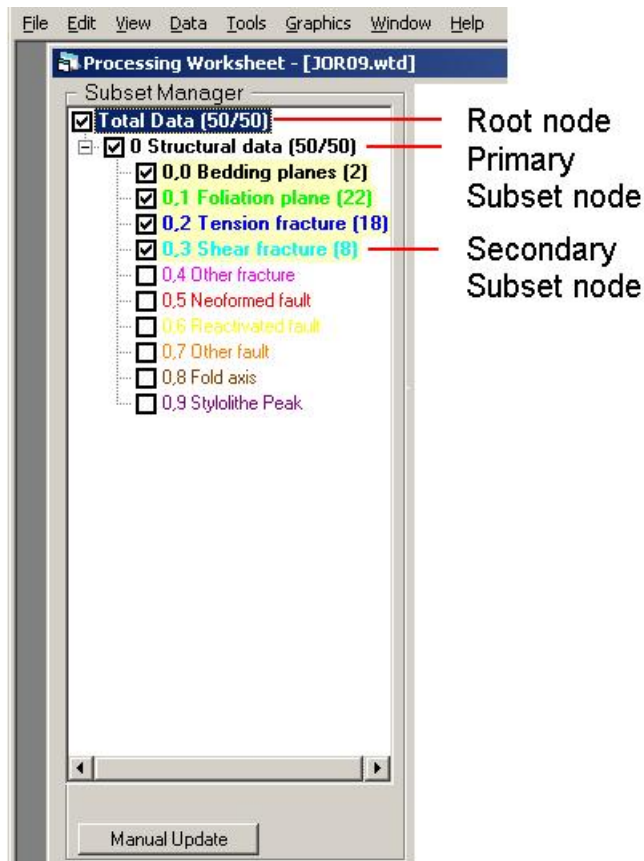
Subset Manager

The Subset Manager is the central component which primarily allows selecting the subsets which will be used for processing. The composition of each subset is defined according to the value of the Subset Working Index associated to each structural data and managed in the data table of the Processing Worksheet. The Subset Manager is also used for labelling the subsets and controlling the colors of the related data in the graphic outputs. The color of the data displayed in the graphics is set according to the font color of the label of its subset. The subset properties (subset index, labels, color and value of the check box) are stored at the end of the data file (.wtd, .csv, .txt and .xls) and are loaded together with the data when re-opening the file. Only properties that different from the default ones (void labels, black color and unchecked value for the check boxes) are stored. This is a new feature implemented in Win-Tensor since version 1.4.12. It is not included in the Dos version and therefore the subset properties are not recorded in the .dat file format to ensure compatibility with the old Tensor program for Dos.

Figure 3: Subset manager in the Processing Worksheet with definition of the TreeView nodes. Standard display mode with the Check box, the Subset Index, the Subset label in its color and the number of data contained in the subset in brackets.

Caution is requested here as the earlier versions of WinTensor do not include this feature. The files types *.wtd, *.csv, and *.txt generated by version 4.4.12 or above will no more be fully compatible with the earlier versions. If they are loaded using an earlier version, the subset properties at the end for the file will also be loaded as a normal structural data and this will generate warning errors. To avoid this, it is advised to delete the data corresponding to the subset properties at the end of the file before loading it. This can be done within the Data Worksheet of Win-Tensor also after loading the file but then a warning message will appear for each line of Subset property.

Working with the Subset Manager

Automatic population of the TreeView

When the Data Processing Worksheet is loaded for the fist time after the file has been created, the TreeView of the Subset Manager is only populated with the Primary nodes corresponding to the subset indexes used in the Data Worksheet (defined as Input Indexes), and the check box of all of them will be checked . All the Secondary nodes attached to that Primary node will also be loaded but only those containing data will be checked. If a new subset is defined within the data table of the Processing Sheet (in the Subset Working Index column), the corresponding node will appear in the Subset Manager; but it will remain unchecked. Once loaded, a node will remain present in the TreeView even if no more data are attached to the corresponding subset.

By default, all nodes (with exception of the root node which is labelled "Total Data") are labelled by its Subset Index (Integer for Primary Subset and decimal number Secondary Subset). The label is completed by the number of data contained in the Subset, under brackets. For the root node, a first number corresponding to the total number of data selected in the Data Set is separated by the sign "/" from the total number of data contained in the database. For the Primary Subsets, the numbers corresponds to the data (selected/total) contained in the Subset only. For the Secondary Subsets, the value indicated represents the number of data contained in the Subset, selected or not. No number between brackets appears for subsets that contain no data.

When a file with recorded Subset properties has been loaded, the TreeView of the Subset Manager in the Data Processing Worksheet is populated according to the defined properties (labels, colors and value of the check boxes).

Selection of Subsets

The selection of a subset is done simply by checking the associated check box using a single mouse left click A double click on a Primary Subset node will either deactivate all attached Secondary nodes if the Primary node was already checked. If it was unchecked, then all the Secondary nodes will be checked. Similarly, double clicking on the Root node will affect accordingly all attached Primary Subset nodes.

Changing the node label display mode

In the default mode, long labels are displayed for each node with successively the Subset Index, the subset label and the number of data between brackets as explained above). When loading the data for the first time, no labels are defined, so the number of data between brackets immediately follows the Subset Index. The width of the Subset Manager is also a bit expanded to allow proper display of the long node labels.

Short labels are composed of a prefix "Prim.Sset" or "Sec.Sset" followed by the Subset Index and the number of data.

Changing from one display mode to the other can be done using the Menu / Edit / Short/long subset label command or a right click on the Root node. When the Root node is selected an

option Subset data number appear on the Menu / Edit command and clicking on it removes or adds the number of data between brackets at the end of the node label.

Editing the Subset labels and colors

The dialog box for editing the node properties appears by right-clicking on the node line of a Primary Subset (Fig. 4). It allows to edit the label and the color of the selected Primary Subset and also of all the 10 Secondary Subsets attached. The labels can be edited directly in the text boxes, and the colors can be changed by clicking on the colored buttons. This will cause the color dialog box to appear (Fig. 5). Alternatively, pre-defined labels and colors assembled into a series of Palettes can be loaded by activating the drop-down control. The current palette can be saved as User palette and will be stored in the first item of drop-down control, from which it can be re-loaded. Once the properties are defined, they can be applied to the active data by pressing the Apply button (the Subset properties window will be closed).
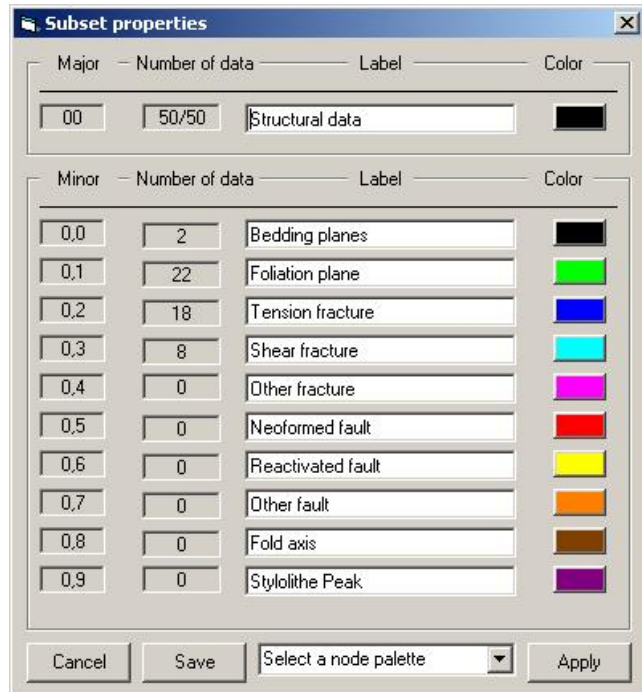
Figure 4: Subset properties dialog box.

The color of each node can also be edited individually by right-clicking on the node line.
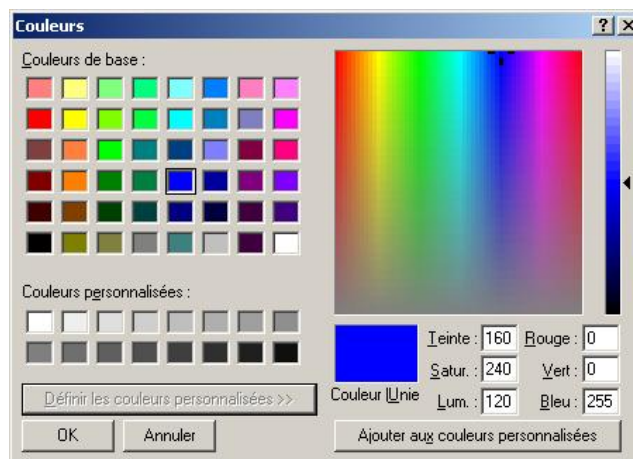
Figure 5: The color box, expanded.

As in many Windows programs where different ways are provided for accessing the same command, the same result can also be obtained using the Menu / Edit commands. When a Primary Subset node is selected the item below the first horizontal bar in the Menu / Edit panel is changed into "Primary Subset Properties". Clicking on it displays the Subset properties dialog box. When a Secondary Subset node is selected, the menu item is changed

into "Minor Subset Color" and clicking on it displays the Color dialog box for the selected Secondary Subset. As said above, when the Root node is selected, the menu item is turned into "Subset data number" and clicking on it removes or adds the number of data at the end of the node label

Primary Subset versus Secondary Subsets colors

To avoid confusion by using a large number of different colors, the colors can be defined either at the Primary Subset level or at the Secondary Subset level, but these definitions are are mutually exclusive. When a color other than black is defined for a Primary Subset, all the related Secondary Subsets will be displayed with this color, even if they have also colors defined other than black. The colors associated to the nodes of the Secondary Subsets will not be affected and resetting the color of the Primary Subset to black will reactivate the normal affectation of colors to the Secondary Subsets.

Subset Legend

A legend of the active Subsets can be placed in some of the graphics in the different panels of the Processing Worksheet. With version 1.4.12, a subset legend can be inserted in the stereonet picture using the Graphic Options dialog box.

Immediate / Delayed application

Every change in the Subset manager, either in the selection (using the checkboxes) or in their properties (color) is reflected in the active processing panel. A change of selection will cause the reprocessing of the data in function of the active panel and a change in color will cause a refreshing of the graphics using the new colors for the data display. For data bases with less than 100 data, this is has immediate effect, while for larger data bases, the data reprocessing and graphic refreshing will be done only after applying on button *Update* at the bottom of the Subset Manager (Fig. 6). The latter will become when a change in the Subset manager has been made. This has been set in order to avoid unnecessary waiting when several modifications has to be made on the Subsets before applying the changes to databases with a large number of data.
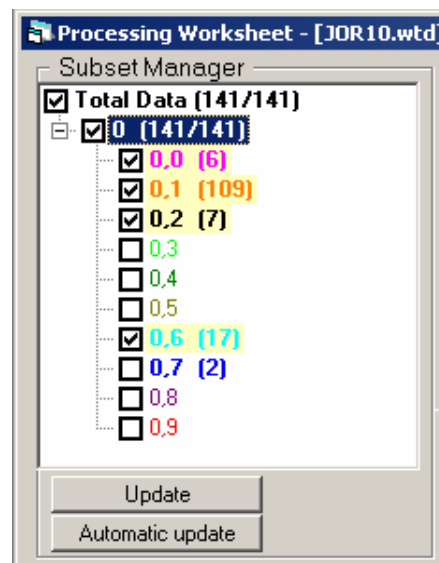

Figure 6: Buttons at the bottom of the Subset Manager controlling the Immediate / Delayed application of the changes in the Subset selection or properties.

The delayed application can be suppressed by clicking on button *Automatic Update*. By doing so, the button *Update* will be no longer visible and the label button *Automatic Update* will be changed into *Manual Update*. The delayed application is reactivated by clicking on button *Manual Update*.