Quality Control of Geochemical Data

After your data has been imported correctly into the system, your master database should contain only pristine data. Practical experience in exploration geochemistry, however, shows that quality control and due diligence responsibilities continue even after initial data import. The Geochemistry system is intended to streamline geochemical quality control by providing a single environment for quickly calculating summary statistics and manipulating data values. In addition, you may want to further evaluate your data by applying math expressions and evaluating results, or by editing individual (problem) data values.

Geosoft provides sample data for you to use when working through these How-To Guides. These data files can be found in your ".../Geosoft/resourcefiles/data/geochemistry" folder.

Calculating Summary and Advanced Statistics

The Geochemistry system quickly calculates and displays summary and advanced statistical analysis of your ASSAY channels. The system also enables the export of the ASSAY channel statistics to a CSV file. The CSV file is saved in your current project directory.

It is assumed that you will carry out summary statistics on individual channels. This approach enables you to quickly see that the minimum and maximum values for each element are within the accepted data range. The system also indicates the number of dummy values (placeholder) within a channel. This information is useful for recognizing missing data, or for determining where the import routine has replaced negative values (assumed to be the lower detection limit) with a positive value at half the lower detection limit (or negative value). The advanced statistics function provides a detailed statistical analysis of each individual channel.

Calculate and View ASSAY Statistics

1. On the *GeochemAnalysis* menu, select **Summary Statistics**.

   The *Summary Statistics* dialog appears.
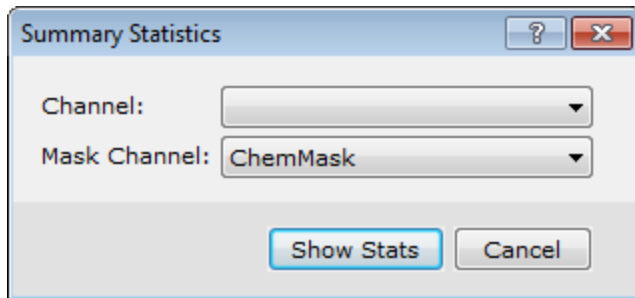
2. From the *Channel* dropdown menu, select the channel you wish to calculate.

3. From the *Mask Channel* dropdown menu, select a mask channel (optional).

4. Click the **Show Stats** button to view the basic stats.

   Key fields you should have a careful look at include the number of items and the number of dummies. The following table describes the summary statistical values that are reported.

Element	Name of the channel.
Number of items	Number of "non-dummy" items in the channel.
Number of dummies	Number of blank items. Dummies are ignored in the calculation.
Minimum	Minimum value.
Maximum	Maximum value.
Range	Range of values.
Mean	Arithmetic mean or average value. This is equal to the sum of all values divided by the number of items.
Standard Deviation	Standard deviation in the data. This is equal to the square root of the variance in the data, and is an estimator for the "width" of the data.

5. To view the extended statistics for a channel click the **More** button. The *Advanced Statistics* dialog is displayed.

6. To return to the *Summary Statistics* dialog box click **OK**. You can now view the statistics of another assay channel or cancel by clicking the **Cancel** button.

Saving ASSAY Statistics

1. On the *GeochemAnalysis* menu, select **Save Statistics**.
   The Output ASSAY channel statistics to a file dialog appears.
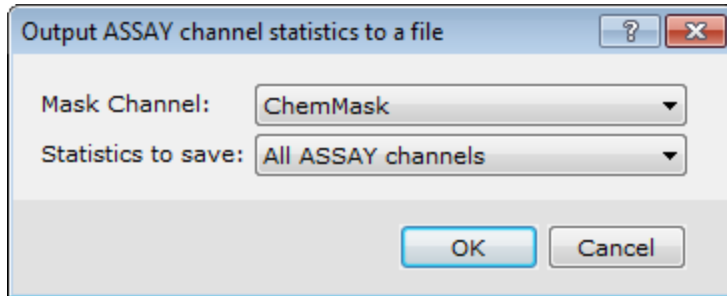
2. From the *Mask Channel* dropdown list, select a mask channel (optional).

   Select an item from the *Statistics to save* dropdown list.

3. Click **OK**. The ASSAY statistics file will be saved in your current project directory. This file can be opened in Excel (or any text editor) for viewing or printing.

Dynamic Data Links

Dynamic links are virtual geo-referenced connections between spreadsheets, profiles, and maps. These links work in two directions, so that you can select any window and see the equivalent data in all other windows you are using.

Shadow Cursor Tool

The *Shadow Cursor Tool* enables Desktop Dynamic Linking, which is designed for visualizing data for a specific line or for visualizing multiple maps simultaneously — the system enables you to place a special *Shadow Cursor* link ( ) on one or more maps to locate a specific part of the data. When you open a map, the system automatically turns on a map link and activates the Shadow Cursor button on the Map Tools Bar. If you select a value in the database or a point in a profile, you see a Link cursor that moves to the same location on the map.

 Shadow Cursor Data Linking Tool

The *Shadow Cursor Data Linking Tool* enables Database Dynamic Linking, which is designed for visualizing data for an entire database — the system automatically scrolls through the database to show the corresponding data each time you select a line/group on a map ( ). You can only use Database Dynamic Linking if your map contains group linking items (i.e. survey lines, symbols, down hole sections or ArcGIS LYR files). The advantage of database linking is that you can quickly navigate through your database, using all three views (spreadsheet, profile and map) to select data for further processing or to analyze and interpret the dataset.

⚠️ *For more information on the Shadow Cursor Tool (desktop dynamic linking) and Shadow Cursor Data Linking Tool (database dynamic linking) see the Oasis montaj Online Help system.*

Search and Replace Values in Database Channels

Search and replace enables you to quickly search any *Geochemistry* database channel for a specific value or text string and replace all found targets with a new value or text string.

Searching and Replacing Values in a Database

1. From the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Search/Replace**.
   The *Search and replace values in a channel* dialog appears.
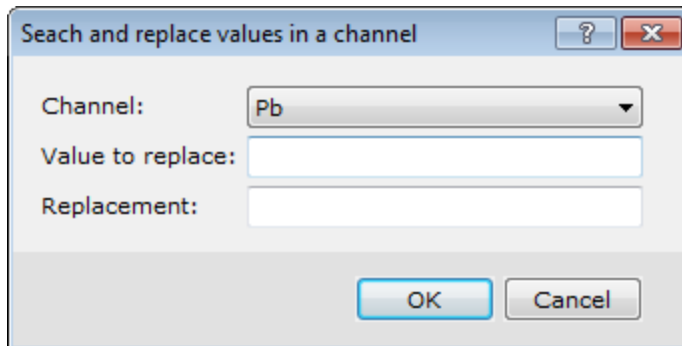
2. Using the dropdown list, select the *Channel* to search.

3. Then in the *Value to replace* box, type the value you wish to replace.

4. In the *Replacement* box, type the value you wish to replace it with.

5. Click **OK**. The system replaces the values in the channel.

Sort Data Based on One Channel

A common procedure is to sort data channels to examine extreme high or low values in detail. From a practical standpoint, this capability can be useful if you have a data value that is significantly higher than all other values; high values affect symbol plotting and gridding. This type of quality control can help you decide whether you want to edit certain values so they do not distort symbol plots and grids.

Sorting Data Based on One Reference Channel

1. On the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Sort All by 1 Channel**.
   The *Sort all channels based on a reference channel* dialog appears.
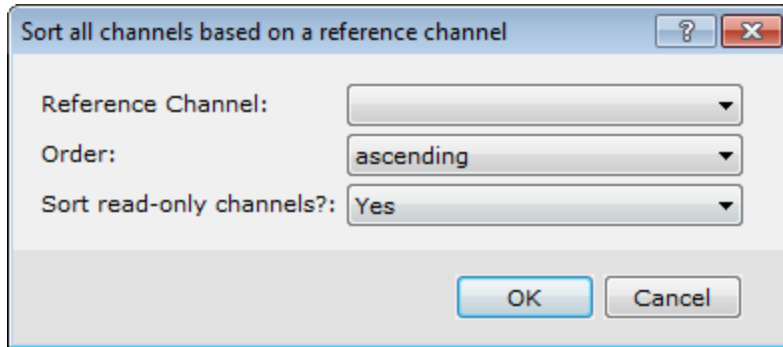
2. On the *Reference Channel* dropdown list, select a channel, for example, "Cu".

3. On the *Order* dropdown list, select *ascending* or *descending*.

⚠️ *If read-only (protected) channels are not sorted, the fiducial relationships will be lost, and data integrity is lost.*

4. Click **OK**. The system sorts the database based on the selected channel (so that the data integrity is preserved).

Sort Data Based on Two Channels

Another common procedure is to sort your data based on two channels. You may want to sort by batch or sample number and an element to check data quality or sort by two related elements (e.g., Ni and Cr) to compare highs and lows. The data are first sorted by the primary reference channel, then by the secondary channel.

Sorting Data Based on Two Reference Channels

1. On the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Sort All by 2 Channels**.
   The *Sort all channels based on two reference channels* dialog appears.
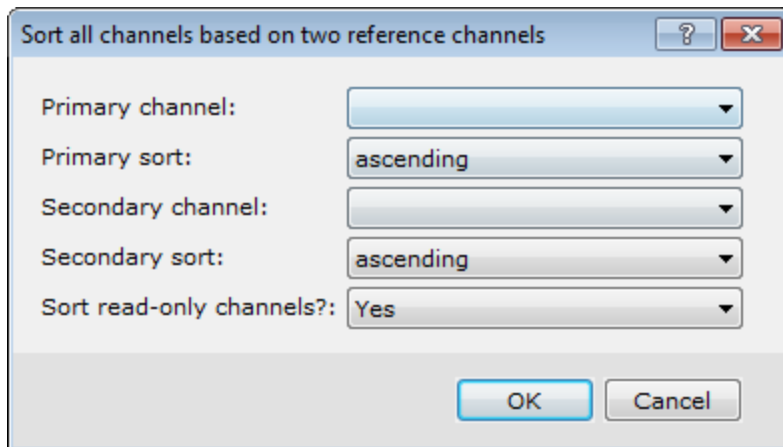
2. Use the *Primary channel* dropdown list to select the primary reference channel for sorting.

3. Use the *Primary sort* dropdown list to select either *ascending* or *descending*.

4. Use the *Secondary channel* dropdown list to select the secondary reference channel for sorting.

Use the *Secondary sort* dropdown lists to select either *ascending* or *descending*.

5.

⚠️ *If read-only (protected) channels are not sorted, the fiducial relationships will be lost, and data integrity is lost.*

6. Click **OK**. The system sorts the database first by the primary channel, then by the secondary channel.

Work with Channels

We recommend that you perform all data editing in a "new" channel, so that the original data remain unchanged. In the following procedure, you first copy a channel to a new channel (the new channel opens automatically), display the channel, and edit the last value in the channel.

Copying a Channel

1. On the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Copy Channel**. The *Copy a channel* dialog appears.

2. Use the *Copy FROM* dropdown list to select the channel to copy the content from, for example "Cu".

3. In the *TO* field, type a name for the new channel, for example "Cu_edit".

4. You can usually leave the remaining fields to the default values, if you are neither renumbering the channel nor changing the sample increment.

5. Click **OK**. The system copies the channel (column).

⚠️ *You must now display the new channel in your database spreadsheet.*

Hiding a Channel

1. Select a channel header cell (for example, Cu).

2. Press the space bar.

The system removes the channel header from the spreadsheet. This does not affect the actual data. You can redisplay the channel at any time as described above.

Colour Channel Values

Geochemistry now has the ability to colour channel values based on user defined colour zones. Colour zones can be defined by using the *Define Colour Zones* Tool or by applying a pre-existing *.ITR file to the channel. Note that, the *.ITR file must be in the form "_Channel.ITR" with "Channel" being the name of the Channel to colour.

⚠️ *Histograms derived from coloured channels are automatically coloured using the same colour *.ITR file.*
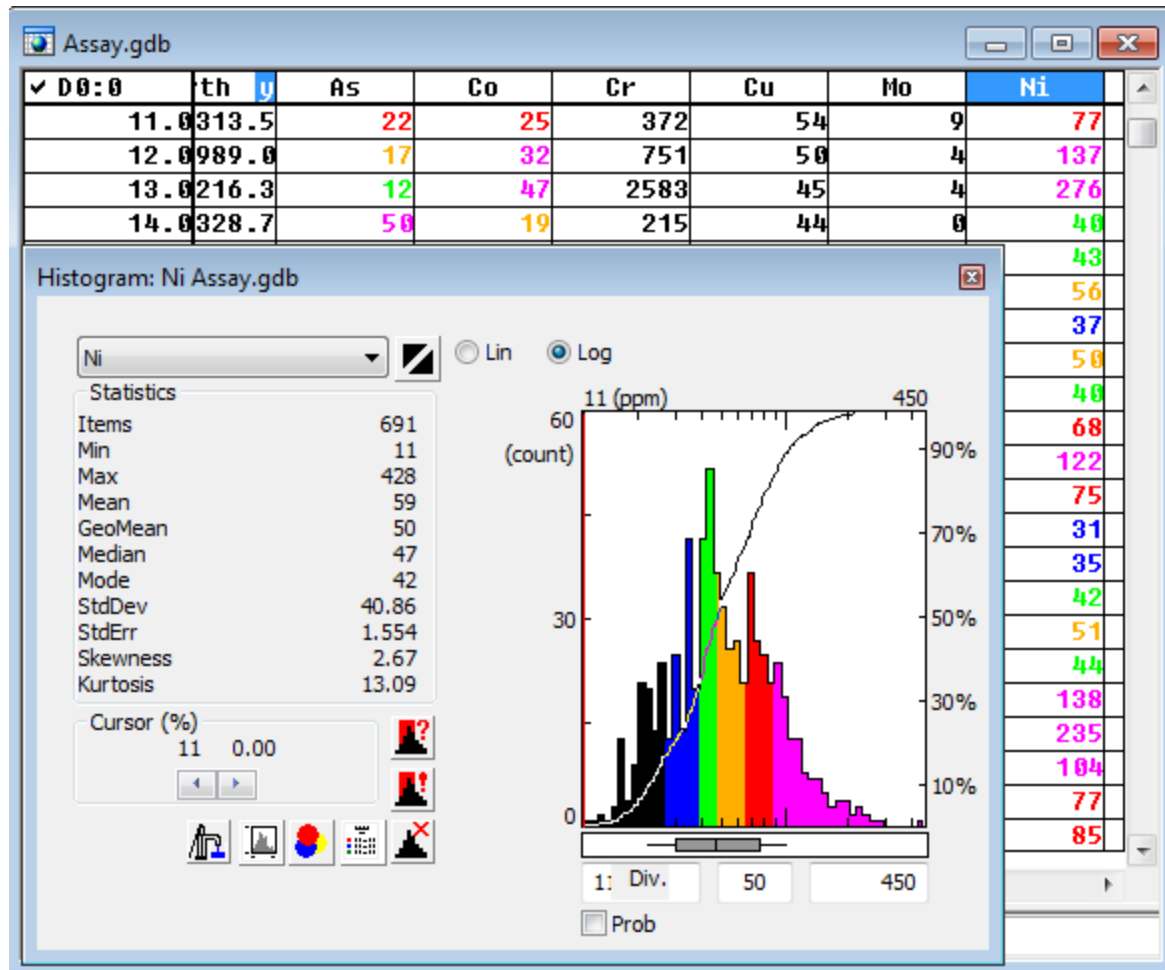
Colouring Channels using the Define Colour Zones Tool

The *Set Channel Colours* menu item can be accessed by right clicking on a selected (highlighted) channel.

1. Select (highlight) the channel you want to colour.

2. *Right-click* and from the pop-up menu select **Set Channel Colours**.
   The *Define Colour Zones* tool will be displayed.
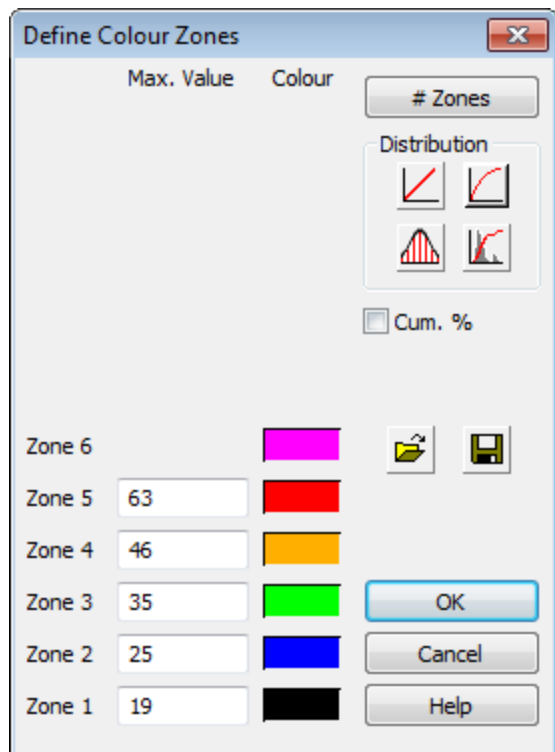
⚠️ *If you have an *.ITR file in the form "_Channel.ITR" (with "Channel" the name of the selected channel) in your project directory this file will be retrieved and used to colour the channel.*

3. The channel statistics are calculated using equal-area. Note that, the channel statistics are calculated for all the currently selected lines.

4. Using the *Define Colour Zones* tool you can set the number of zones (up to 8), set the transform type (linear, log, equal-area, normal). The percent button enables you to see the equivalent percent values.

5. Click **OK** and the channel values are automatically coloured.

⚠️ *Histograms derived from the coloured channel will automatically be coloured using the channel colours.*

Colouring a Channel Using a Zone File

Database channels can be coloured using a predefined zone file (*.ITR, *.ZON, *.TBL, or *.LUT). Zone files map ranges of values to specific colours. To apply a zone file to a database channel, the file must be in the current project directory, in the form _Channel.ITR, where "Channel" is the name of the channel to be coloured, and ".ITR" is the zone file extension. This file is automatically retrieved by Geochemistry and used to colour the channel.

1. From the *Define Colour Zones* tool, click the **Load from file** button.

2. Navigate to the zone file (.itr, .zon, .tbl, .lut) and click **Open**.

⚠️ *To save your current colour zones to a file, click the Save button.*

Removing Colours from a Channel

1. Select (highlight) the channel you want the colour removed.
2. *Right-click* and from the pop-up menu, select **Remove Channel Colours**. The colours will be removed from the channel.

Protecting Channels

The channels you create are unprotected (protected channels are indicated by a black triangle in the left-hand corner of the channel name cell). It is recommended that you protect your channels on a regular basis so that interim results are not overwritten accidentally.

1. *Right-click* on the channel header cell.
2. From the pop-up menu, select **Protected**.

To Remove Protection from a Channel

1. *Right-click* on the protected channel header cell.
2. From the pop-up menu, deselect **Protected**.

⚠ *You can also protect all of the channels by right-clicking on a channel header cell and selecting **Protect All**. You can also remove the protection from all of your channels by right-clicking on a channel header cell and select **Protect None** from the pop-up menu.*

Applying Geochemical Math Expressions

Often you may want to apply a math expression (such as Log10) to evaluate the Log distribution of the data. You can apply simple or more complex expressions as required.

The Oasis montaj system enables you to define and apply mathematical expressions in three ways:

➢ Using the Channel Math Expression Builder
➢ From an expression file
➢ Interactively in the Spreadsheet window

⚠ *Before you begin, we recommend that you create a new channel for storing any changes you make to the original imported data.*

Using the Channel Math Expression Builder

1. On the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Channel Math**.
   The *Channel Math Expression Builder* dialog appears.
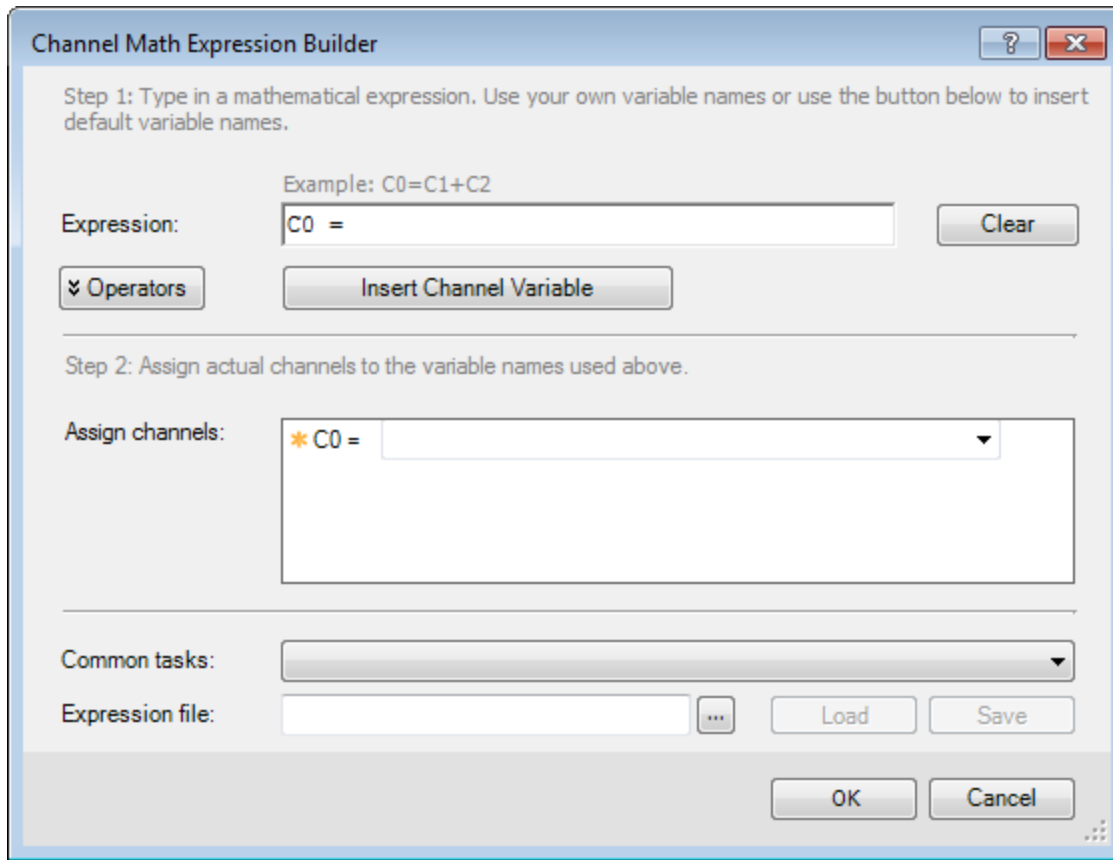
2. In the *Expression* box, add an expression where the variables are 'existing' channel names (for example, Cu_ edited/1000).
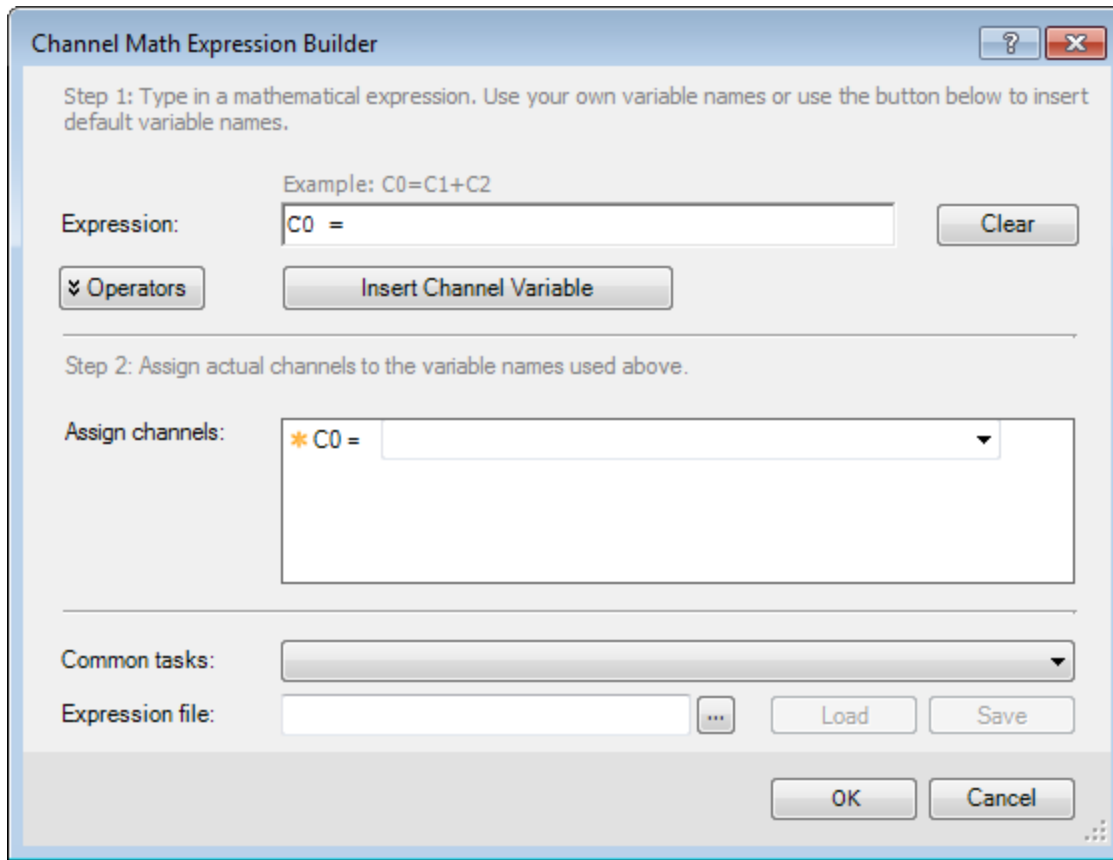
⚠️ *For detailed instructions on how to use the Math Expression Builder, click the **Help** button.*

3. Click **OK**. The result of the expression is placed in the currently specified cells.

To Apply a Mathematical Expression from a File

1. On the *GeochemAnalysis* menu, select *GeochemUtilities* and then select **Channel Math**.
   The *Channel Math Expression Builder* dialog appears..

2. Using the *Expression file***Browse** button, locate the math expression file (*exp) that contains the math expression you want to apply.

3. Click **OK** to apply the math expression the to specified cells.

To Create and Save a Mathematical Expression file

A math expression file can be created in your default text editor program (eg. Notepad). A math expression file may contain many lines of expressions and blank or comment lines. Comment lines are indicated by a / character as the first character of the line. For example, following is an expression file for squaring the sum of two channels:

/Channel C will be the square of the sum of channel A and channel B:

@sum = A + B; C = @sum * @sum;

⚠ *Each expression must end with a semi-colon (;) character. You can place more than one expression on a line as long as each expression ends with a semi-colon.*

Save your math expression file with an (.EXP) extension. *Oasis montaj* will display files with the (.EXP) extension in the *Expression file* browse dialog.

Applying Math Expression Using the Formula = Line Feature

1. Create a new channel for storing your results (e.g., Cu_Edit), and display that channel in the spreadsheet. Set the channel's Class attribute to "ASSAY"

2. Copy data from the original channel to the new channel.

3. Select the channel you want to change by clicking three times on the channel header cell so that the channel and header are highlighted.

4. Press the **=** sign on the keyboard. This changes the edit line of the screen to "*Formula*=" (bottom left hand corner of the screen).

5. In the Formula= entry field, type an expression where the variables are existing channel names (for example, Cu_edited*1000.

6. Press **Enter**. The result of the expression is placed in the currently selected cells.

To Compose Geochemical Math Expressions

The following list provides examples of simple math expressions.

1.) Multiplying or dividing values (changing ppm to ppb or ppb to ppm)

Cu_edited*1000
Cu_edited/1000

2.) Determining Ratio values

Cu_edited/Pb

3.) Replacing outliers with dummy

Cu_edited>200?dummy:Cu_edited
i.e., if Cu_edited is greater than 200, then put dummy, else put Cu_edited
(Cu_edited>100&&Cu_edited<200)?(new=1):(new=DUMMY)
i.e., if Cu_edited greater than 100 and less than 200, then put new = 1, else put new = dummy

4.) Window data channel

Window(Cu_edited,0,100)
i.e. all values < 0 and >100 will = dummy

5.) Log transformations

log(Cu_edited)
i.e Natural log
Log10(Cu_edited)
i.e. Log base10

6.) Calculating Distance between each point and the origin of the local coordinate system

sqrt(X*X+Y*Y)
i.e. expression assumes that "X" and "Y" are coordinate channel names.

Publication Date: December-17-15