
CurveExpert Basic

Release 1.4

Daniel G. Hyams

October 05, 2010

CONTENTS

1 Quick-Start Guide	3
1.1 Features	3
1.2 CurveExpert is Shareware	4
2 Usage Overview	5
2.1 Menu Bar	6
2.2 Tool Bar	6
2.3 Status Bar	6
2.4 Data Window	6
2.5 Ranking Chart	6
2.6 Graph Preview Area	6
2.7 While CurveExpert is Working	7
2.8 Printing	8
3 Working with Data	9
3.1 Introduction	9
3.2 Opening and Saving Data Files	9
3.3 Importing Data Files	10
3.4 Overview: Data Manipulation	12
3.5 Data Scaling	13
3.6 Data Translation	13
3.7 Data Clipping	13
3.8 Data Sorting	13
3.9 Getting Information about your Data	14
3.10 Using the Built-In Spreadsheet	15
4 Modeling your data	17
4.1 Introduction	17
4.2 Linear Regression	17
4.3 Nonlinear Regression	18
4.4 Spline Interpolation	23
4.5 Lagrangian Interpolation	24
4.6 Using CurveFinder	25
4.7 Calculating Groups of Models	26
5 Graphing	29
5.1 Graphing Window Overview	29
5.2 Graphing Controls	30
5.3 Graphing Menu and Shortcuts	30

5.4	Graph Properties	31
5.5	Graph Input and Output	36
6	Querying Results	37
6.1	Model Information	37
6.2	Analyze	40
6.3	Creating a Table	42
6.4	File Export	44
7	Defining Custom Models	47
7.1	Overview	47
7.2	User Model Dialog	47
7.3	Custom Model Reference	49
8	Computational Performance	51
8.1	Increasing Performance	51
8.2	Floating Point Exceptions	52
9	CurveExpert Options	55
9.1	Application Options	55
9.2	Regression Options	57
9.3	Printing Options	58
10	Appendix: Built-In Nonlinear Regression Models	59
10.1	Exponential Family	59
10.2	Power Family	59
10.3	Yield-Density Models	60
10.4	Growth Family	60
10.5	Sigmoidal Family	60
10.6	Miscellaneous Family	60
11	Appendix: Legal Information	63
11.1	Copyright	63
11.2	Software End User License Agreement	64
12	About the Author	67
13	Contacting the Author	69
14	Frequently Asked Questions	71
14.1	Can CurveExpert work with more than one independent variable?	71
14.2	How can I cite CurveExpert in a publication?	71
14.3	Why does CurveExpert suddenly erase my curve fits off of the ranking chart?	71
14.4	How can I increase the speed at which the graphs are updated?	71
14.5	Why does the chi-square reported in the Information dialog not match the standard error?	72
14.6	Why is there a huge number for standard error and correlation coefficient sometimes? Is this a bug?	72
14.7	As nonlinear regression progresses, the indicator for standard error (S) sometimes freezes, although the iteration counter keeps on ticking. What is going on?	72
14.8	A spline fit never matches my data at the points like it is supposed to, and sometimes has spikes and/or wildly oscillating tendencies. Is there something amiss?	72
14.9	Lagrangian interpolation sometimes bombs with a floating point exception. What gives?	73
14.10	I am using a Postscript printer driver, and my plots will simply not print in the orientation (portrait or landscape) that I tell it to. Is there anything that I can do?	73
14.11	CurveExpert saves the previous window position, and I don't want it to. How do I turn this behavior off?	73

14.12 How do I obtain a debugging report to send to the author?	73
14.13 Why are grid lines drawn as solid lines (instead of dotted) on some computers?	73
14.14 The Internet links pointed to by the menu items under Help Internet are now broken. Is there any way that I can change the addresses to the correct home page?	74
14.15 The file import dialog never appears, now that I need it. How do I get it back?	74
15 References	75
16 Indices and tables	77

Contents:

QUICK-START GUIDE

This guide leads you through the bare minimum of information you need to know to run CurveExpert successfully.

1. Load data into CurveExpert. In this quickstart, we will simply read a sample text file supplied with CurveExpert. Choose File|Open (see *Opening Data Files*), and double click the file BEANROOT.DAT. When the File Import (see *Importing Data Files*) dialog appears, simply click OK.
2. Perform all curve fits in CurveExpert's library. In most cases, you can select a particular model from the Apply Fit menu, but in this demo, we will just calculate them all. Choose Tools|CurveFinder (see *Using CurveFinder*), and then press OK in the following dialog.
3. Examine the curve fit information. A graph (see *Graphing*) for the Weibull model should now be displayed on your screen. Right click anywhere on the graph and select Information (see *Model Information*). Browse through the tabs as desired, and press Close when finished.
4. Export a table. Again right click anywhere on the graph and select Generate Table (see *Creating a Table*). Navigate to a folder that you have permission to write a file to, and also type in the filename "weibull.tab", and set the minimum to 0, maximum to 2, and the increment to 0.01. Click OK. This will generate a text file with the Weibull model evaluated within the range just given.

At this point, you can go back to the main window and double click any curve fit listed in the ranking, which is in the top right area of the main window. Double clicking these entries will display a graph for that model with the data, and you can perform similar operations to examine the coefficients, analyze the curve, or generate more tables, among other things.

1.1 Features

- Over 35 built-in regression models to choose from, along with over 15 additional models provided as user-defined models
- User-defined models allow you to input, design, and manage your own regression functions with up to 19 parameters
- Toolbox of fitting techniques includes linear/nonlinear regression and many types of interpolation/spline fitting.
- Uses the Levenberg-Marquardt algorithm for quick and robust nonlinear regression performance.
- Support for uncertainties in each data point
- Unlimited number of data points.
- As curve fits are applied, a ranking chart keeps track of the best to the worst models
- CurveFinder examines every possible regression model for your data set and returns the best one

- Full-featured graphing capability with customizable, dynamic graphs gives immediate feedback on curve fit quality
- User has the ability to terminate a computation at any time
- Customizable plots can be copied directly to the clipboard for use in another Windows application
- Easy to use point-and-click interface
- Reads simple ASCII data files and can ignore comments or text in the file; file import features enable the reading of a wide variety of files
- Data set can be scaled, translated, sorted, the outliers removed, or edited by hand similarly to a spreadsheet
- Analysis tools allow the location of exact points on the curve fit as well as differentiation and integration of any model.
- Full featured toolbar and status bar increase your productivity

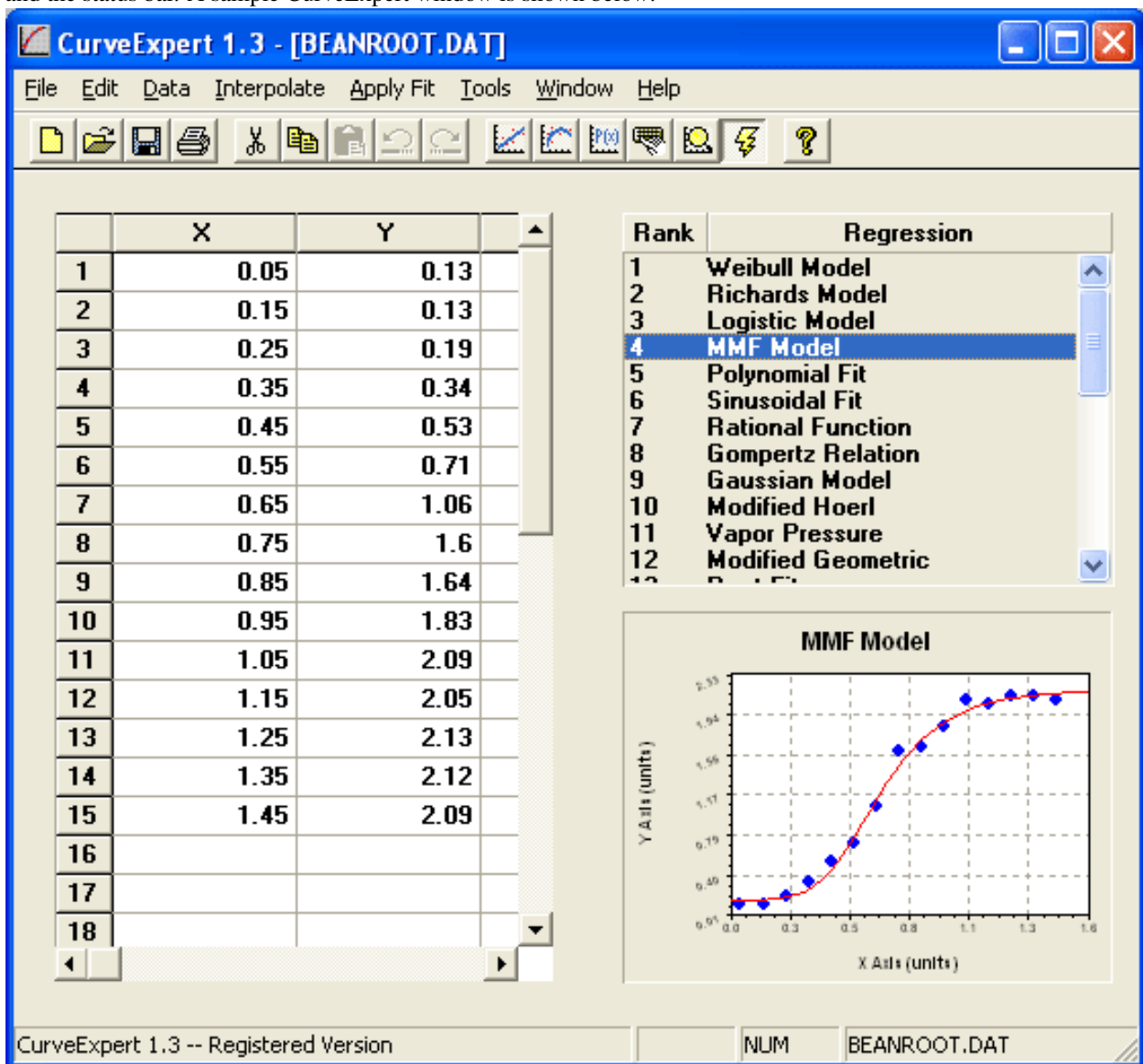
1.2 CurveExpert is Shareware

If you intend to keep this program, please register it so that I can continue development and support of this program. Registration is quick, easy, and inexpensive – the fee is only \$45.00.

See <http://www.curveexpert.net> for more information on the platforms supported by CurveExpert. High school or college students are eligible for a student discount; simply include a copy of your university or high school ID. Send your registration today! You may also read over the benefits of registering section for extra information on registering. If you have already registered, I extend a warm welcome and thank you. If you have any suggestions, or there is any regression model that you would like to see added to the repertoire, please send me mail or e-mail right away. I will gladly to listen to your comments. Development of this software has taken countless hours of study, programming, and debugging. Please support the shareware concept and register the software if you intend to keep and use it! The future of this product depends on the response that I receive from you, the user. If you would like for this project to continue and grow, please send suggestions and register, if you use this program.

USAGE OVERVIEW

The CurveExpert main window is divided into five main parts: the data window, ranking chart, the toolbar, menu bar, and the status bar. A sample CurveExpert window is shown below:



2.1 Menu Bar

Like all Windows applications, you can select items off of the menu bar to accomplish certain tasks. To find out what each menu item does, [click here](#).

2.2 Tool Bar

The toolbar gives access to frequently used commands. Watch the status bar (at the bottom of the main window) as you move the mouse across each button to find the function of each one of these buttons.

2.3 Status Bar

The status bar shows you the current status of CurveExpert and gives you help on the function of the toolbar buttons.

2.4 Data Window

This window is the spreadsheet-like interface that occupies the left side of the CurveExpert main window. The data window holds all of the data that is to be modeled. The first column of data is the x data, while the second column holds the y data. As of this version of CurveExpert, the third column is not used. To edit a data point, simply click the cell that it is in and type in the new value.

2.5 Ranking Chart

The ranking chart ranks all of the valid regression models from best to worst. The standard error and correlation coefficient (described in detail in the Data Modeling section) are used to determine these rankings. Selecting an entry in this chart causes it to be previewed in the Graph Preview Area, and double-clicking an entry opens a graphing window for that particular fit. Note that interpolations are not included on this list.

2.6 Graph Preview Area

The graph preview area shows the currently selected plot in the ranking chart (above), as a small thumbnail, so that you can quickly visualize different curve fits without opening a large number of graphing windows. This preview window also gives the capability of watching a regression as it progresses to convergence.

The graph preview area on the CurveExpert main window gives you an easy way to quickly visualize various curve fits that have been performed. As a regression is completed, an entry appears for it in the ranking chart (also on the CurveExpert main window). Clicking an entry in the ranking chart causes the plot preview area to generate a graph of that particular regression. If no regression is selected, the preview area simply shows the data present in the built-in spreadsheet. Note that the preview area will update every time the data set is changed, so it always reflects the latest changes to the data set.

The preview area also gives the capability to visualize regressions as they progress to convergence. At the end of each iteration, the preview is redrawn with the latest estimates of the parameters. The result is an animation of the curve fit as it attempts to achieve the best estimate of your data. For animations to occur, the Enable Animations menu item must be checked (see below). Note that the colors, axes, etc. of the plot preview area are determined by the default

plot setup, and cannot be customized directly. To open a plot that is customizable, double-click the plot preview, and a full graphing window will appear which can be manipulated as desired.

Right-clicking on the preview area shows a menu that allows you to customize the behavior of the preview. The items are listed below:

2.6.1 Enable Preview

Turn the preview area on or off. When on, the preview area has the behavior described in the documentation above. When off, the preview does not draw any graphs. If the data and/or curve fit takes a significant amount of time to regenerate, you might want to turn previews off. Default: On

2.6.2 Enable Animations

Turn preview animations on or off. When on, the preview area graphically shows the progression of a regression to convergence during a calculation. When off, no updates occur during a calculation. If fast performance is a key issue, then you might want to disable animations obviously, time is added to each iteration to redraw the preview area. Default: On

2.6.3 Open Plot

Open a graphing window for the plot shown in the preview area. This can also be accomplished by double-clicking the preview area.

2.6.4 Help

Get help on how to use the graph preview area.

2.7 While CurveExpert is Working

While the calculation proceeds, you can watch the status bar to see what regression is being calculated, the iteration number, and (most importantly) the current chi-square value of the model as the nonlinear regression algorithm progresses. The chi square lets you know how well the model is currently fitting the data, and it should be steadily decreasing, in an optimum situation. When CurveExpert is working, the toolbar and menus are disabled except for two new buttons on the far right hand side of the toolbar. Note that for interpolations and linear regressions, each button has the same meaning. They will both terminate the calculation immediately and no results will be shown. Also note that if background processing is disabled, these buttons will also be disabled; in effect, the ability to stop the calculation (in a graceful manner) is lost if background processing is disabled.

2.7.1 Yellow Check: (Keyboard Equivalent: ESCAPE)

Press this button to stop the computation for a particular curve fit. The parameters are accepted as they are; this button is handy when the standard error has dropped to a stable value, but CurveExpert is continually “tweaking” the values of the parameters to arrive within the tolerance bounds set in the File|Preferences dialog. Selecting this button will accept the current parameters as they are. The same operation can be accomplished by pressing Esc on the keyboard.

2.7.2 Red Stop Sign: (Keyboard Equivalent: SHIFT+ESCAPE)

Press this button to abort all calculations immediately. The same operation can be accomplished by pressing Shift+Esc on the keyboard.

2.8 Printing

2.8.1 Dataset

Choose “Print” from the main menu to generate a hardcopy of the data set. The data set will be printed in two columns of x and y data, and each point is labeled sequentially. If the printer is not set as desired, press the Setup button and configure the printer. Pressing OK will commence printing, and pressing the Cancel button will return you to the CurveExpert main window.

2.8.2 Graphs

To print a graph, right click on a graph, and select Print. Selecting this menu item prints a hardcopy of the current plot. A print dialog will be shown that will allow you to select printers and change the print setup. If the printer is not set as desired, press the Setup button and configure the printer. Pressing OK will commence printing, and pressing the Cancel button will return you to the CurveExpert graph window.

Note that you may select whether the printout will be in color or black and white by selecting the Black and White and Color radio buttons at the bottom of the dialog. If you have a black and white printer and select a color plot, then the plot will be grayscaled before it is sent to the printer.

WORKING WITH DATA

3.1 Introduction

CurveExpert supports a flexible text file format for easy data import and supports four types of data manipulation after the file has been read.

3.2 Opening and Saving Data Files

Data files in CurveExpert are simply two-column text files with the first column for the x data, and the second column for the y data. See Data Requirements for more information on the permissible format of a CurveExpert data file.

3.2.1 Opening Data Files

Simply choose the file that you want to use, and CurveExpert will read that file, examine it, and display the data in the main window for manipulation. If necessary, you may manually edit the data or use the data manipulation features included in the Data menu. You are now ready to perform some curve fits on the data using the *Apply Fit* or *Interpolate* menu.

Note that you may select the Read Only checkbox to prevent the file that you have opened from being written to at any time during a CurveExpert session. This keeps accidental saves from overwriting possibly valuable comments and data in the file that was initially opened. A protected/read-only file's name will appear in the status bar in red.

While CurveExpert reads a file, the status bar shows the state of the file being read. The first section of the status bar shows the file name, the second section shows the number of lines parsed, and the third section shows the number of data points read. Finally, the last section shows the action being taken (SCANNING or READING). On each file read, CurveExpert scans the file on the first pass, and reads the file on the second pass.

3.2.2 Saving Data Files

Data files will be written in two column form with no headers or footers. A tab separates the first and second columns. Note that if a data file that contains comments is read into CurveExpert and subsequently saved over the same file, the original comments are lost.

While CurveExpert saves a file, the status bar shows the state of the file being written to. The first section of the status bar shows the file name, the second section shows the number of lines written, and the third section shows the number of data points written. Since CurveExpert writes one data point per line, these numbers will be identical. Finally, the last section shows the action being taken (WRITING).

3.2.3 Data File Requirements

CurveExpert reads a plain ASCII text file, which consists of columns. You can produce these text files by hand, by spreadsheet, or any other method that can produce two columns of data. Note that the type of delimiter in the data file is irrelevant – the columns need only be separated by a nonnumeric character. Therefore, the data file may have tab, space, or comma delimited data.

Comments may be interspersed freely within the data files, with the only restriction being that comments must either be placed on a line by themselves or after any data on that line. Note again that there must be two numbers on a line for the data to be read correctly, and that if a line starts with a number, it is assumed to be a data line. If it starts with a nonnumeric character, the line is assumed to be a comment. Here are two examples, where the user wants all of the numeric data in the file to be read into CurveExpert:

Incorrect:

```
1.0
2.5
2.0
2.7
2.5
2.9
this is text 2.9
3.1
3.1
3.2
```

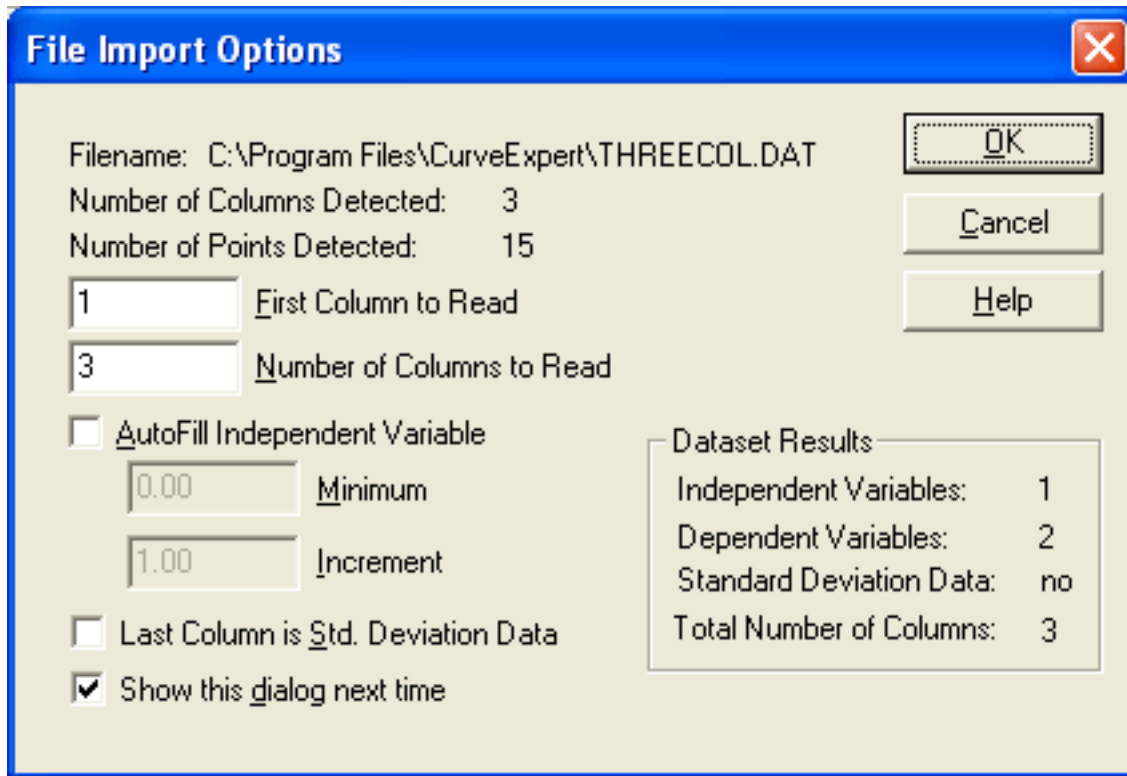
Correct:

```
1.0
2.5
2.0
2.7
2.5
2.9
2.9
3.1 this is text
3.1
3.2
```

When CurveExpert reads the first example, it assumes that the line with ‘this is text’ is a comment, and that line is discarded. In the second example, the line with the comment starts with a number and is read correctly.

3.3 Importing Data Files

CurveExpert allows a large amount of flexibility when reading data files. A column can be added to the data set, columns can be skipped. As soon as a file is selected to be read [note that for this to be true, the “Always Show Import Dialog” setting in Tools|Options must be checked] , the file import dialog is displayed, allowing the user to select the manner in which the file should be read. If the file is a simple two column data file, then the default settings in each field are adequate. The File Import dialog is shown below, and explanations of each item contained within it follow.



3.3.1 File Import Control Buttons

- OK: Press the OK button when all options have been set to your liking, and you are ready for the file to be read.
- Cancel: Press the Cancel button to abort the reading of a file. The previous data set will still be intact.
- Help: Press Help to get help on using the File Import dialog.

3.3.2 Summary Information

The full path to the file selected is stated first in the “Filename” field. Then, the number of data points and columns that CurveExpert has detected in the data file is specified. This information should be checked whenever a file is read to ensure that the file contains the expected data in a valid format. See Data File Requirements for more information on this topic.

3.3.3 User-Selectable File Options

Along the left side of the dialog, the user-selectable options are listed. Place the first column that should be read from the data file in the First Column to Read field. Specify the number of columns to read in the Number of Columns to Read field. The default values for these fields are to start on the first column and read all columns in the data file. If the AutoFill Independent Variable option is selected, then CurveExpert will automatically fill the independent variable (x) data column with a linear sequence. The first number in this sequence is specified in the Minimum field, and the increment is specified in the Increment field. Autofilling adds one column to the data set. The provided column will be the length specified in the “Number of Points Detected” information field. If the data set contains a column that should be interpreted as standard deviation data (uncertainty data) for the dependent variable (y), then select the Last Column is Standard Deviation Data checkbox. CurveExpert will use the last column read as the standard deviation data.

3.3.4 Dataset Results

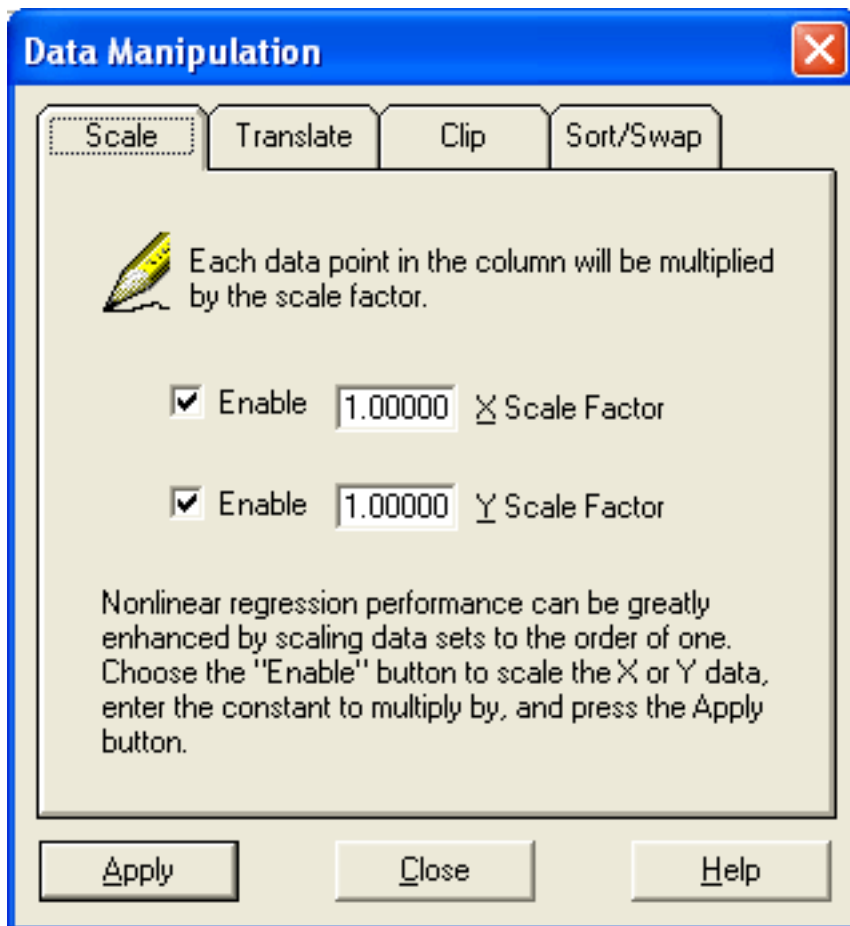
The region labeled “Dataset Results” gives the result of the combination of the read data file and the user-selectable options. The number of independent variables, dependent variables, and columns is specified here. Also, whether the data set contains standard deviation data is also given. If the file is not yet valid to be read, asterisks appear in these fields. Adjust the user-selectable items until the desired values appear in the dataset results. Note that CurveExpert can only accept one independent variable and one independent variable at this time.

3.3.5 Show This Dialog Next Time

If you are always reading simple two-column text files, then this dialog can be suppressed. Deselect the Show This Dialog Next Time checkbox before exiting this dialog to accomplish this. To regain access to this dialog at a later time, select “Tools|Options” and re-check “Always show file import dialog”.

3.4 Overview: Data Manipulation

Manipulating your data using the built-in features of CurveExpert can be very beneficial to curve fit performance. Please review the ‘Avoiding Floating Point Exceptions <fpe.html>’ section. All data manipulation options are organized into a dialog that is obtained via the DataManipulate menu.



At any time, you can always click on the numbers in the built-in spreadsheet and modify the numbers. Also, if you click on the row one past the end of the dataset (or press return while modifying the last row of a dataset), CurveExpert

will allow you to add a new data point by typing directly into the spreadsheet.

3.5 Data Scaling

To obtain the best curve fitting performance, you should scale your data to order one if the data is extremely large or small. Review the Floating Point Exceptions and Performance Issues sections for reasons why you might want to scale your data.

The Data Manipulation dialog box is the control that you use to multiply your data set by a scaling factor. After opening this box via the DataManipulate menu choice, press the “Scale” tab. To scale your data set, type in the desired scale factor into the scaling fields, and select the “Enable” buttons if you want a scale factor to be applied to the indicated column of data. Leaving the “Enable” checkboxes unselected means that CurveExpert will ignore the factor in the corresponding field. Finally, click the “Apply” button; the results of the scaling will immediately be apparent in the main window’s spreadsheet.

Note that after scaling the data, CurveExpert does not automatically get rid of the current curve fits (to prevent work that you have done from being lost). If you wish to clear the curve fit cache, then select Tools|Clear Curvefits.

3.6 Data Translation

Translating your data by a fixed additive factor can remove a known bias from the data set, as well as allow a better choice of a function to model the data. The Data Manipulation dialog box is the control that you use to add a factor to your data set. After opening this box via the DataManipulate menu choice, press the “Translate” tab.

To translate your data set, type in the desired translation factor into the scaling fields, and select the “Enable” buttons if you want translation to be applied to the indicated column of data. Leaving the “Enable” checkboxes unselected means that CurveExpert will ignore the factor in the corresponding field. Finally, click the “Apply” button; the results of the translation will immediately be apparent in the main window’s spreadsheet.

Note that after translating the data, CurveExpert does not automatically get rid of the current curve fits (to prevent work that you have done from being lost). If you wish to clear the curvefit cache, then select Tools|Clear Curvefits.

3.7 Data Clipping

Clipping (or pruning) your data means to remove any outliers that might contaminate a curve fit which is applied to your data set. This operation clips the data at a given x/y maximum and minimum. The Data Manipulation dialog box is the control that you use to clip outliers from the data set. After opening this box via the DataManipulate menu choice, press the “Clip” tab. Fill in the desired maximum and minimum for the x/y data, and press the Apply button. The results of the data clipping operation will be immediately apparent in the main window’s spreadsheet.

Note that after clipping the data, CurveExpert does not automatically get rid of the current curve fits (to prevent work that you have done from being lost). If you wish to clear the curve fit cache, then select Tools|Clear Curvefits.

3.8 Data Sorting

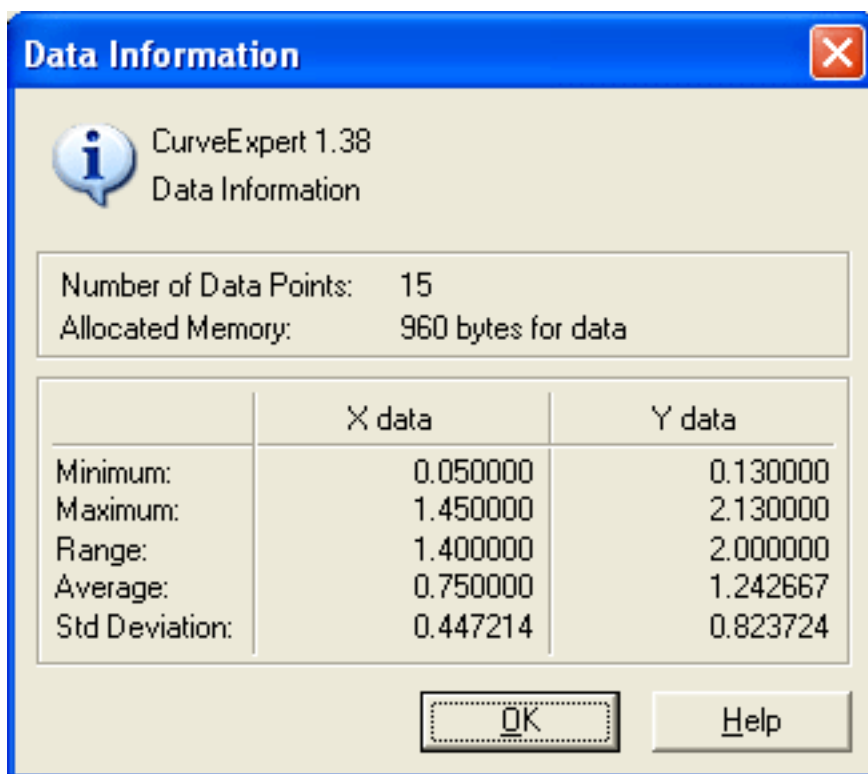
CurveExpert can sort the entire data set, using either the independent variable (x) or the dependent variable (y) as the key. When this operation finishes, all of your data will be in order of the key that was chosen. This operation is crucial for interpolation-type data fits. The Data Manipulation dialog box is the control that you use to sort the data set. After opening this box via the DataManipulate menu choice, press the “Sort” tab. To sort the data set, choose the appropriate sort key (x or y) and press the Apply button. The results of the sorting operation will be immediately

apparent in the main window's spreadsheet. If the independent variable (x) is used as the key, the result of the sorting will be a data set that is in order with increasing x. Conversely, if the dependent variable (y) is used, the data set will be in order with increasing y.

Note again that sorting the data set on x is crucial for interpolation type curve fits.

3.9 Getting Information about your Data

The data information dialog gives you statistical information about the current data set in CurveExpert; select DataInformation from the main menu. First, the number of data points in set is given, along with the memory requirement of storing that data. Note that the "memory allocated" statistic shows the total amount of memory allocated for data, curve fits, and open graphing windows.



Other statistics given by this dialog are the average, minimum, maximum, and standard deviation from the mean of both the x and y data. The average is a simple arithmetic average, which is defined as

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

The standard deviation from the mean describes the scatter of the data set around the average (higher standard deviation means more scatter):

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}}$$

3.10 Using the Built-In Spreadsheet

3.10.1 Cutting from Another Application and Pasting into the Spreadsheet

The spreadsheet supports cut/paste operations between CurveExpert and commercial spreadsheets and/or applications; this expedites the moving of data from a spreadsheet to CurveExpert, or from CurveExpert to the spreadsheet.

For best results, if you are pasting in an entire dataset into an empty spreadsheet, do not select any cells in the spreadsheet. Just select Edit|Paste, and CurveExpert will read the clipboard data into the built-in spreadsheet.

Warning: Not following the rules above might result in an undesired (0,0) pair inserted at the end of your dataset, and this data point WILL affect any curve fit calculations. So, watch for this occurrence, after you finish entering data and/or pasting in data. If the (0,0) pair is there, simply click on its row header and choose Edit|Cut to get rid of the unwanted point.

CurveExpert features a built-in spreadsheet for editing your data. This spreadsheet strongly resembles commercial packages such as Excel in operation and in appearance. To edit a particular point, simply click in its cell; a black selection rectangle will appear along with a blinking caret. The data in the cell may then be edited. To select a range of data, click the left mouse button and drag the mouse until the desired data is highlighted.

The headings across the top of the spreadsheet indicate the type of data in each column. The “X” column denotes the independent variable, the “Y” column denotes the dependent variable, and the “STD” column (if present) contains standard deviation data for each data point. The row headings down the left side of the spreadsheet specify the data point number. To select the entire spreadsheet, click the button in the upper left corner. To select an entire row or column, click the row or column header, respectively. The horizontal and vertical scrollbars at the bottom and left side of the spreadsheet allow you to scroll to different areas of the spreadsheet, if necessary. To navigate within the spreadsheet, use the arrow keys. Pressing the TAB or ENTER keys will intelligently navigate the selector to the next position in the spreadsheet in a Z-pattern (across, then down). Holding SHIFT while pressing either the TAB or ENTER keys will navigate in exactly the opposite manner. Note that you may use the TAB or ENTER to add new points to the data set. To add a column, click on the empty column directly to the right of the rightmost column. To delete a column, select an entire column and pick “Cut” or “Clear” from the Edit menu. Note that column removal is only available if the data set still has a dependent and independent variable; likewise, column addition is only available if standard deviation data is not yet present. Addition of a data point to the data set is as simple as picking the row below the bottommost row in the data set. All of the operations in the Edit menu will operate on the current selection in the spreadsheet. A multi-level undo/redo allows you to remove editing mistakes from the data set. Undoable actions include replacing a data point, range clears, and range inserts. Any action that can be undone can also be redone.

- Undo <item>: Undo the last edit in the built-in spreadsheet. The <item> clause identifies which action will be undone. Note that this undo mechanism is a multi-level; you may undo several actions in sequence. Items that can be undone include replacing a data point, typing in the edit control, clearing a range, or inserting a range. If no undo action is available, this option will be dimmed.
- Redo <item>: Redo the last undone action in the built-in spreadsheet. The <item> clause identifies which undoable action will be redone; this will always be the last action that was undone. Note that this redo mechanism is a multi-level; you may redo several undo actions in sequence. Any item that can be undone can be redone. If no redo action is available, this option will be dimmed.
- Cut: Remove the selection in the built-in spreadsheet and place it in the clipboard. This is functionally equivalent to selecting “Copy”, and then “Clear”. If no selection is valid in the spreadsheet, then this option is dimmed.
- Copy: Copy the selection in the built-in spreadsheet to the clipboard. If the edit control in the spreadsheet is active, only the selected text is copied. If a selection has been made in the spreadsheet, then the block is copied, in tab delimited format, to the clipboard. Note that this can be pasted directly into another region in the built-in spreadsheet, word processing applications, or commercial spreadsheet applications. If no selection is valid in the spreadsheet, then this option is dimmed.

- **Paste:** Paste the contents of the clipboard at the caret, if the edit control in the spreadsheet is active and only one value is detected on the clipboard. If more than one point is detected on the clipboard, then the selector serves as the insertion point for the paste operation. If the number of columns on the clipboard does not match the number of columns in the data set, then zeros are filled in the space to the side(s) of the pasted columns. If one row is selected at the time of the paste, the row is deleted, and the contents of the clipboard are pasted in its place. If a range selection is active in the built-in spreadsheet at the time of the paste, then the extent of the range must equal the number of rows and columns on the clipboard. The data is simply pasted over the selection if they match. If they do not match, then the paste operation fails. If no text is available from the clipboard, then this option is dimmed.
- **Clear:** Remove the selection in the built-in spreadsheet. If the edit control is active, only the selected text is removed. If a range selection is active, then it is erased from the data set, and the appropriate rows are shifted upward to take the selection's place.

MODELING YOUR DATA

4.1 Introduction

Given a set of data points, often called “observations,” a common need is to condense the data by fitting it to a model in the form of a parametric equation. This “model equation” can be anything that the user desires – it can range from a simple polynomial to an extremely complex model with many parameters. One important consideration in selecting an appropriate model is the underlying law that the data represents. Optimally, the model should be chosen to reflect that law so that the parameters in the curve fit have physical interpretation and meaning.

Data modeling may be performed by several techniques: interpolation, regression, or data smoothing. Interpolation and regression are the types supported by CurveExpert. Interpolation guarantees that the fitted curve will pass through each and every data point. Regression simply ensures that the “merit function”, which is an arbitrary function that measures the disagreement between the data and the model, is minimized. In this approach, the model parameters are adjusted until the merit function becomes as small as possible. One of the features of CurveExpert is the large number of nonlinear regression models that are available. This type of model is the most commonly used in real-world applications.

4.2 Linear Regression

All linear regressions are accessible in the first half of the “Apply Fit” menu.

Models which consist of a linear combination of a particular set of functions X_k are called linear models, and linear regression can be used to minimize the difference between the model and data. The general form of this kind of model is

$$y(x) = \sum a_k X_k(x)$$

where $X_k(x)$ are fixed functions of x that are called the basis functions, and a_k are the free parameters. Note that “linear” refers only to dependence of the model on the parameters a_k ; the functions $X_k(x)$ may be nonlinear. Minimization of the above linear model is performed with respect to the merit function

$$S(a) = \sum_{i=1}^n \left[y_i - \sum_{k=1}^{np} a_k X_k(x_i) \right]^2$$

The minimum of the above equation occurs where the derivative of S with respect to the parameters disappears. Substituting the linear model into this function, taking the first derivative, and setting this equal to zero gives the normal equations that can be solved directly for the parameters a_k .

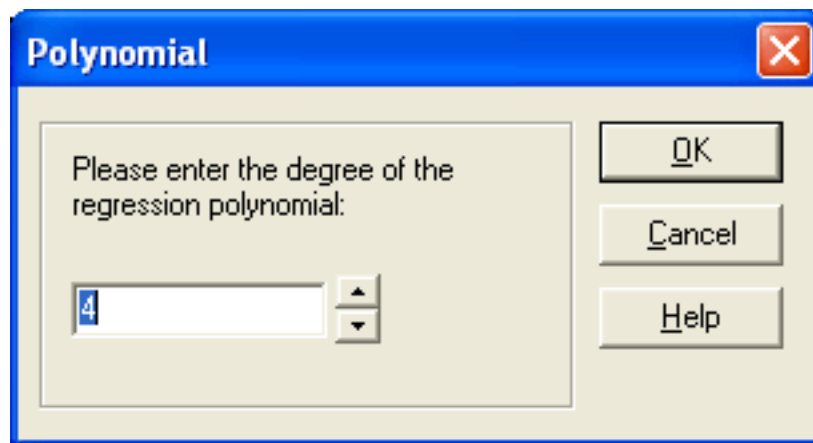
4.2.1 Linear Regression Models

In CurveExpert, all of the linear models implemented are linear combinations of x_n . These models are listed below:

- Linear Fit $y = a+bx$
- Quadratic Fit $y = a+bx+cx^2$
- Polynomial Fit $y = a+bx+cx^2+dx^3+...$

Setting the Polynomial Degree

For a polynomial regression, you must supply what degree of polynomial that you would like; CurveExpert will prompt you for that information. Polynomials usually tend to give decent curve fits to almost any data set, but higher order polynomials (as a rule of thumb, over 9) tend to oscillate badly. Also, polynomials offer no insight into the model that governs the data.



Note that you may not choose a polynomial degree greater than or equal to the number of points in the data set minus one. A polynomial degree equal to the number of points in the data set minus one is simply interpolation, and that type of curve fitting can be obtained via the Interpolation menu. Degrees greater than the number of data points cause an overconstrained problem, to which the solution is undefined. Note also that you may not choose a polynomial degree of one or two; use the built-in linear and quadratic curve fits instead. In short, the polynomial degree that you can enter in this dialog box is limited to $2 < \text{degree} < n-2$.

4.3 Nonlinear Regression

All nonlinear regressions are accessible in the last half of the “Apply Fit” menu. These models include the CurveExpert built-in models, as well as any user defined models that you have designed.

This program uses the Levenberg-Marquardt method to solve nonlinear regressions. This method combines the steepest-descent method and a Taylor series based method to obtain a fast, reliable technique for nonlinear optimization. This statement will be repeated later, but can bear stating at the beginning of this discussion: neither of the above optimization methods are ideal all of the time; the steepest descent method works best far away from the minimum, and the Taylor series method works best close to the minimum. The Levenberg-Marquardt (LM) algorithm allows for a smooth transition between these two methods as the iteration proceeds. In general, the data modeling equation (with one independent variable) can be written as follows:

$$y = y(x; \vec{a})$$

The above expression simply states that the dependent variable y can be expressed as a function of the independent variable x and vector of parameters a of arbitrary length. Note that using the ML method, any nonlinear equation with an arbitrary number of parameters can be used as the data modeling equation. Then, the merit function we are trying to minimize is

$$\chi^2(\vec{a}) = \sum_{i=1}^n \left(\frac{y_i - y(x_i; \vec{a})}{\sigma_i} \right)^2$$

where N is the number of data points, x_i denotes the x data points, y_i denotes the y data points, σ_i is the standard deviation (uncertainty) at point i , and $y(x_i; a)$ is an arbitrary nonlinear model evaluated at the i th data point. This merit function simply measures the agreement between the data points and the parametric model; a smaller value for the merit function denotes better agreement. Commonly, this merit function is called the chi-square.

Now, we will take a step back to provide a framework of optimization methods. From the area of pure optimization, which will not be explained here, two basic ways of finding a function minimum are a Taylor series based method and the steepest-descent method. The Taylor series method states that sufficiently close to the minimum, the function can be approximated as a quadratic. Without detailed explanation of that method, a step from the current parameters a to the best parameters a_{min} can be written as

$$\vec{a}_{min} = \vec{a}_{cur} + H^{-1} [-\Delta\chi^2(\vec{a}_{cur})]$$

where H is the Hessian matrix (a matrix of second derivatives). If the approximation of the function as a quadratic is a poor one, then we might instead use the steepest-descent method, where a step to the best parameters from the current parameters is

$$\vec{a}_{min} = \vec{a}_{cur} - c\Delta\chi^2(\vec{a}_{cur})$$

This equation simply states that the next guess for the parameters is a step down the gradient of the merit function. The constant c is forced to be small enough that a small step is taken and the gradient is accurate in the region that the step is taken.

Since we know the chi-square function, we can directly differentiate to obtain the gradient vector and the Hessian matrix. Taking the partial derivatives of the merit function with respect to a gives

$$\frac{\partial\chi^2}{\partial a_k} = -2 \sum_{i=1}^n \frac{y_i - y(x_i; \vec{a})}{\sigma_i^2} \frac{\partial y(x_i; \vec{a})}{\partial a_k}$$

To obtain the Hessian matrix, take the gradient of the gradient above (so that we have a matrix of partial second derivatives):

$$\frac{\partial^2\chi^2}{\partial a_k \partial a_l} = -2 \sum_{i=1}^n \left[\frac{1}{\sigma_i^2} \frac{\partial y(x_i; \vec{a})}{\partial a_k} \frac{\partial y(x_i; \vec{a})}{\partial a_l} - \frac{y_i - y(x_i; \vec{a})}{\sigma_i^2} \frac{\partial^2 y(x_i; \vec{a})}{\partial a_l \partial a_k} \right]$$

Now, for convenience, define the gradient vector and the curvature matrix as

$$G_k = -\frac{1}{2} \frac{\partial\chi^2}{\partial a_k} = \sum_{i=1}^n \left[\frac{y_i - y(x_i; \vec{a})}{\sigma_i^2} \frac{\partial y(x_i; \vec{a})}{\partial a_k} \right]$$

$$C_{kl} = \frac{\partial^2\chi^2}{\partial a_k \partial a_l} = \sum_{i=1}^n \left[\frac{1}{\sigma_i^2} \frac{\partial y(x_i; \vec{a})}{\partial a_k} \frac{\partial y(x_i; \vec{a})}{\partial a_l} \right]$$

[Note that the second derivative term in C will be ignored because of two reasons: it tends to be small because it is multiplied by $(y-y_i)$, and it tends to destabilize the algorithm for badly fitting models or data sets contaminated with outliers. This action in no way affects the minimum found by the algorithm; it only affects the route in getting there.] So, the Taylor series method (inverse Hessian method) can be written as the following set of linear equations:

$$\sum_{k=1}^{np} C_{kl} \delta a_l = G_k [1]$$

where NP is the number of parameters in the model that is being optimized. This linear matrix will be our workhorse for this method after some modification; it can be solved for the increments δa that, when added to the current approximation for the parameters, gives the next approximation. Likewise, we can substitute our “convenient” definitions into the steepest descent formula to obtain

$$\delta a_l = cG_l[2]$$

Neither of the aforementioned optimization methods are ideal all of the time; the steepest descent method works best far away from the minimum, and the Taylor series method works best close to the minimum. The Levenberg-Marquardt (LM) algorithm allows for a smooth transition between these two methods as the iteration proceeds.

The first issue in deriving the LM method is to attach some sort of scale to the constant c in the steepest-gradient method (equation 2). Typically, there is no obvious way to determine this number, even within an order of magnitude. However, in this case, we have access to the Hessian matrix; examining its members, we see that the scale on this constant must be $1/C_{ll}$. But, that still may be too large, so let’s divide that scale by a nondimensional factor (l) and plan on setting this much larger than one so that the step will be reduced (for safety and stability).

The second issue to formulate the LM method is noting that the steepest-descent and Taylor series methods may be combined if we define a new matrix M_{ij} by the following:

$$M_{ii} = C_{ii}(1 + \lambda), i = j$$

$$M_{ij} = C_{ij}, i \neq j$$

This matrix combines equations [1] and [2] into a convenient and compact form. So finally, we have a means of calculating the step δa in the parameters by the following system of linear equations:

$$\sum_{k=1}^{np} M_{kl}\delta a_l = G_k[3]$$

Note that when l is large, the matrix M is forced to be diagonally dominant; consequently, the above equation is equivalent to the steepest descent method (equation 2). Conversely, when the parameter l goes to zero, the above equation is equivalent to the Taylor series method (equation 1). So, we vary l to switch between the two methods, continually calculating a parameter correction δa that we apply to our most recent guess for the parameter vector. The steps that are taken in the LM algorithm are as follows:

1. Compute $\chi^2(a)$
2. Pick a conservative value for λ (0.001 in CurveExpert)
3. Solve the linear equations (equation 3) for δa
4. Evaluate $\chi^2(a + \delta a)$
5. If $\chi^2(a + \delta a) \geq \chi^2(a)$, increase λ by a factor (10 in CurveExpert) and go back to step [3].
6. If $\chi^2(a + \delta a) < \chi^2(a)$, decrease λ by a factor (10 in CurveExpert), correct the parameter vector by $a = a + \delta a$, and go back to step [3].

Iteration is stopped when $|\chi^2(a + \delta a) - \chi^2(a)| < \epsilon$. This tolerance ϵ is the value specified in the FilePreferences dialog.

4.3.1 Setting Initial Guesses

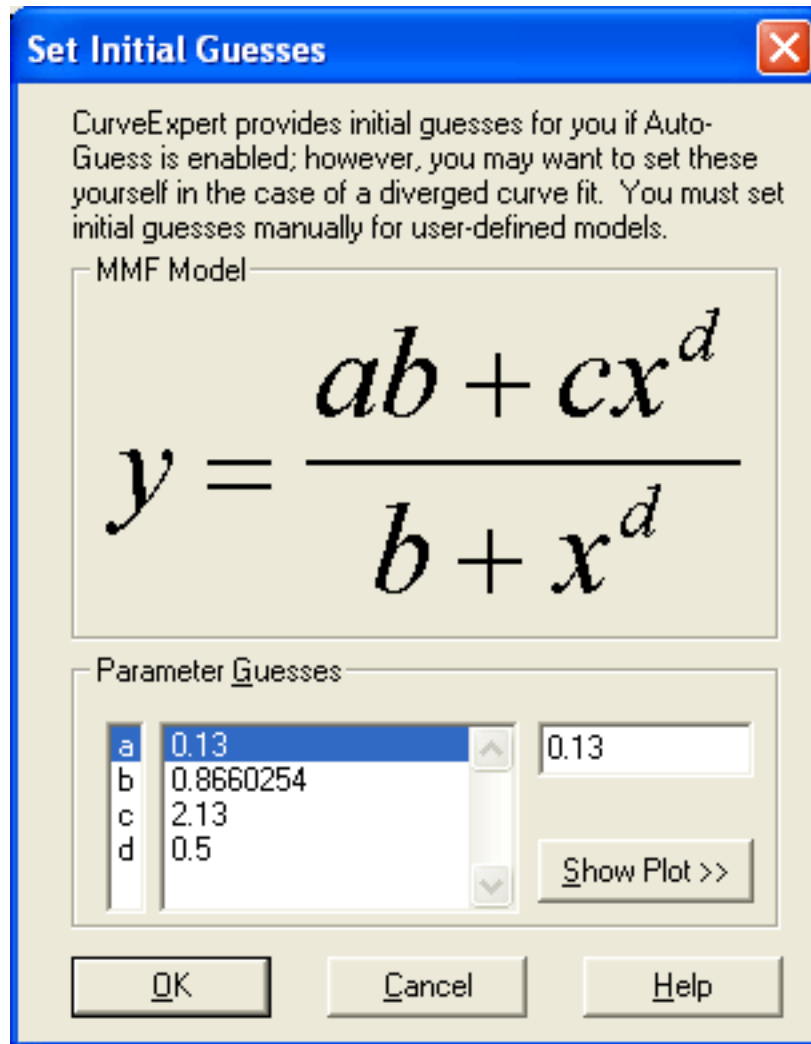
All nonlinear regression algorithms require initial guesses to commence the search for the optimum regression parameters. As long as AutoGuess is enabled, CurveExpert normally will calculate initial guesses for all built-in nonlinear regressions (linear regressions do not require initial guesses) for you automatically. However, in the case of a poorly

performing curve fit, you might want to set these initial guesses manually. Also, you must set these initial guesses manually if you are using a user-defined nonlinear regression model.

In many cases, the physical interpretation of the individual parameters in a regression model may not be clear; hence, an inordinate amount of time can be spent simply trying to obtain appropriate initial guesses. To facilitate this process, CurveExpert provides a dynamic view plot which adjusts itself every time you adjust the initial parameters. This allows you to see the behavior of the curve, and adjust the parameters by hand in such a way that the curve is fairly close to the data points. To display this plot, simply click the Show Plot button. Subsequently, any changes you make to the initial guesses will be shown immediately on the dynamic view plot.

Hint: The plot drawing speed is increased automatically by CurveExpert; it increases the “Resolution” parameter. To increase speed even more, right click on the plot window, choose “Attributes” off of the popup menu, and increase the “Resolution” parameter further.

The initial guess window is shown below:



Initial Guess Control Buttons

- OK: Press OK to accept the initial guesses entered in the Parameter Guesses section and let CurveExpert perform the nonlinear regression.
- Cancel: Press Cancel to abort the regression and return to the main CurveExpert window.

- Help: Press Help to get CurveExpert help on setting initial guesses for a nonlinear regression.

Model Display

To ensure that the meaning of the parameters is clear, this display shows you the equation for the model that you are providing initial guesses for.

Parameter Guesses

- Guess Listbox/Edit Control: Enter an initial guess for each parameter in the corresponding field. Refer to the Model Display for the use of each of the parameters in your nonlinear regression model. The number of parameters required for you to enter will match the number of parameters that are in the model, and you may scroll up or down as necessary.
- Show Plot: Click this button to dynamically view the changes that you make to the regression parameters. This will greatly assist you in obtaining good initial estimates for the parameters.

4.3.2 Regression Weighting Schemes

CurveExpert implements five different types of regression weighting schemes, so that you have a range of methods to choose from. Weighting allows the user to define how much influence each data point will have on the final set of fitted parameters; a large weight denotes that a particular point influences the parameters more (pulls the curve toward itself strongly), and a small weighting on a data point decreases its effect on the parameters. Select “Tools\Weighting Schemes” to pick the weighting to use for a particular curve fit. Only one can be selected at a time; the currently selected scheme is denoted by a check in the menu mentioned above.

The particular type of weighting used in a curve fit is reported in the Graph Information page of the Graph Properties dialog. Also note that weighting only has influence on regressions; any item from the Interpolation menu will not be affected by weighting, since (by definition) interpolations must go through each data point. The types of weighting offered by CurveExpert are the following:

None

The regression is performed as is, and no weighting is applied during the curve fitting process. This is the most common type of regression analysis; each point has equal influence on the final converged curve. No weighting is equivalent to weighting each data point by 1.0.

By Uncertainty

The regression is weighted based on the supplied uncertainties in each data point. The uncertainties (in the STD column in the built-in spreadsheet) are interpreted as one standard deviation around the data point. So, during the regression, the calculation will be weighted by

$$\frac{1}{\sigma_i^2}$$

The data points with high uncertainty will contribute less to the converged curve, and data points with low uncertainty will contribute more. This menu choice is not available if CurveExpert detects that no standard deviation data is present in the built-in spreadsheet. Also, if a file that has standard deviation data is read in, CurveExpert automatically selects uncertainty weighting for you.

By $1/x$

This type of weighting is rarely used, but weights each data point by the inverse of its x value. Be wary to *not* use this type of weighting when negative x data exists in the data set.

By $1/x^2$

This type of weighting is rarely used, but weights each data point by the inverse of its x value squared. The farther the point is from $x=0$, the less influence it will have in the final parameter set. This type of weighting is sometimes used when it is known that the uncertainty is a percentage of the independent variables value.

By $1/y$

This type of weighting is rarely used, but weights each data point by the inverse of its y value. Be wary to *not* use this type of weighting when negative y data exists in the data set.

By $1/y^2$

This type of weighting is sometimes used when uncertainties in each data point are not known, but when the uncertainty tends to be a percentage of the measured y value. The larger the y value is, the less weight a particular point has.

4.4 Spline Interpolation

CurveExpert implements four types of spline interpolation: linear/quadratic/cubic splines, and tension splines.

Warning: The data must be sorted for spline interpolations to complete successfully. To sort your data, choose DataManipulation from the main menu and select the “Sort” tab.

4.4.1 Linear Splines

Linear splines are the simplest form of spline – they simply approximate the range between data points by a straight line. The purest definition of a linear spline is that it simply ensures that each discrete spline matches the data points. These splines are simply a “connect-the-dot” approach that can be useful in increasing the resolution of a given table of values by using linear interpolation between each point. However, if this is the application, cubic splines provide a much more accurate representation between the data points.

Note also that the data must be sorted for the spline operation to complete successfully. To sort your data, choose DataManipulate and select the Sort tab.

4.4.2 Quadratic Splines

Quadratic splines interpolate the range between data points by a 2nd order polynomial (quadratic). Quadratic spline curve fitting ensures that each spline is equal to the data points, and that the 1st derivative of the splines are continuous, even at the knots. Quadratic splines perform better than simple linear splines, but are often characterized by “swinging” too high or too low past the data points.

Quadratic splines must have one free condition specified for the first spline. CurveExpert follows common practice, which is to specify that the 2nd derivative of the first spline is zero – this has the effect of connecting the first and second points with a straight line.

The implementation of quadratic splines used in CurveExpert requires a good deal of memory and crunching time (even more than that of the cubic spline). Do not use this curve fit on an excessive number of data points! Use caution when applying this class of splines.

Note also that the data must be sorted for the spline operation to complete successfully. To sort your data, choose DataManipulate and select the Sort tab.

4.4.3 Cubic Splines

Cubic splines interpolate the range between data points by 3rd order polynomials (cubics). Cubic spline curve fitting ensures that each spline is equal to the data points, the 1st derivatives are continuous at the knots, and the 2nd derivatives are continuous at the knots. Since discontinuities in 3rd and higher derivatives cannot be visually detected, cubic splines have an “aesthetically pleasing” appearance. In fact, cubic splines mimic the behavior of drafting splines used in olden days. Various end conditions can be used for cubic splines. Common practice is to use “natural” end conditions by specifying that the 2nd derivative of the end splines are zero at the endpoints. This gives the visual appearance of “straightening out” at the ends, which is the behavior of a physical drafting spline. This behavior is the type implemented in this version of CurveExpert. Note also that the data must be sorted for the spline operation to complete successfully. To sort your data, choose DataManipulate and select the Sort tab.

4.4.4 Tension Splines

Tension splines behave similarly to cubic splines, with the addition of a parameter (the tension) which specifies how “tightly” the spline forms to each knot. A low tension parameter will yield a spline similar to a cubic spline; whereas at high tension, the spline will approach a linear spline. The Set Tension dialog box allows you to set this parameter before the tension spline curve fit is performed. You can experiment with the tension to obtain the visual quality that you want in your spline.

Note that there is no easy way to specify a tension spline, except by creating a table of its values. It is also very ill-behaved – expect a few complaints from CurveExpert about floating point errors. To alleviate this type of error, make sure that your data is scaled to an order of one, and give intermediate tension parameters (0.1 to 20).

Note also that the data must be sorted for the spline operation to complete successfully. To sort your data, choose DataManipulate and select the Sort tab.

4.5 Lagrangian Interpolation

Warning: The data must be sorted for spline interpolations to complete successfully. To sort your data, choose DataManipulation from the main menu and select the “Sort” tab.

Lagrangian interpolation is simply interpolating the data set with a polynomial. The order of the polynomial is determined by the number of points in the data set – the polynomial must have enough degrees of freedom to fit the data set at every point. Therefore, in a data set with n points, the interpolating polynomial will be of $n-1$ degree.

Lagrangian polynomials are particularly well suited for data sets with a low number of data points (8 or less). Preferably, the data points should not have too much scatter. Like all polynomials, the Lagrangian interpolating polynomial exhibits more oscillatory behavior as the degree increases. Therefore, the use of Lagrangian interpolations should be avoided if the number of data points is large. Currently, CurveExpert limits the use of this type of interpolations to data sets with 13 points or less. As the name implies, extrapolating beyond the scope of the data points with an interpolating polynomial can be very detrimental. Interpolations only provide accuracy between the data points – using them to obtain y data outside of the range of the data points must be done very carefully.

Note also that the data must be sorted for the interpolation to complete successfully. To sort your data, choose DataManipulate and select the Sort tab.

4.6 Using CurveFinder

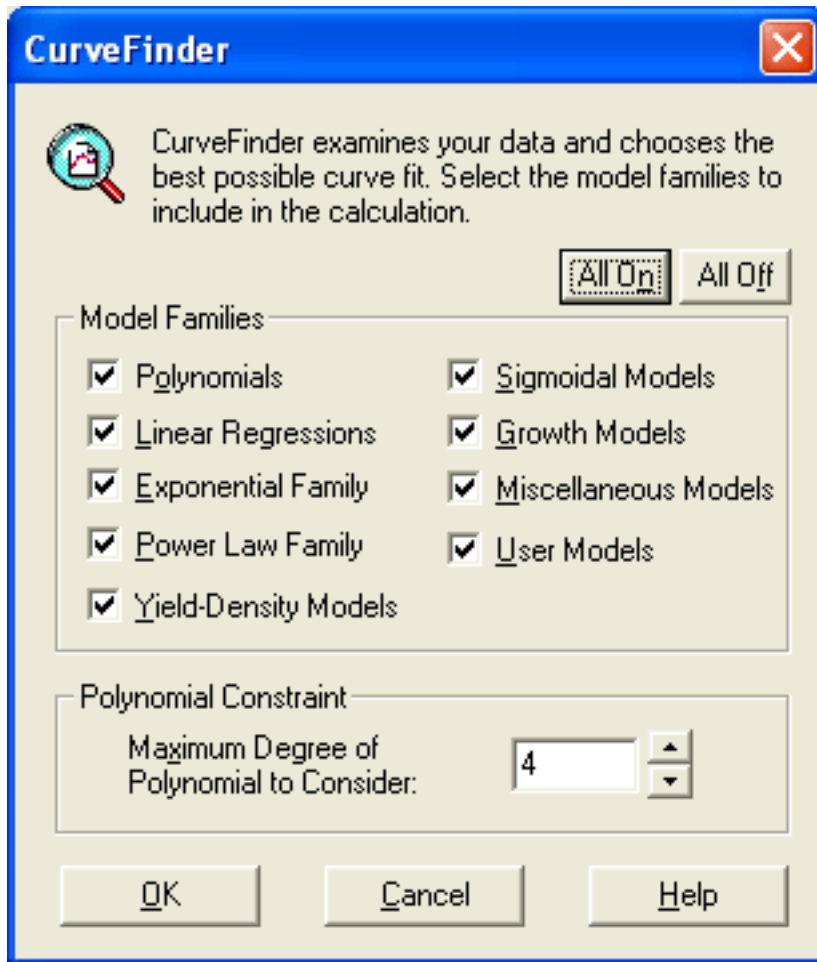
Using the built-in CurveFinder essentially does all of the hard work for you. You provide the data file, and the CurveFinder does the rest. It sifts through every possible curve fit, ranks the fits from best to worst, and presents you with the best one. Choosing CurveFinder from the Data menu will invoke a dialog box that will ask you for input on which families of regression models to search. Also, you may use the “all on” and “all off” buttons to check or uncheck all of the families at once. See Regression Models and Linear Regression for details on each built-in model. If the polynomial family is included for consideration, you must specify the maximum degree of the polynomial that CurveFinder will consider in the “Polynomial Constraint” area. Of course, if the polynomial family is not included, the polynomial constraint will simply be ignored.

Once you are finished selecting which model families CurveFinder should search, press the OK button. The calculation will then proceed, and each curve fit is ranked according to its standard error and correlation coefficient (see the Data Modeling with CurveExpert section for the appropriate equations), and the best fit is then displayed in the graphing window. That's all there is to it!

Note that CurveFinder considers ONLY regressions (linear and nonlinear) in its search for the best fit. Further, as described above, you can select only certain families to be considered in the search for the best curve fit. See Regression Models for a description of the model families. Interpolations are guaranteed to pass through each of the data point, so an evaluation of their performance relative to a regression-type curve fit is rather difficult. Therefore, interpolations are excluded from the CurveFinders search; however, you may still decide to try these curve fits manually when CurveFinder is done.

This dialog will automatically generate initial guesses for the nonlinear regression models, whether AutoGuess is enabled or not. If a particular curve fit diverges or generates a floating point exception, it is excluded from the calculation process. In the case that an important curve fit in the group is not calculated, you must select that curve fit manually from the Apply Fit menu and adjust the initial guesses, if necessary.

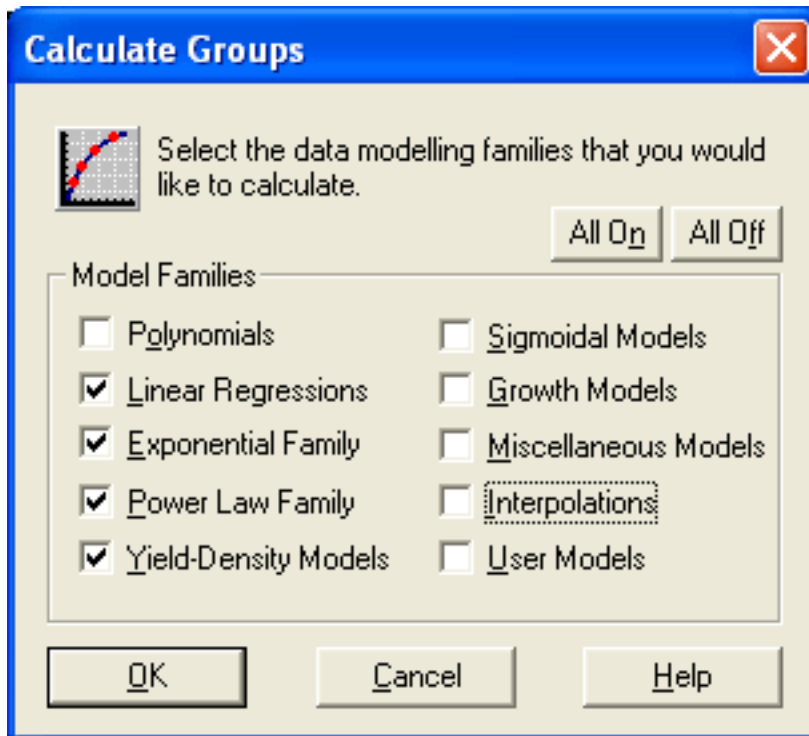
The CurveFinder dialog is shown below:



4.7 Calculating Groups of Models

The Calculate Groups dialog allows you to automate the process of finding the right data model. This option is most helpful when you have data that follows a certain trend that matches a particular model family or families. Using Calculate Groups, you can select those families that you would like to see, and CurveExpert will automatically calculate all of the curve fits in that group, rank them on the main windows ranking chart, and open a graphing window for each of the curve fits. See Regression Models for a description of the model families.

This dialog will automatically generate initial guesses for the nonlinear regression models, whether AutoGuess is enabled or not. If a particular curve fit diverges or generates a floating point exception, it is excluded from the calculation process. In the case that an important curve fit in the group is not calculated, you must select that curve fit manually from the Apply Fit menu and adjust the initial guesses, if necessary. The Calculate Groups dialog is shown below:



4.7.1 All On

Press this button to select all of the families listed in this dialog.

4.7.2 All Off

Press this button to deselect all of the families listed in this dialog.

4.7.3 Model Families

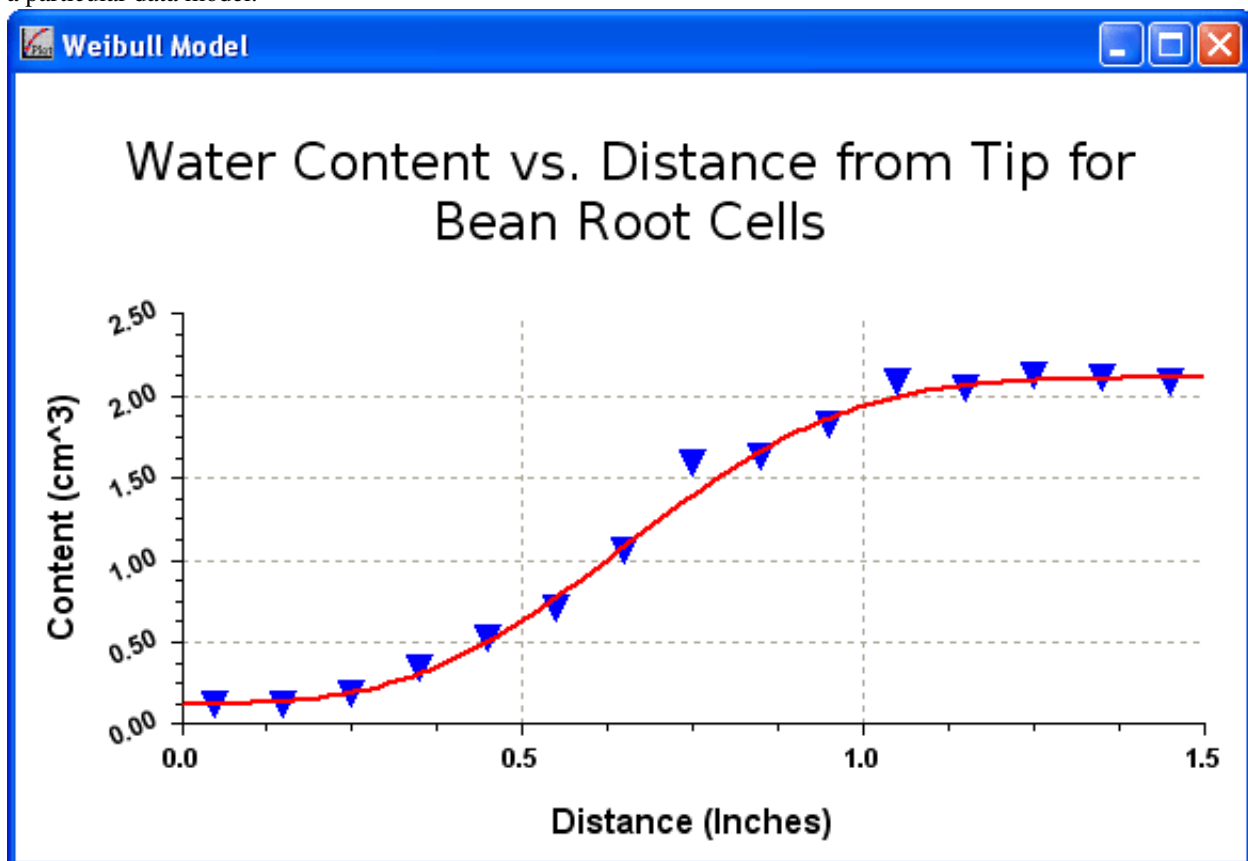
Click the checkboxes beside each model family to include them in the calculation. Families that are not marked by an X will not be included.

4.7.4 Calculate Groups Control Buttons

- **OK:** Click the OK button to signify that you are finished choosing regression families, and that CurveExpert should begin calculating the families that you selected.
- **Cancel:** Click the Cancel button to abort and not calculate any curve fits.
- **Help:** Click the Help button to find out how to use the Calculate Groups dialog.

GRAPHING

CurveExpert has extensive graphing capability. This capability gives you immediate feedback on the performance of a particular data model.



5.1 Graphing Window Overview

CurveExpert allows you to open many graphing windows at one time so that you may compare different curve fits. Note that changes that you might make to a particular graphing window will affect only that window. See the Graph Properties dialog to see how to apply your changes to all plots, as well as saving your customizations for future use.

Each graph has a consistent user interface. Press the right mouse button to obtain the graphing menu, which gives you access to all of the functions available for plots, such as customizing the plot, obtaining model information, analyzing

the model, or printing the plot. Zooming and panning are controlled by dragging a rectangle across the graph pane and/or pressing the arrow keys. Any plot can be autoscaled by pressing **Shift+A**.

The most vital piece of information that you might want to access are the values of the model parameters. Press the Info button, select *Information* from the graphing menu (by clicking the right mouse button), or press Ctrl-I to access this (as well as other) information. Note that if the graph title is displayed, the Info button is not available; choose *Information* from the graphing menu or press Ctrl-I to access the curve fit information.

Graphs can be customized by selecting *Graph Properties* from the graphing menu.

It is possible that CurveExpert will not be able to properly draw your graph. For example, the y-extents might be negative on a graph that is being drawn with log axes. Or, the extents might be set such that a singularity in the model exists somewhere in the domain. CurveExpert will notify you of these occurrences with a error message printed on the graphing surface. In these cases, reset the graphing extents using the XY Axis Properties section of the Graph Properties dialog (accessed via the popup menu described below).

5.2 Graphing Controls

- **Zoom:** You may zoom in on a region on the graph by holding down the left mouse button, dragging over the area of interest from top left to bottom right, and then releasing the button. Likewise, zoom out by holding down the left mouse button, dragging from bottom right to top left, and then releasing the button. Zooming out will cause the entire graphing area to be enclosed in the rectangle that was made by dragging “backwards.” Note: Zooming is not available on log or semilog plots. Also note that any zoom that starts or ends outside of the graphing pane will be ignored.
- **Pan:** To pan the graph in any direction, press the arrow keys. This will pan the graph in the direction that you choose by a magnitude of the major tick length. Holding SHIFT while pressing the arrow keys gives you finer control by only panning one minor tick length. Note: panning in the direction of a log scaled axis is not supported.
- **AutoScale:** Press SHIFT-A to autoscale the graph. Autoscaling guarantees that every data point on the plot will be shown, which is the default for a newly opened graphing window.

5.3 Graphing Menu and Shortcuts

By pressing the right mouse button or by pressing Ctrl-M in any graphing window, you can pick from the graphing menu (keyboard equivalents are listed also):

- [Ctrl-W] Find Main Window
- [Ctrl-G] Graph Properties
- [Ctrl-I] Information
- [Ctrl-T] Generate Table
- [Ctrl-L] Analyze
- [Ctrl-C] Copy
- [Ctrl-P] Print
- [Ctrl-X] Close/Exit
- F1 Help

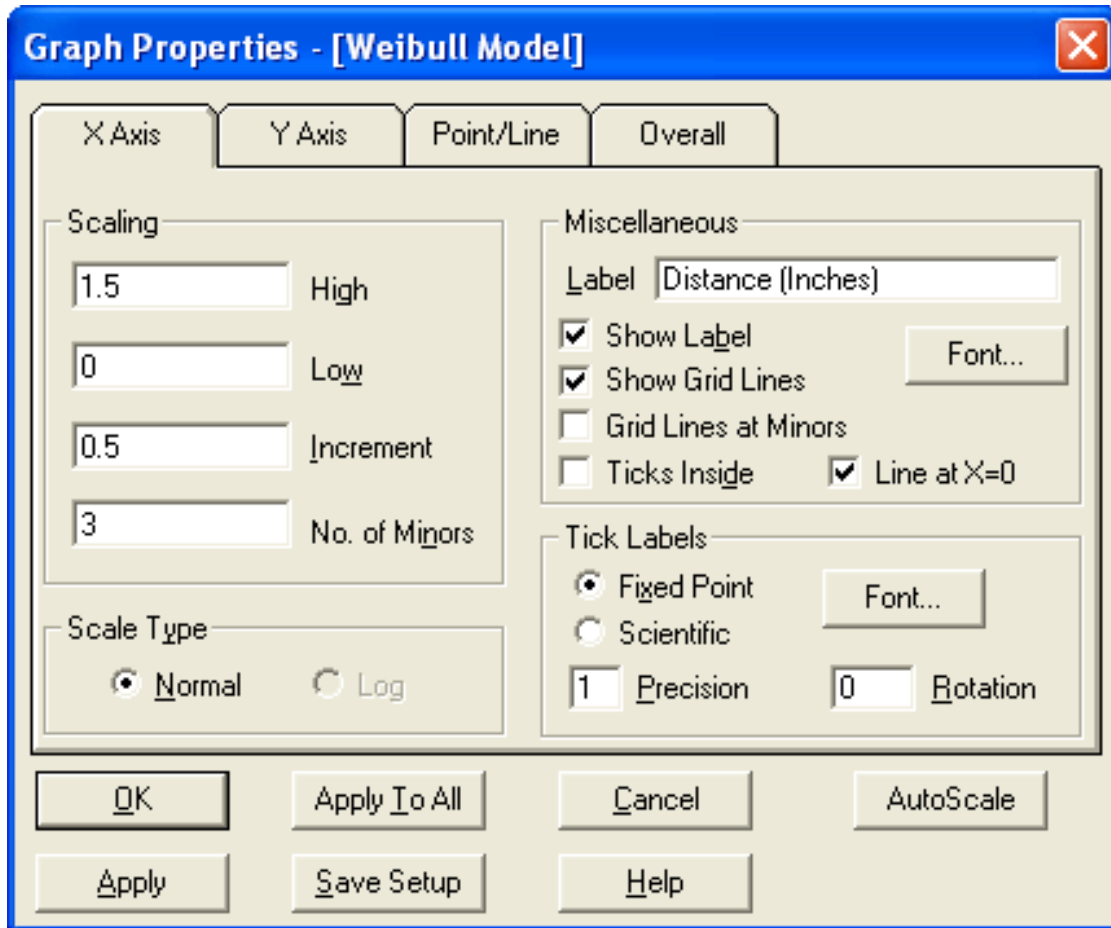
5.4 Graph Properties

The Graph Properties dialog allows you to modify all of the scaling and characteristics of the current graph.

5.4.1 Graph Property Control Buttons

- **OK:** Press OK to apply all of your changes to the current CurveExpert graph and quit the dialog.
- **Apply:** Press Apply to “try out” your graph property changes. The graph will be updated immediately, but you may still press Cancel to discard your changes.
- **Apply to All:** Press Apply to All to immediately apply all of the current graph properties to the current plot and all open plots. All attributes as well as the scaling for the current plot are copied to all other plots; hence, all graphs will have exactly the same view, colors, line thicknesses, etc. This action CANNOT be undone by pressing the Cancel button.
- **Save as Default:** If you would like to save your graph attribute and scaling setup permanently, click this button. This information will then be used on every plot that is drawn by CurveExpert.
- **Cancel:** Press Cancel to discard any changes that you have made to the graph properties and quit the dialog.
- **Help:** Press Help to get help on the currently selected tab of the Graph Properties dialog.
- **Autoscale:** Press this button to have CurveExpert automatically set the graph axes and tick mark increments. This is the default view for a newly opened graphing window, and ensures that all data points will be shown on the graphing window. This button is only available on the X and Y axis tabs.

If an invalid entry has been entered in any of the tabbed pages (for example, entering 0 for the Resolution parameter on the Overall page), then the offending entry will be selected, and the computer's speaker will sound.



5.4.2 X and Y Axis Properties

The X axis and Y axis property sheets modify all graph properties that are associated with the x/y axes, such as the scaling, tick marks, and axis labels.

Scaling

Axis scaling is determined by the settings in the High and Low edit fields. As the names imply, the graph will be displayed over the range set by the Low value and the High value. Major tick marks will be placed at increments specified in the Increment edit field. Also, if the corresponding Show Grid Lines box is checked, grid lines will be displayed at each major tick. The number of minor ticks displayed between the major ticks is specified in the No. of Minors field.

Scale Type

The Scale Type section allows you to specify whether the axis should be scaled on a normal (linear) scale or a logarithmic scale. To use normal (linear) axes, select the Normal radio button. Likewise, to use logarithmic axes, select the Log radio button.

Using logarithmic scales has limitations – a plot showing zero or negative numbers on the axis in question cannot be scaled logarithmically. Of course, you can always reset the High and Low values to positive values to be able to scale the axis logarithmically. Note that the High, Low, and Increment fields are not represented exactly when using logarithmic scaling – the closest multiple of ten is chosen for the High and Low entered, and multiples of the next lower major tick are used for the minor increments.

Tick Labels

The tick labels for both axes can be customized by using the Tick Labels section of the Graph Properties dialog. The following options are available:

- **Fixed Point:** Select this radio button if you would like the numbers at every major tick to have fixed point format. The number of decimal places displayed is set in the Precision edit box below this radio button.
- **Scientific:** Select this radio button if you would like the numbers at every major tick to have a scientific/engineering format. The number of decimal places displayed is set in the Precision edit box below this radio button.
- **Precision:** The number in this field indicates the number of decimal places that will be displayed on the tick labels along the axes.
- **Rotation:** The tick labels can be rotated from 0 to 360 degrees. Enter the desired number of degrees to rotate these labels in the Rotation field.
- **Font:** The font used for the tick labels can be modified by clicking this button. Any Truetype font and any color can be chosen; refer to the Changing Graph Fonts section for details.

Miscellaneous

Axis properties that do not fall under one of the above categories are documented below.

- **Label:** Type in the axis label that is desired for the axis in question.
- **Font:** The font used for the axis label can be modified by clicking this button. Any Truetype font and any color can be chosen; refer to the Changing Graph Fonts section for details.
- **Show Label:** Select this checkbox if the axis label should be displayed. If this box is not checked, then the axis label will not be drawn.
- **Show Grid Lines:** Select this checkbox if grid lines (associated with the axis in question) should be shown at every major tick mark. If this box is not checked, then the grid lines will not be drawn.
- **Grid Lines at Minors:** Normally, grid lines are drawn only at the major ticks. Checking this button will cause the grid lines to be drawn at every minor tick also.
- **Ticks Inside:** Select this checkbox if all major and minor ticks should be drawn inside of the graph pane. If this box is not selected, all ticks will be drawn outside of the graph pane.

5.4.3 Point and Line Properties

The Point/Line property sheet modifies all graph properties associated with the appearance of the data point markers and the curve representing the model prediction.

Data Point Properties

- **Show points:** This checkbox will determine whether data points are drawn on the plot or not. For residual plots, this choice will read Show Residuals as Points. When this checkbox is selected, all residuals will appear as points on the graph. When left unselected, the residuals will be drawn as bars. Note that the properties of the residual points and/or bars can be controlled by the following properties, just as normal data points can.
- **Fill points:** This checkbox determines whether the points will be drawn as a solid color or an outline. For example, if the 'circle' symbol is chosen and the 'Fill points' options is checked, a solid disk will be drawn at each data point. The color used to fill the point is the same as the color used to draw the outline.
- **Show Error Bars:** This checkbox determines whether error bars will be shown on the data points, if standard deviation data is present in your data set. If this option is checked, then a rectangular error bar will be shown on each point. Note that this option will be disabled if your data set does not contain standard deviation data for each data point.
- **Error Bar Scale:** By default, the length of the error bar is the value of the standard deviation for the point is drawn on (this is unscaled, or scale=1.0). If you would like to scale the length of the error bar, then change the value of this item. A common use for this is to show a (theoretical) 95% confidence bound for the data point; this corresponds to a scale factor of 1.96, based on the Gaussian distribution. This item will not be enabled unless standard deviation data is present in the current data set.
- **Point size:** Enter the size of the symbol that will denote the data points, in pixels. This size may not be negative. The up/down buttons beside the edit control may also be used to change the size.
- **Point color:** The current point color is displayed beside this button. To change this color, press this button and choose a new color.
- **Symbol:** Press the 'Symbol' button to obtain a list of possible markers that can be used for each data point. The currently selected symbol will have a checkmark beside it on the symbol menu. Currently, the supported symbols are the circle, square, triangle, inverted triangle, hourglass, plus, and cross.

Line Properties

- **Line Width:** Enter the width, in pixels, of the curve that will be drawn as the curve fit. This size may not be negative. You may also use the up/down buttons beside the edit control to change the size. Note that the Line Width setting is only valid for solid lines; if any other style of line is chosen, then the width of the line will be one.
- **Resolution:** Enter the stepsize, in pixels, that the curve will be updated. The curve is made of straight line segments that appear as a curve when the resolution is small enough; however, the finer the resolution (lower number) the slower the curve will be drawn. For optimum performance, set the resolution to the highest number possible without degradation of plot quality.
- **Line Color:** The current line color is displayed beside this button. The line color is used to draw the curve fit line. To change this color, press this button and select a new color.
- **Line Style:** Press this button to display a menu with the available line types. Note that the Line Width setting is only valid for solid lines; if any other style of line is chosen, then the width of the line will be one.

Sample

This area of the Point/Line property dialog shows a sample of the currently selected point and line combination. Note that the background color is the currently selected Fill Color from the Overall tab. Also note that this sample is dynamically updated as you change the settings on this tabbed page.

5.4.4 Overall Graph Properties

The Overall property sheet modifies all properties associated with the plot as a whole, such as the title, graph pane, graph box, and graph coloring.

Show Header Instead of Title

Selecting this checkbox will display the header instead of the graph title. The header consists of the information button (which you can press to invoke the Information Dialog), directions concerning how to invoke the graphing menu, and the value of the standard error and the correlation coefficient for the current model. If this checkbox is left unselected, the title that is defined by the user in the Graph Title field will be centered in the window/page in the title font.

Graph Title

Enter the text for the graph title in this field. The string may be up to 255 characters long; if the title is too long for one line, then room will be made for it on the plot and it will be broken up into multiple lines. Note that the title will not be displayed unless the Show Header Instead of Title checkbox (above) is left unselected.

Graph Title Font

Press this button to select a font for the title. See also Changing Graph Fonts for further details concerning how to use the font selection dialog.

Graph Pane

- **Border Buttons:** Select or deselect the pane border checkboxes to show or hide the lines around the graph pane.
- **Lock Frame Color:** If this button is checked, then all text drawn on the graph will automatically follow the selected Frame Color (see below). When this button is unchecked, the text colors can be individually selected through the Font selections in the Graph Properties dialog. Note that the text colors selected through the font dialogs will not be used unless this checkbox is unselected.
- **Fill Color:** The current fill color is displayed beside this button. The fill color is the color that will be used to shade in the background of the graph. To change this color, press this button and choose a new color.
- **Grid Color:** The current grid color is displayed beside this button. The grid color will be used to draw the dotted grid lines that are drawn at every major tick. To change this color, press this button and choose a new color.
- **Frame Color:** The current frame color is displayed beside this button. The frame color is used to draw the border around the graph as well as the tick marks. To change this color, press this button and choose a new color.
- **Outside Fill:** The current outside fill color is displayed beside this button. The outside fill color is the color that is painted outside of the graphing frame. Since the axes labels and various other printed items are drawn in this area, this color should not be the same as the Frame color. Remember that if the font color for the title, axes labels, etc. are the same color as the Outside Fill color, then they will be invisible! See the Lock Frame Color switch above for more information concerning this.

Graph Box

- **Show Graph Box:** To show or hide the graph box, select or deselect this checkbox, respectively. The graph box is the box that encloses the entire graph. This box is not very visible for on-screen plots; therefore, this option is most relevant for printed or copied plots.

- **Box Width:** Enter the line width, in pixels, that the graph box will be drawn with. Click the corresponding up/down arrows to increase or decrease this line width.
- **Box Color:** Click this button to select the color with which the graph box will be drawn.
- **Rectangular Corners:** Selecting this radio button will draw the graph box as a simple rectangle.
- **Rounded Corners:** Selecting this radio button will draw the graph box as a rectangle with rounded corners.

5.5 Graph Input and Output

5.5.1 Copy to Clipboard

Selecting this menu item copies the current plot to the clipboard so that you can paste it into another application. Best results when copying are obtained by resizing the graphing window to the actual size that you want to paste into the other application, and then selecting “Copy Plot.”

5.5.2 Print

Selecting this menu item prints a hardcopy of the current plot. A print dialog will be shown that will allow you to select printers and change the print setup. If the printer is not set as desired, press the Setup button and configure the printer. Pressing OK will commence printing, and pressing the Cancel button will return you to the CurveExpert graph window.

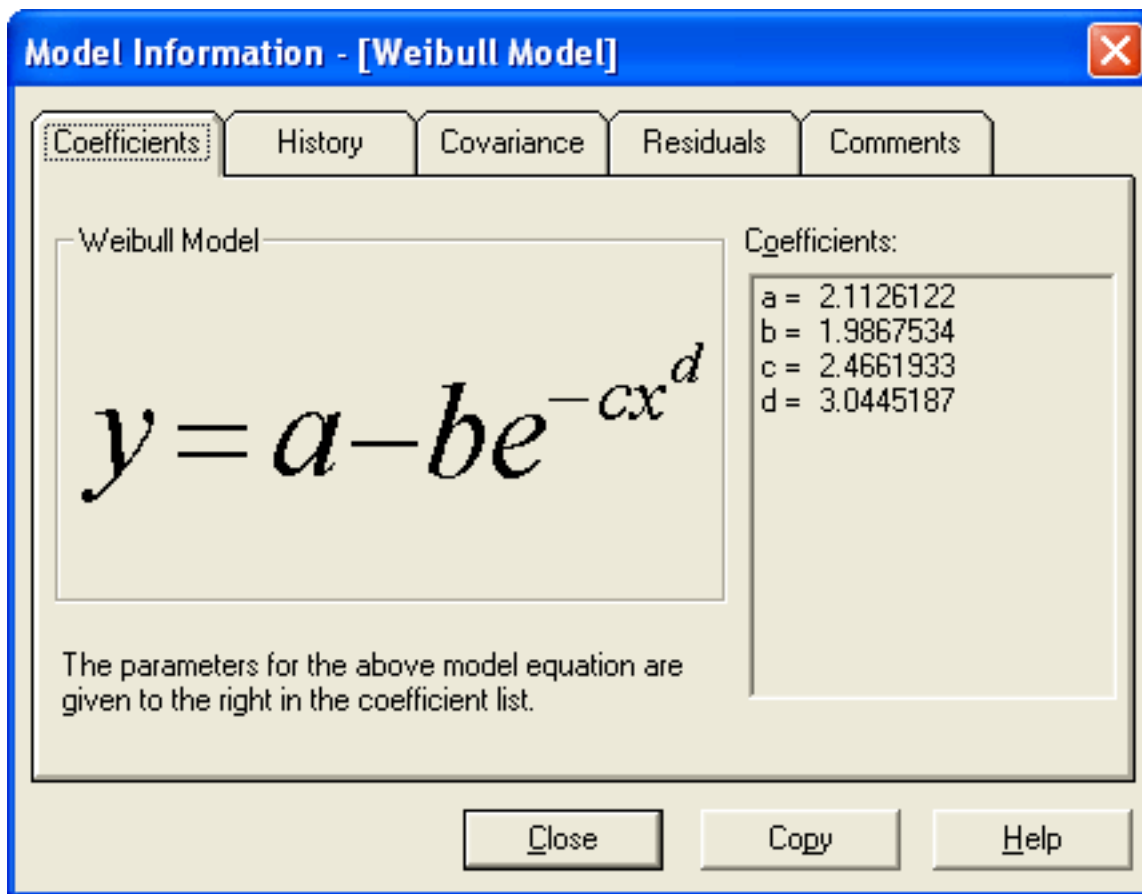
Note that you may select whether the printout will be in color or black and white by selecting the Black and White and Color radio buttons at the bottom of the dialog. If you have a black and white printer and select a color plot, then the plot will be grayscaled before it is sent to the printer.

QUERYING RESULTS

All result query operations are available through the graph. Right click on the graph in order to access the query operations.

6.1 Model Information

The Model Information dialog gives all of the information needed to apply a curve fit and some supplemental information about the performance of the given model. The most important information contained in this panel are the values of the parameters of the model.



6.1.1 Information Buttons

- OK: Press OK to close this window.
- Help: Press Help to get CurveExpert help on how to interpret the items in the this dialog.
- Copy: Press the copy button to copy the displayed information to the clipboard. See the individual sections below for the specific data that will be copied.

6.1.2 Coefficients

The Coefficients page gives the most vital information about the model in the current graph. The values of the coefficients are given here as well as the model equation, so that meaning of the parameters is clear.

Model

The Model section gives the data modeling equation so that the interpretation of the coefficients is clear. The coefficients a,b,c, etc. in the Coefficients section match the corresponding parameters shown in the model.

Coefficients

The Coefficients section gives the values of the relevant parameters for the current model – these coefficients are always expressed as a, b, c, d, and so on. These coefficients match the model shown in the Model section. If the list of coefficients exceeds the extent of the window that they are shown in, a scrollbar will appear to allow you to scroll through the parameters.

Copy

If the Copy button is pressed when the coefficient tab is active, the list of coefficients is copied to the clipboard.

6.1.3 History

The History page shows the value of the chi-square and each parameter in the current curve fit as a function of the iteration number. These values give information on the path that the nonlinear regression algorithm took to optimize the parameters (where optimization means reducing the chi-square value as much as possible). This page is divided into two sections: the Chi-Square History, and the Parameter Histories.

The Chi-Square history section shows the iteration number in the left column, and the value of the chi square at that iteration on the right. At iteration zero, the associated chi-square is computed with the initial guesses. To view the history for any parameter, use the list box beside Param 1 or Param 2. Two lists are provided so that the user can compare changes in two parameters at once.

Note that the History page is only meaningful for nonlinear regression type curve fits. Linear regressions and interpolations can be solved directly, so there is no iteration involved in their solution.

Copy

If the Copy button is pushed when the history tab is active, the complete history for the chi-square and then each parameter is copied to the clipboard in successive columns.

6.1.4 Covariance Matrix

The covariance matrix gives further information about the relation between the model parameters and their uncertainties. The entries along the diagonal are directly related to the uncertainty in each parameter – for example, the entry in matrix position (2,2) corresponds to the uncertainty in the second parameter (which is b). The off-diagonal components are the covariances between parameters. For example, the entry in matrix position (3,4) is the covariance between parameters 3 and 4 (which are c and d). The covariance matrix can be used in further statistical analysis of the regression model and its goodness-of-fit. At this time, CurveExpert offers this matrix as additional information to the user, and does not draw any conclusions based on its entries. Note also that the inverse of the covariance matrix is the curvature matrix.

Copy

If the Copy button is pushed when the covariance tab is active, the entire covariance matrix, in nxn format, is copied to the clipboard.

6.1.5 Residuals

The residual plot graphically depicts the difference between the data points and the model evaluated at the data points. The residual at point i is defined by

$$R_i = y_i - f(x_i)$$

where y_i is the measured value at x_i , and $f(x_i)$ is the predicted value at x_i . These distances are shown as bars or points on the residual plot; the magnitudes of the data points are simply replaced by the residual defined above. If the residual is positive, then the data point is above the model's prediction; likewise, if the residual is negative, then the data point is below the model's prediction. The residuals can provide an indication of a particular model's performance. If there are runs of like-signed residuals, then a better model for the data is likely to exist.

Optimally, the residuals should exhibit a random scatter around zero, which indicates that the data points are randomly distributed around the curve. A run is a sequence of like-signed residuals, which stand out on the residual plot. A large number of runs indicates that data systematically deviates from the curve.

Open/Close Plot

Press this button to open or close a large residual plot. Opening the plot simply shows the residual plot in its own window in which you may resize, zoom, print, and alter the properties just as with any other graph. While the residual plot window is open, this button will change to read Close Plot. Clicking this button will then close the large residual graph.

Copy

If the Copy button is pressed when the residual tab is active, a table of residuals is copied to the clipboard. This table is in a two-column format separated by a tab character. The first column is simply the x data, and the second column is the difference between the data point and the model evaluated at that point ($y_i - f(x_i)$).

If a copy of the plot is desired, simply open a plot window of the residuals (by clicking the Open Plot>> button and then select Copy from the graphing menu of that plot).

6.1.6 Comments

The comments page is meant to give additional information on the error involved in the curve fit, and also gives other miscellaneous information that could apply to your particular model.

Errors

The Error section gives you information on the curve fit performance. Two quantities are used to express the “goodness” of a particular curve fit – the correlation coefficient and the standard error of the estimate. In general, the correlation coefficient will range from 0 to 1, with a correlation coefficient of 1 being the best. In some peculiar circumstances, CurveExpert will report a correlation coefficient greater than one; this is indicative of a very poor data model. The standard error will be strictly positive, with a smaller standard error representing the better curve fit. For more information, refer to the Data Modeling with CurveExpert section.

Comments

The Comments section gives you important information about the current model, such as information regarding the characteristics of interpolation, or the performance of a nonlinear regression. For nonlinear regression, the message will inform you if the fit converged, the number of iterations required for convergence, and the tolerance imposed to determine convergence.

Copy

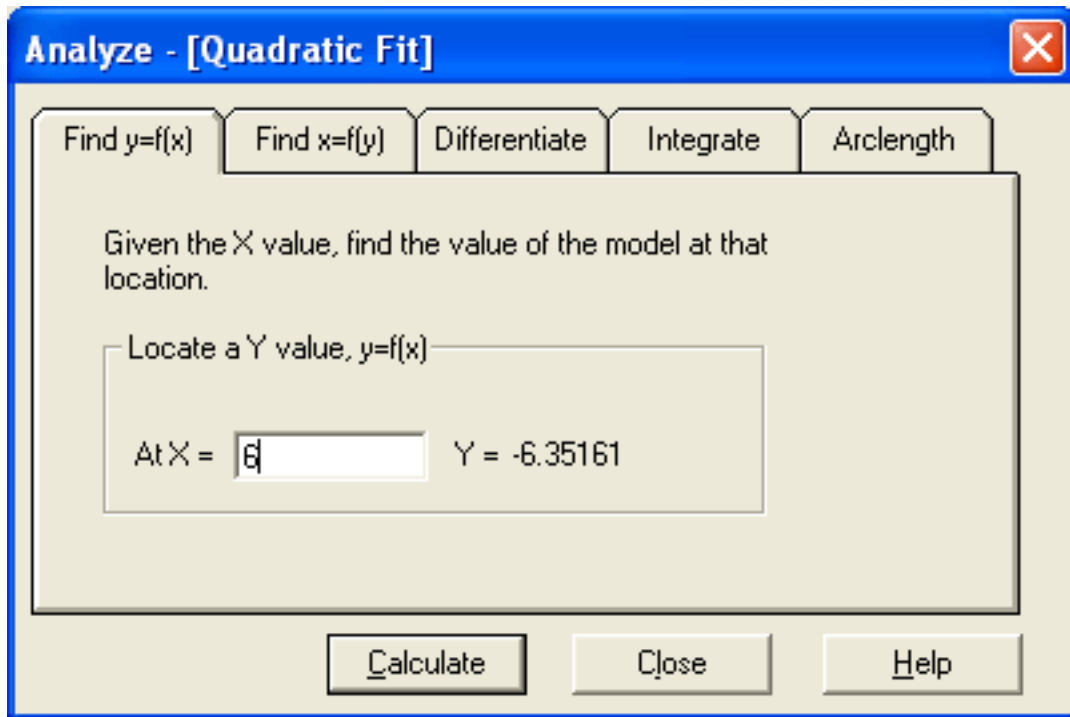
If the Copy button is pushed when the comments page is active, the error information and the comments section is copied to the clipboard.

6.2 Analyze

The Analyze Fit feature allows easy location of x/y points, differentiation, and integration of a curve fit that has been performed. Since this feature may only be used from the graphing window (by choosing Analyze from the graphing menu or pressing Ctrl-L), all calculations will be based on the model from the invoking window.

6.2.1 Analyze Fit Control Buttons

- Calculate: Press calculate to compute the information on the active page. The item calculated could be the location of an x point, location of a y point, a derivative, or an integral (depending on the active page).
- Close: Press Close to indicate that you are through analyzing the current model. The dialog box will be closed.
- Help: Press Help to get CurveExpert help on how to use the Analyze Fit feature to locate discrete points, differentiate, and integrate the model function.



6.2.2 Find a Y Value, $y=f(x)$

CurveExpert can evaluate the curve fit at a given x to give you the projected y value by the current curve fit. Just type in the point at which you want to evaluate the curve fit (the x value) in the “At X=” field and press Apply or <Enter> on your keyboard. The calculator will then show the corresponding y value. If CurveExpert cannot evaluate the curve fit at the point that you specified, then the word ERR is printed in the result field.

6.2.3 Find an X Value $x=f(y)$

CurveExpert can evaluate the curve fit at the given y to give you the projected x value. Just type in the point at which you want the corresponding x in the “At Y=” field and press OK or <Enter> on your keyboard. The calculator will then show the corresponding x value. If CurveExpert cannot find the corresponding x value, then the word ERR is printed in the result field.

NOTE that the results obtained for the x value should be, at the least, visually checked with the graph. There is a chance the root (x -value) returned matches a y value on the graph at a location other than what was intended. This occurs simply because functions may have multiple roots, and CurveExpert may locate any one of them (since all of them are just as valid as the other) without assistance from the user. Consider a simple quadratic $y=x^2$: if you want the x value corresponding to $y=4$, there are two valid answers to this question (2 and -2). So, although this search facility in CurveExpert is very powerful, it should be used with care. The optional Initial Guess for X field should be used wisely so that the correct root is located.

Most of the time, the initial guess field is not needed; however, as mentioned above, the desired x -value may not be returned. In this case, you must set the Initial Guess for X to a value reasonably close to the root (x -value) that you are trying to find. This greatly assists CurveExpert in finding the correct root. Note that the number shown in this field when the Analyze Tool is invoked is the initial guess generated by CurveExpert, which will subsequently be used in the computation of the x -value.

6.2.4 Differentiating at a point

Simply fill in the point at which you want to take a derivative into the “At X=” edit field. Then press the “Calculate” button, and the result will be displayed beside the “dY/dX=” entry. The default value at which the derivative will be computed is the current midpoint of the graphing window. If a derivative calculation fails, the word ERR will appear in place of the result. The derivatives are taken using a central difference combined with Richardson extrapolation to refine the derivative. This refinement is terminated when successive refinements yield less than 0.01% difference in the derivative.

6.2.5 Integrating over an interval

To integrate a function on the interval [a,b], Fill in the value of a in the “From X=” field, and the value of b in the “to X=” field. Then press the “Calculate” button, and the result will be displayed beside the “I(ydx)=” entry. The default interval over which the integral will be taken is the current extent of the invoking graphing window. If an integral calculation fails, the word ERR will appear in place of the integral result.

The integral of the model function is taken using Simpson’s 1/3 rule coupled with the Romberg integration method. The interval is divided into successively finer subintervals which are integrated with Simpson’s rule and summed, and these results are refined with the Romberg method. Refinement is terminated when successive iterations yield less than 0.01% difference in the integral or when the interval has been divided into over 32767 subsegments. If the integral did not converge before the maximum number of segments is exceeded, a note will appear informing you of this occurrence. Use the result of an unconverged integral with due caution.

6.2.6 Finding Arclength

The length of the graph from point a to point b is defined as the arclength. This distance is equivalent to the distance traveled by an observer moving along the curve from the location (a, f(a)) to (b, f(b)).

Exercise caution when finding the arclength of a given plot. The formula to calculate arclength is

$$L = \int_a^b \sqrt{1 + \left[\frac{dy}{dx} \right]^2} dx$$

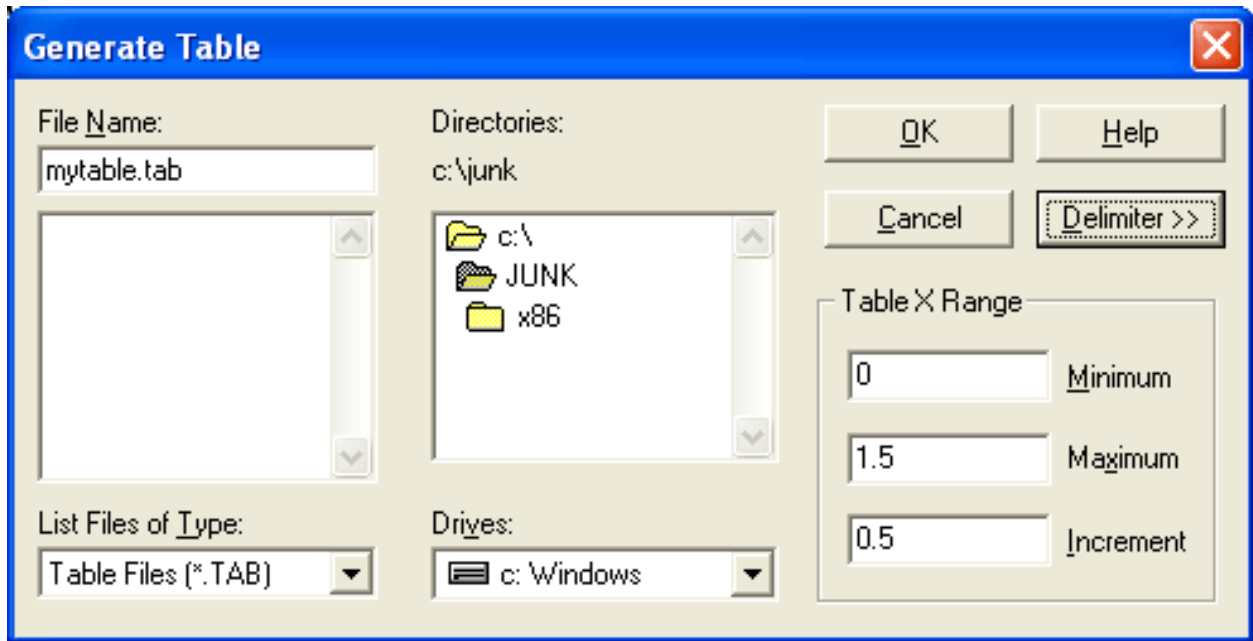
So, if the derivative dy/dx becomes very small (numerically), this calculation can fail. Check your curve using the Differentiation tab to make sure that the derivative over the interval in question is not small (i.e., less than 1e-5).

6.3 Creating a Table

The CurveExpert table generation mechanism generates a data table based on the current data model. This table will contain the calculated model values(y) corresponding to the range of x data points that you specify. To create a table, CurveExpert must know the following information:

1. The starting x value
2. The ending x value
3. The resolution (increment, step size) of the table

Simply fill this information into the provided spaces on this dialog box, enter a filename, and press the ‘OK’ button. By default, the tables generated by CurveExpert have the extension ‘.TAB.’ The generated table will simply be two columns of data: the first column is the x data, and the second column is the calculated y data from the model that you specified.



6.3.1 Table Control Buttons

- **OK:** Press OK after you have entered all of the necessary information and you are ready to save your table.
- **Cancel:** Press Cancel to abort this operation and not create a table.
- **Delimiter:** Press this button to specify the type of delimiter, or separator, to use in the table. You may choose a space, tab, or comma to separate the x and y columns. The default delimiter is a tab.
- **Help:** Press Help to learn on how to generate a table from a given data model.

6.3.2 Table X Range

This region of the dialog box is where you can specify the starting and ending x value and the resolution of the table, as discussed above.

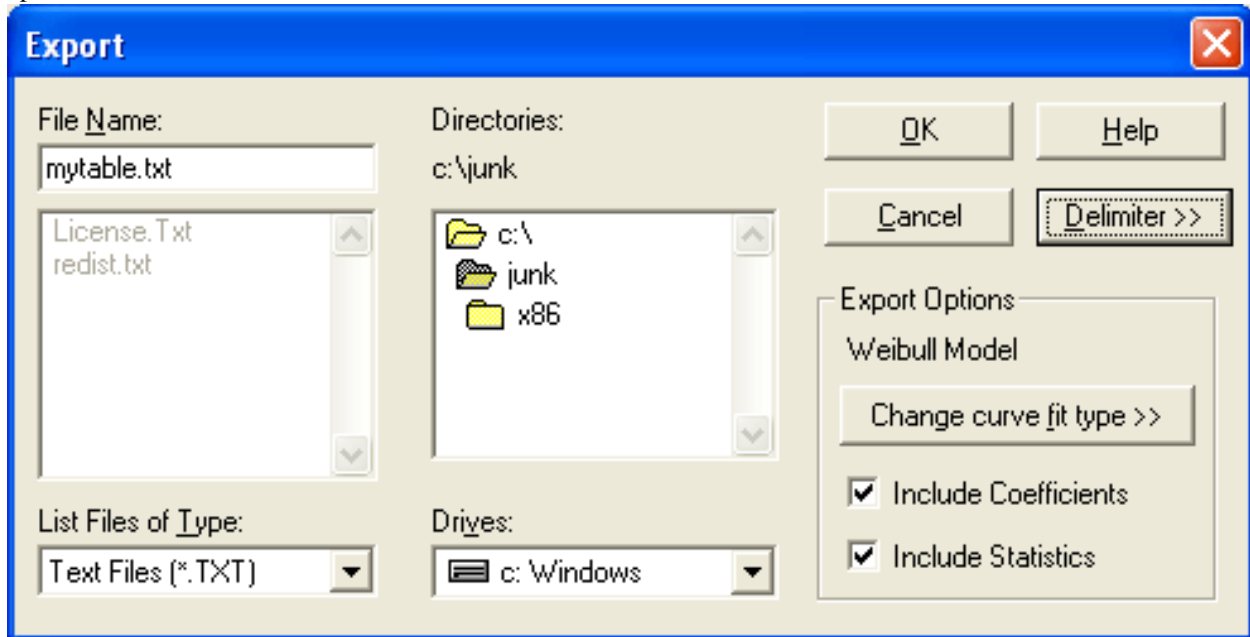
- **Minimum:** Enter the starting x value of the table in this edit field.
- **Maximum:** Enter the ending x value of the table in this edit field.
- **Increment:** Enter the interval in x that you want the data model to be calculated. Note that this number must be the same sign as (maximum-minimum). By specifying a negative Increment and a minimum greater than the maximum, it is possible to generate reverse-order tables.

6.3.3 Example

For example, suppose that you wanted to generate a table file named “mydata.tab” for a range of $x=0.0, 0.5, 1.0, 1.5\dots 10.0$ with comma separators between columns. You would specify 0.0 in the ‘Minimum’ field, 10.0 in the ‘Maximum’ field, and 0.5 in the ‘Increment’ field. Then press the ‘Delimiter’ button and choose ‘Comma’ from the delimiter menu. Finally, you would enter the filename mydata.tab in the File Name field and press the ‘OK’ button.

6.4 File Export

Exporting a file allows you to easily transfer data out of CurveExpert, most often to a spreadsheet. The exported file will contain the data points along with the specified curve fit's approximation to the data evaluated at every data point. Also, if you desire, the exported file can contain the specified model's coefficients (a,b,c,d...) and statistics describing your data set such as the mean and standard deviation. The filename for the exported file is entered in the File Name field as usual. The extra buttons on the right hand side of the Export dialog are special settings for the exporting operation.



6.4.1 Delimiters

The Delimiter button allows you to choose what type of delimiter you would like to use to separate the data columns. The default is a tab delimiter, but you may also use a single space or a comma.

6.4.2 Specifying Curve fit Type

The curve fit that you will export is, by default, set to the last curve fit that was performed. Its name is shown in the Export Options group. If you wish to change this curve fit, click the Change Curvefit Type>> button – a list of curve fits that have already been calculated will appear, and you may select from that list. Note that a curve fit must already be calculated to be valid for exporting. If the curve fit you desire is not on the list, cancel the Export dialog, calculate the curve fit via the Apply Fit or Interpolation menus, and then reopen the Export dialog.

6.4.3 Include Coefficients

The checkbox Include Coefficients signifies whether you would like the curve fit coefficients to be present in the exported file. Note that for spline interpolations, the coefficients are given as the coefficients of EACH piecewise polynomial used between each knot/data point. The default is ON.

6.4.4 Include Statistics

Lastly, the checkbox Include Statistics signifies whether you would like relevant statistics to be present in the exported file. The statistics that will be included are: X and Y data minimum, X and Y data maximum, X and Y data mean, and the X and Y data standard deviation. The default is ON.

DEFINING CUSTOM MODELS

7.1 Overview

CurveExpert allows you to define your own nonlinear regression models; all you have to do is type in the equation. When you choose Define User Model from the Apply Fit menu, you will be presented with a dialog in which you can type or choose your model. Type your model into the edit field provided, or choose a model from the scrolling list. You may use operators and functions in your model that are defined at the end of this help page. Parameters should start at letter 'a' and continue sequentially. CurveExpert allows you to use up to 19 parameters, lettered a..s. All arithmetic operations are valid, as well as the symbol ^ to raise a number to a power. Here is an example of what you could enter in the Define User Model dialog box:

```
a*x*log(b+x)+c^x
```

Note that there is no need to enter the 'y=' before the function – only the right hand side of the model is entered.

Logical operators are also supported in your user-defined regression models. The operators that are less than (<), greater than (>), less than or equal to (<=), or greater than or equal to (>=). Each of these expressions evaluate to one if the expression is TRUE, or zero if the operation is FALSE. In this way, you can implement logical expression into your model. For example, if you would like to construct a model of two separate lines to where one line is applicable to the region of $x \leq 5$, and one line is applicable to $x > 5$, then you might construct the following four-parameter model:

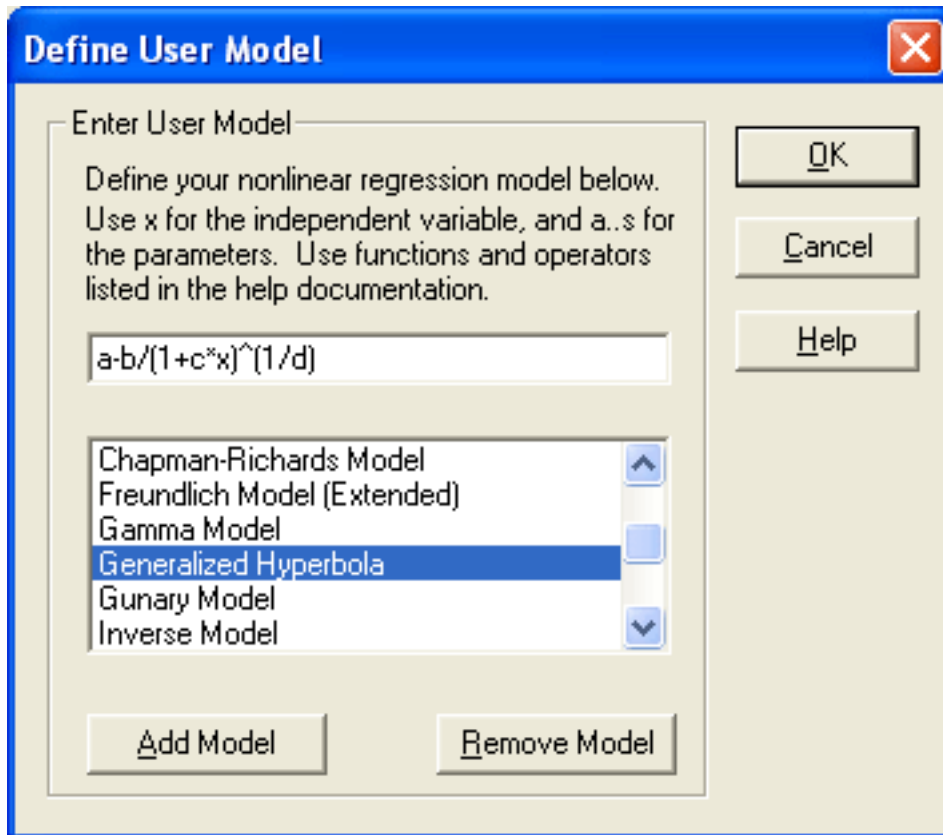
```
(a+b*x)*(x<=5)+(c+d*x)*(x>5)
```

These logical operators can be combined in any way to create your user defined model.

See Custom Model Reference for a complete list of the supported arithmetic and logical operators.

7.2 User Model Dialog

A picture of the User Model dialog is shown below, with a description of the controls/operations following:



7.2.1 User Model Control Buttons

- **OK:** Press OK to use the model that you have entered to fit the data set. The nonlinear regression calculation will begin. If there is a syntax error with your model, you will be notified at this time.
- **Cancel:** Press Cancel to abort this operation and return to the main window.
- **Help:** Press Help to learn how to apply custom regression models.

7.2.2 Model Edit Field

Type your custom regression model into this field, using a...s for the model parameters. Parameters should start at letter 'a' and continue sequentially. CurveExpert allows you to use up to 19 parameters, lettered a..s. All arithmetic operations are valid, as well as the symbol ^ to raise a number to a power. Note that there is no need to enter the 'y=' before the function – only the right hand side of the model is entered. Also, you may edit the text in this box at any time, even after you have selected a model from the Model List.

The normal cut, copy, and paste keys can be used in this edit field; i.e., Ctrl+X for cut, Ctrl+C for copy, and Ctrl+V for paste.

7.2.3 Model List

User models that have been saved will be listed in this box. Scroll up or down to see them all. Selecting an item from this box will simply copy its formula to the Model Edit Field so that you can either click OK to calculate that model, or modify it as necessary. Double-clicking a model in this list will select that model for a regression immediately.

7.2.4 Add Model

Clicking on this button will add the model that is in the Model Edit Field to the Model List. This operation permanently saves a model that you have typed in. You will be asked what to call the model so that a name (not a formula) will be added to the list.

7.2.5 Remove Model

Clicking on this button will remove the highlighted model from the Model List permanently. A dialog will ask you to verify this operation.

7.3 Custom Model Reference

7.3.1 Operators

The following operators are supported by CurveExpert to be used in your custom regression model: * + addition * - subtraction * - unary minus (negation) * * multiplication * / division * ^ power * < less than * > greater than * <= less than or equal to * >= greater than or equal to * != not equal * == equal

7.3.2 Special Variables

The following special variables are supported by CurveExpert may be used in your custom regression model: * pi the constant 3.14159..... * avgx average of the x column * avgy average of the y column * stdx standard deviation of the x column * stdy standard deviation of the y column * minx minimum of the x column * miny minimum of the y column * maxx maximum of the x column * maxy maximum of the y column

7.3.3 Functions

The following functions are supported by CurveExpert to be used in your custom regression model:

- abs absolute value
- fact factorial
- sqrt square root
- exp exponential
- log natural logarithm
- log10 base-10 logarithm
- sin sine
- cos cosine
- tan tangent
- asin arcsine
- acos arccosine atan
- arctangent sinh
- hyperbolic sine

- cosh hyperbolic cosine
- tanh hyperbolic tangent
- ceil first integer greater than value
- floor first integer less than value
- j0 Bessel function of the first kind, 0th order
- j1 Bessel function of the first kind, 1st order
- y0 Bessel function of the second kind, 0th order
- y1 Bessel function of the second kind, 1st order

COMPUTATIONAL PERFORMANCE

8.1 Increasing Performance

CurveExpert has been optimized to perform regressions and interpolations as quickly as possible. Additional speed gains are sure to come in later versions; however, you can greatly speed up the operation and reliability of CurveExpert by carefully following certain guidelines:

8.1.1 Disable graph preview animations

Right click on the graph preview window (the graphing area in the main window of CurveExpert), and make sure that “Enable Animations” is off. By default, CurveExpert ships with this option enabled in order to provide live feedback on changes being made to the curve fit as it is being calculated.

8.1.2 Scale your datasets

Always, always, always scale your data sets to the order of one. You may do this prior to reading in your data file, or after your data file has been imported using the scaling feature in CurveExpert. This point cannot be overemphasized! If you contact me with a problem concerning a situation where CurveExpert will not fit a certain set of data, this is the first question that I will ask.

Imagine a data set with x values ranging from 1000 to 10000 and a regression model where the term $\exp(a*x)$ is involved. Unless a very good initial guess for a is given, chances are that an exponential to a very large power (eg. $\exp(1000)$) will be taken in the course of the nonlinear regression algorithm. The calculation will overflow, and the regression will fail as a result. Even if the regression happens not to fail, the regression algorithm will have an exceedingly tough time finding the correct parameters, since a small change in the free parameters will cause a tremendous change in the size of the term. The moral of the story: always scale your data! Note that CurveExpert is able to perform correctly on data in any scale, as long as the calculations do not overflow or underflow. So, if a data set is giving problems, scaling it should be the first action to take.

8.1.3 Enable background processing (obsolete)

Enabling background processing (in the FilePreferences dialog) slows down computation considerably, since the calculation must pause for Windows to process pending messages for all running programs. If you are running CurveExpert by itself, or you don't care if another program is ignored when CurveExpert is calculating, then turn this option off! Calculations speed will increase greatly. The CALC flag in the status bar will be written in red when background processing is disabled to remind you that every other program is halted while CurveExpert works.

8.1.4 Set the tolerance parameter reasonably

Don't set the tolerance parameter (in the FilePreferences dialog) too low. In regression modeling, not much advantage is to be gained by setting a very strict tolerance. Its main purpose in life is to prevent the nonlinear regression algorithm from converging on local minima.

8.1.5 Use single precision calculations (registered users only)

CurveExpert uses double-precision floating point numbers to do all calculations. The registered version of CurveExpert comes with two executables: one version using single-precision calculations (seven sig figs), and the second using double-precision (15 sig figs). Using a single-precision (theoretically) greatly speeds up performance, at the cost of accuracy. Especially with extremely large data sets, the single-precision version should provide a speed and memory advantage. By registering, you can make the choice yourself. See the sample speed test data below for an indication of the performance of the different packages. You may check what type of data representation is being used by CurveExpert by selecting Help>About.

8.2 Floating Point Exceptions

CurveExpert takes every precaution to avoid math errors, a.k.a. floating point exceptions. Here are some ways that you can avoid these errors:

8.2.1 Scaling of your data is important

Take care to scale both x and y data to a magnitude of one. For example, if you have data that describes atmospheric pressure at different elevations, you might have (in metric units) a data set that looks like:

```
x=[-100, 0, 100, 500, 1000, 4000] meters  
y=[102000,101325,101000,100500,100000,99900] Pascals
```

Using the scaling feature in the Data menu of CurveExpert, you can scale this data using a scale factor of 0.001 on the x data, and 0.00001 on the y data. So, you would then have the following data set:

```
x=[-0.1, 0, 0.1, 0.500, 1, 4] kilometers  
y=[1.02,1.01325,1.01,1.005,1.0,0.999] bars
```

The second example is much more likely to allow nonlinear regressions to converge, and also will allow higher order polynomial fits to be performed with more accuracy. If you have a data set that seems particularly ill-behaved, scaling can help solve this problem.

8.2.2 Data should be appropriate to the model

Make sure that the data is appropriate to the model. Especially look out for using logarithmic or exponential families of models with data that contains zeros or negatives. For example, it is not possible, in any shape or form, to obtain a negative or zero with the basic exponential model ($y=ae^{(bx)}$, assuming a is nonzero).

So, it is not wise to use a model that cannot reflect the trends in the data. CurveExpert should disallow the use of certain regression models with data that is negative or zero; however, in the case that one should slip through or you are using a user-defined model, you should be aware of the characteristics of the data modeling equations so that choices of an incorrect model can be avoided.

8.2.3 For interpolations, data points must be sorted

If you are using an interpolation-type curve fit, the data points **MUST BE SORTED** by ascending x ! Use the DataManipulate feature in CurveExpert to sort your data points correctly. It is quick and painless!

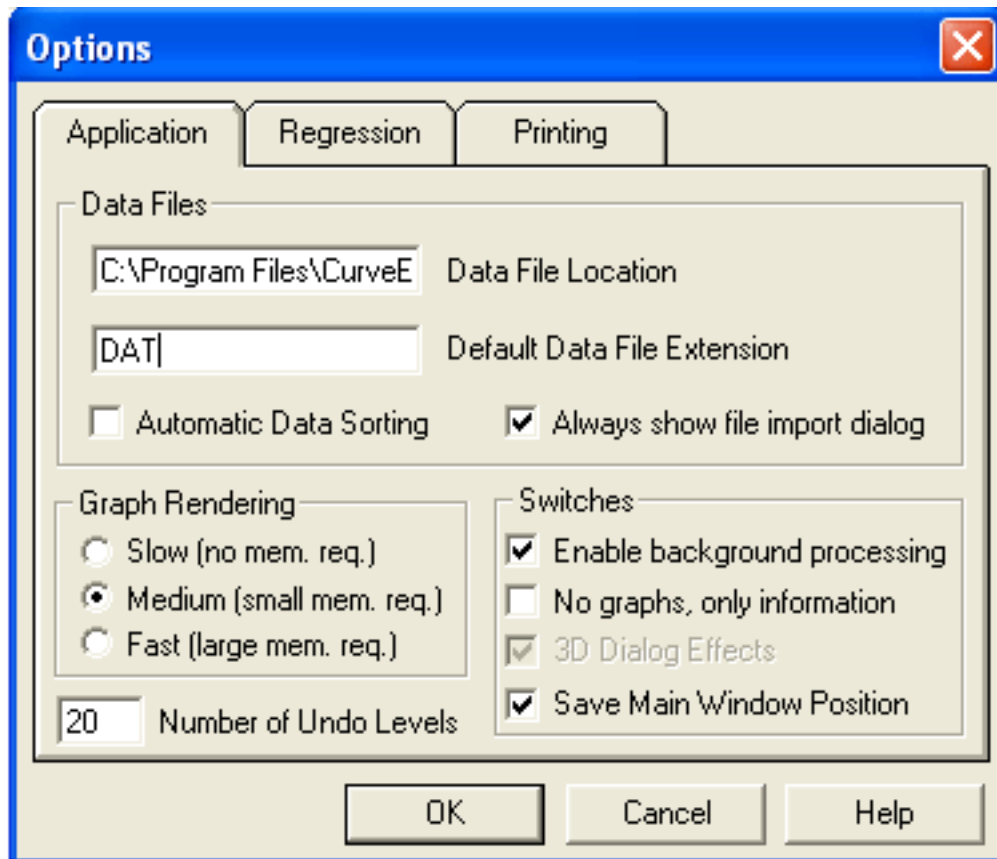
8.2.4 Sometimes, it's unavoidable

Some curve fits are simply ill-behaved, i.e. prone to divergence. For some data sets, it may not be possible to converge certain nonlinear regression models.

CURVEEXPERT OPTIONS

The Options dialog allows you to set important parameters that CurveExpert will use during operation. Two separate sets of user preferences are maintained for the single and double precision versions (registered users only).

If an invalid entry has been entered in any of the tabbed pages, then the offending entry will be selected, and the computers speaker will sound.



9.1 Application Options

The application options tabbed page sets user preferences that control the overall behavior of CurveExpert in the Windows environment.

9.1.1 Data Files

- **Data File Location:** The directory in which you store your data files can be input in this edit field. For convenience, each time CurveExpert is started, the directory will be automatically changed to this working directory. By default, this directory is set to be the CurveExpert home directory.
- **Default Data File Extension:** If a file extension is not supplied in the FileOpen or FileSave dialog box, then this extension is automatically added. Default: DAT
- **Automatic Data Sorting:** Selecting this checkbox will automatically sort any data file as it is read into CurveExpert. Default: Off
- **Always show file import dialog:** Selecting this checkbox will cause the file import dialog to always appear when a file is selected to read into CurveExpert. This dialog will give access to all of CurveExpert features for parsing data files. When this checkbox is cleared, CurveExpert will attempt to intelligently read the data file based on the number of columns that it detects.

9.1.2 Graph Rendering

The graph rendering radio buttons may be set to Slow, Medium, or Fast. Each setting has advantages and disadvantages.

- The “Slow” setting requires no extra memory, but all graphs are drawn from scratch when their window is resized or uncovered.
- The “Medium” setting requires a small amount of extra memory and behaves like the “Slow” setting, with the exception that any user-model graphs will be drawn using “Fast” rendering. See below; this is the default.
- The “Fast” setting requires a large amount of extra memory, because the contents of every graph window are cached for quick redraws. The graph is only drawn the first time it is opened, and each time it is resized. Any other graph viewing will take a minimal amount of time to redraw, since the image is stored in memory. This option is most useful for users who have extremely large data sets that take a long time to draw.

9.1.3 Switches

- **Enable Background Processing:** When CurveExpert is iterating, it can take complete control over Windows (faster), or do its processing in the background (slower). Checking this checkbox will force CurveExpert to do its calculations in the background. This is recommended, since certain windows processes need attention often. On the other hand, if CurveExpert is the only application running, then the best performance will be obtained by leaving this box unchecked. Default: Enabled
- **No graphs, only information:** CurveExpert gives you the option of viewing the graph first, or the fit information first. If this checkbox is selected, the model information (coefficients, etc.) will be shown directly, instead of a graph. Default: Disabled
- **3D Dialog Effects:** Attractive (in the opinion of the author) gray dialogs with 3D effects are the default in CurveExpert. If you prefer the white dialogs with no 3D effects, deselect this checkbox. Note that this options is largely obsolete; as all versions of Windows later than Windows 95 have 3D dialog effects by default. Default: Enabled.
- **Save Main Window Position:** CurveExpert saves the position of the main window when the application is exited, so that when it is opened again, it assumes the same position. If this behavior is undesirable, deselect this checkbox. Default: Enabled.

9.1.4 Setting the number of Undo Levels

- **Number of Undo Levels:** CurveExpert implements a multilevel undo system for the built-in spreadsheet. The number of actions that can be undone can be set in this edit field. The number of levels of undo must range from 1 to 99. If your system has low amounts of memory, then a low number of undo levels will be conserve this memory. Default: 20.

9.2 Regression Options

The regression options tabbed page sets user preferences that control the regression and mathematical operation of CurveExpert.

9.2.1 Linear Regression

- **Maximum Polynomial Degree:** Enter the maximum degree of a polynomial allowed by CurveExpert in this field. This setting determines the number of polynomials examined by CurveFinder in a search sweep, and it also prevents choosing a wildly oscillating polynomial by accident. This number may not exceed 19. Default: 10
- **Default Polynomial Degree:** Enter the default degree of polynomials used in CurveExpert in this field. The default polynomial degree is the polynomial that will be calculated by the Calculate Groups command, and the degree that will appear by default when you select a polynomial from the Apply Fit menu. Default: 4

9.2.2 Nonlinear Regression

- **Regression Tolerance:** Enter the desired tolerance for nonlinear regression convergence in this field. The nonlinear algorithm uses this number to determine whether or not it has converged on the correct model parameters. Mathematically speaking, the difference between the standard errors in two consecutive iterations must be smaller than this tolerance for the computation to terminate. The smaller the tolerance, the more accurate the parameters will be, but the algorithm will take much more time to converge. Default: 1e-6
- **Diverging Iterations:** Enter the number of iterations that the standard error (mentioned above) must increase for the algorithm to declare divergence. Default: 10
- **Maximum Iterations:** Enter the maximum number of iterations for the nonlinear algorithm to perform before it is forced to terminate. Default: 100

9.2.3 General Regression

- **Weighting:** CurveExpert can either perform no weighting for each point, or weight each point by $1/y^2$ as the nonlinear regression progresses. Using no weighting is most appropriate when no known relationship exists between the variances of the data around the fitted curve and the y values of the points. However, it is often the case in experiments that it is known that the uncertainty is a certain percentage of the y value; in this case, weighting the points by $1/y^2$ can prevent the larger scatter at higher y from adversely affecting the regression model. Default: No Weighting
- **Automatic Initial Guesses:** Enabling this checkbox will let CurveExpert generate initial guesses for built-in nonlinear regressions for you. This option, in effect, enables AutoGuess upon startup. Default: Enabled
- **Quick Derivatives:** The partial derivatives of the model with respect with each of its parameters is required in the Levenberg-Marquardt algorithm for nonlinear regression. CurveExpert uses central differencing to approximate

these partial derivatives; however, more precise approximations can be accomplished using Richardson's extrapolation. Checking the "Quick Derivatives" checkbox uses simple central differencing; leaving this checkbox unselected forces CurveExpert to use Richardson's extrapolation for the numerical derivatives. The latter is more accurate, but much slower than using central differences. Default: Enabled

9.3 Printing Options

The printing preferences tabbed page sets user preferences that control printing settings and margins that CurveExpert will use for graph hardcopies and data tables.

- **Current Printer Info:** This section shows all information for the currently selected printer. The name, driver version, orientation, paper size, and resolution are shown.
- **Default Print Type:** The default print type determines whether the "Black and White" or "Color" printing selection will be automatically selected when a graph is printed. Of course, the user may change the selection at print-time. This option is included for the convenience of color printer users. Default: Black and White.
- **Print Margins:** The top, left, bottom, and right margins (in inches) can be set in each of these edit fields. Note that no margin can be less than 0 inches, and no corresponding margins can add to higher than the paper extent plus the unprintable margin.
- **Print Setup:** Press the Print Setup button to change the current printer and/or the printer configuration.

APPENDIX: BUILT-IN NONLINEAR REGRESSION MODELS

Nonlinear regression models come in all shapes and sizes – in CurveExpert, an attempt has been made to include as many models as possible, while at the same time providing relevant models to a wide range of applications. Also, keep in mind that if the model that you desire is not built-in to CurveExpert, you may define your own using the Define User Models facility. In CurveExpert, the nonlinear models has been divided into families based on their characteristic behavior. These families and their members are enumerated below:

10.1 Exponential Family

Exponential models have the exponential or logarithmic functions involved. They are generally convex or concave curves, but some models in this group are able to have an inflection point and a maximum or minimum.

- Exponential $y = a * \exp(b * x)$
- Modified Exponential $y = a * \exp(b/x)$
- Logarithm $y = a + b * \ln(x)$
- Reciprocal Logarithm $y = 1 / (a + b * \ln(x))$
- Vapor Pressure Model $y = \exp(a + b/x + c * \ln(x))$

10.2 Power Family

The Power Family involves raising one or more parameters to the power of the independent variable, or raising the dependent variable to the power of a given parameter. This family is generally a set of convex or concave curves with no inflection points or maxima/minima.

- Power Fit $y = a * x^b$
- Modified Power $y = a * b^x$
- Shifted Power $y = a * (x - b)^c$
- Geometric $y = a * x^{(b * x)}$
- Modified Geometric $y = a * x^{(b/x)}$
- Root Fit $y = a^{(1/x)}$
- Hoerl Model $y = a * (b^x) * (x^c)$

- Modified Hoerl Model $y = a \cdot b^{(1/x)} \cdot (x^c)$

10.3 Yield-Density Models

The yield-density models are widely used, especially in agricultural applications. These models historically have been used to model the relationship between the yield of a crop and the spacing or density or planting. Essentially two types of response are observed in practice: the “asymptotic” and “parabolic” yield-density relations. If the response is such that as density (x) increases, but the yield (y) approaches a fixed value, the relationship is asymptotic. If the response is such that there is a distinct optimum as the density increases, the relationship is parabolic. Of course, these types of relationships occur commonly in other scientific areas; therefore, this family of models is very useful.

- Reciprocal Model $y = 1 / (a + bx)$
- Reciprocal Quadratic $y = 1 / (a + bx + cx^2)$
- Bleasdale Model $y = (a + bx)^{-1/c}$
- Harris Model $y = 1 / (a + bx^c)$

10.4 Growth Family

Growth models are characterized by a monotonic growth from some fixed value to an asymptote. These models are most common the engineering sciences.

- Exponential Association (2) $y = a \cdot (1 - \exp(-bx))$
- Exponential Association (3) $y = a \cdot (b - \exp(-cx))$
- Saturation Growth $y = ax / (b + x)$

10.5 Sigmoidal Family

Processes producing sigmoidal or “S-shaped” growth curves are common in a wide variety of applications such as biology, engineering, agriculture, and economics. These curves start at a fixed point and increase their growth rate monotonically to reach an inflection point. After this, the growth rate approaches a final value asymptotically. This family is actually a subset of the Growth Family, but are separated in CurveExpert because of their distinctive behavior.

- Gompertz Model $y = a \cdot \exp(-\exp(b - cx))$
- Logistic Model $y = a / (1 + \exp(b - cx))$
- Richards Model $y = a / (1 + \exp(b - cx))^{1/d}$
- MMF Model $y = (ab + cx^d) / (b + x^d)$
- Weibull Model $y = a - b \cdot \exp(-cx^d)$

10.6 Miscellaneous Family

As with many things in life, some things just don't fit into nice categories. The miscellaneous family is the one in which these “different” nonlinear regression models live.

- Sinusoidal Fit $y = a + b \cdot \cos(c \cdot x + d)$

- Gaussian Model $y = a \cdot \exp(-((x - b)^2)/(2 \cdot c^2))$
- Hyperbolic Fit $y = a + b/x$
- Heat-Capacity Model $y = a + bx + c/x^2$
- Rational Function $y = (a + bx) / (1 + cx + dx^2)$

APPENDIX: LEGAL INFORMATION

11.1 Copyright

CurveExpert 1.4 is composed of the following files: 16-bit version:

- CURXPT.EXE
- CURXPTS.EXE (registered only)
- CURXPT.HLP
- CURVLIB.DLL
- CURVRES.DLL
- DIAGEXT.DLL
- HKLIB.DLL
- HYAPI.DLL
- INSTLIB.DLL
- CURVES.INI
- README.TXT
- VENDOR.TXT
- HISTORY.TXT
- FILE_ID.DIZ

32-bit version:

- CVXPT32.EXE
- CVXPTS32.EXE (registered only)
- CURXPT4.HLP
- CURXPT5.CHM
- CURXPT4.CNT
- CURVRES3.DLL
- DIAGEXT3.DLL
- INSTLIB3.DLL
- HKLIB3.DLL

- HYAPI3.DLL
- GLOBAL.INI
- CURVES.INI
- README.TXT
- VENDOR.TXT
- FILE_ID.DIZ

CurveExpert Copyright 1995-2010 Daniel Hyams. All Rights Reserved.

The DIAGEXT.DLL tab-control extensions were modified from original source by Freeman-Teresa Software and Mr. Edward McCreary:

TabCtrl v1.0 Copyright 1995 Freeman-Teresa Software TabDlg3.DLL is Copyright 1994 Edward McCreary

Microsoft and Windows are registered trademarks of Microsoft Corporation.

Any other proper names are trademarks of their respective companies.

11.2 Software End User License Agreement

IMPORTANT READ CAREFULLY: This End-User License Agreement (EULA) is a legal agreement between you (either an individual or single entity) and Mr. Daniel Hyams of 1698 Chadwick Court, Hixson, TN 37343 (Mr. Hyams) for the software product CurveExpert 1.4, which will be designated by the term Software Product. The Software Product consists computer software and online electronic documentation. By installing, copying, or otherwise using the Software Product, you agree to be bound by the terms of this EULA. If you do not agree to the terms of this EULA, do not install or use the Software Product.

1. **GRANT OF LICENSE.** Mr. Hyams grants to you as an individual, a personal, nonexclusive license to make and use copies of the Software Product. Copies must be made only for backup or archival purposes.
2. **SHAREWARE VERSION.** You are hereby granted a non-exclusive license to use the Software Product for an evaluation period of 30 days, after which you must pay the registration fee to continue using the Software Product. However, use of the Software Product during the 30-day evaluation period is free of charge. Unregistered use of the Software Product after the 30-day evaluation period is in violation of Federal copyright laws and the terms of this EULA. Violators will be prosecuted to the maximum extent under the law. You may make exact copies and distribute the shareware version of the Software Product and its documentation so long as there is no charge applied to distribution of these unmodified copies of the Software Product. You are specifically prohibited from charging or requesting donations for ANY and ALL copies of this Software Product or its documentation. You are prohibited from distributing the Software Product or any portion of the Software Product or its documentation as part of another product (commercial or otherwise) without PRIOR expressed written consent from Mr. Hyams.
3. **REGISTERED VERSION.** If you have registered the Software Product with Mr. Hyams (meaning to remit payment according to the order form contained in the online documentation), the 30-day evaluation period in section 2 is waived. Mr. Hyams agrees to provide you with a registered version of the Software Product, copied onto media deemed suitable, within 30 days of receipt of payment.
4. **REVERSE ENGINEERING.** You may not reverse engineer, decompile, or disassemble the Software Product.
5. **UPGRADES.** All subsequent versions of the software product fall under the terms of this EULA. Upgrades are designated as software packages of the same name but with an incremented version number.
6. **COPYRIGHT.** All title and copyrights in and to the Software Product are owned by Mr. Hyams. The Software Product is protected by copyright laws and international treaty provisions.

7. **WARRANTY DISCLAIMER.** The Software Product and the accompanying files are sold as is without warranty of any kind. All warranties, either express or implied, including, but not limited to, implied warranties of merchantability and fitness for a particular purpose, are expressly and specifically disclaimed.

8. **LIABILITY.** Neither Mr. Hyams or anyone else who has been involved in the creation, production, or delivery of this Software Product shall be liable for any direct, indirect, consequential, special, incidental, or similar damages arising out of the use or inability to use such software even if Mr. Hyams has been advised of the possibility of such damages or claims. In no event shall Mr. Hyams liability for any damages ever exceed the price paid for the license to the Software Product, regardless of the form of the claim. The individual using the Software Product bears all risk as to the quality and performance of the Software Product.

9. **GOVERNING LAW.** This agreement shall be governed by the laws of the State of Tennessee and shall inure to the benefit of Mr. Hyams, his heirs, successors, and assigns. Any action or proceeding brought by either party against the other arising out of or relating to this agreement shall be brought only in a state or federal court of competent jurisdiction located in Chattanooga, Tennessee. The parties hereby consent to in personam jurisdiction of said courts.

ABOUT THE AUTHOR

Hello, my name is Daniel Hyams. I intend to give you the best support possible for this software, so rest assured that I will do my best to take care of any problem that you encounter, or answer any questions that you might have.

I would like to thank several people. First, Charles Petzold for writing such a fantastic book that was worth more than gold throughout this project. Next, the beta testers that provided excellent feedback on the performance of this software, potential uses, and potential problems. A very special thank-you to my wife Michelle, who patiently put up with my endless work hours on this project throughout its multi-year development. Last and most of all, I would like to thank the Lord; without His blessing this project would have never materialized.

The Lord is my shepherd, I shall not want. He makes me lie down in green pastures; He leads me beside quiet waters. He restores my soul. He guides me in the paths of righteousness for His names sake. Even though I walk through the valley of the shadow of death, I fear no evil, for Thou art with me. Thy rod and Thy staff, they comfort me. Thou dost prepare a table before me in the presence of my enemies. Thou hast anointed my head with oil; my cup overflows. Surely goodness and lovingkindness will follow me all the days of my life, and I will dwell in the house of the Lord forever.

Psalm 23, NAS

CONTACTING THE AUTHOR

If you have any questions at all about the use of CurveExpert, or need technical support, feel free to contact the author at the contact points listed below. Also, suggestions for the improvement of this program are always welcome! Also, be sure to read over the Benefits of Registering. Care has been taken to test CurveExpert 1.4 thoroughly, but as we know, bugs may still slip through this process. If you have a problem using this program, experience a bug in the program, or have any suggestions for improvement, please contact me at the address below. For a bug report, the resolution of the problem will be facilitated greatly by sending me a debug report from CurveExpert.

Daniel Hyams attn : CurveExpert 1.4 1698 Chadwick Court Hixson, TN 37343 (423) 843-1861

<http://www.curveexpert.net>

curveexpert@gmail.com

FREQUENTLY ASKED QUESTIONS

14.1 Can CurveExpert work with more than one independent variable?

Unfortunately, no; at this time, CurveExpert only supports one independent variable. For multivariate support, check out CurveExpert Professional.

14.2 How can I cite CurveExpert in a publication?

Hyams, D. G., CurveExpert software, <http://www.curveexpert.net>, 2010.

14.3 Why does CurveExpert suddenly erase my curve fits off of the ranking chart?

Any time that CurveExpert detects that the data set has changed, all curve fits are marked as invalid. The ranking chart shows only those fits that are valid, so the ranking chart is erased. The data set can change by manual editing or using the built-in data manipulation tools in the Data menu.

14.4 How can I increase the speed at which the graphs are updated?

Particularly if one has a large number of data points, the graphing windows are sometimes updated very slowly, as the data point marker must be drawn for every point. However, CurveExpert has a setting in its application defaults, called the Render Level, that can be used to ease the pain of slow redraws. The tradeoff is that a large chunk of precious GDI memory will be used for each graph is open. Choose File|Preferences, and select “Fast Rendering.” Note that when drawing the plot for the first time, or when resizing the window, it will take just as long for the graph to be drawn. The advantage comes about because once this drawing is done, it is saved. Now every time the window is covered/uncovered by another window, or moved to a different location, it does not have to be redrawn from scratch.

14.5 Why does the chi-square reported in the Information dialog not match the standard error?

The standard error and the chi-square values are very similar, but defined differently. The chi square is defined as $X^2 = \sum((y-y_i)^2/\sigma_i^2)$, while the standard error is defined as $S = \sqrt{\sum((y-y_i)^2/\sigma_i^2)/(N-NP)}$. So, if you do the math, you get $S = \sqrt{X^2/(N-NP)}$.

14.6 Why is there a huge number for standard error and correlation coefficient sometimes? Is this a bug?

No, this is not a bug. This is CurveExpert's way of telling you that your data set is badly out of scale – in the course of computing the standard error of the estimate and correlation coefficient, an accumulation of differences between the curve fit and data points and the mean and the data points takes place. If this calculation overflows, which is quite likely when the magnitudes of the x or y data sets are very large, then CurveExpert does the best it can by setting this sum to the maximum allowed by the data representation (single precision or double precision). This usually leads to unrealistic values for the standard error and correlation coefficient. The solution to this problem is to rescale your data to a magnitude of one.

14.7 As nonlinear regression progresses, the indicator for standard error (S) sometimes freezes, although the iteration counter keeps on ticking. What is going on?

Without getting into the internals of the Marquardt-Levenberg method for nonlinear regression (see the Theory section), suffice it to say that everything is still functioning correctly. If the algorithm detects that the standard error (chi square) increases during an iteration, then the new parameters are discarded. This naturally forces the chi square for the current iteration will be the same as the last. However, all is not lost – if this situation occurs, the algorithm makes adjustments to some internal parameters and performs another iteration. It is searching for an 'optimum' value of these internal parameters so that it can reduce the chi square and therefore find better parameters. If the chi square only adjusts itself every other iteration, then the internal parameters are 'on the edge' of optimum and the calculated regression parameters are being discarded every other time.

14.8 A spline fit never matches my data at the points like it is supposed to, and sometimes has spikes and/or wildly oscillating tendencies. Is there something amiss?

Two items will definitely cause this behavior – [1] fitting splines to unsorted data or [2] fitting splines to data with replicate values, i.e., two or more x data points are the same. The solution to problem #1 is easy; simply sort the data using CurveExpert's built-in sort feature. Reapply the spline fit, and everything should work as planned. If problem #2 is the issue, then I have to question the use of a spline to fit your data. I would recommend that one or more of the replicate values be removed, or you will be forced to use a regression-type model. Other reasons that splines might fail is because of the sheer number of data points. Round-off errors in the computations can accumulate, causing incorrect splines to be shown. Be careful when applying splines to very large data sets. Again, splines are usually used on small to moderate size data sets. And as always, as yet another reminder, make sure that your data set is scaled to reasonable numbers (preferably on the order of one). High or low values may be causing excessive round-off error, overflow, or underflow in the computations.

14.9 Lagrangian interpolation sometimes bombs with a floating point exception. What gives?

Check in your data set for a replicate value, i.e., two or more x data points are the same. If this is an issue, then I have to question the use of interpolation to fit your data. I would recommend that one or more of the replicate values be removed, or you will be forced to use a regression-type model. And as always, as yet another reminder, make sure that your data set is scaled to reasonable numbers (preferably on the order of one). High or low values may be causing excessive round-off error, overflow, or underflow in the computations.

14.10 I am using a Postscript printer driver, and my plots will simply not print in the orientation (portrait or landscape) that I tell it to. Is there anything that I can do?

There is a bug in the Postscript printer driver (perhaps it is fixed by now) that causes it to always print in the orientation specified in the Control Panel. Unfortunately, this means that you cannot adjust the orientation in CurveExpert. However, you may open the Control Panel, select Printers, select the Postscript printer, and click the Setup button. Here, you may adjust the orientation. Now plots from CurveExpert will print in the orientation that you have specified through the Control Panel.

14.11 CurveExpert saves the previous window position, and I don't want it to. How do I turn this behavior off?

From the File menu, choose Preferences. Deselect the checkbox that says "Save Window Position."

14.12 How do I obtain a debugging report to send to the author?

The key is running CurveExpert with the `-d` parameter. From the Program Manager or File Manager, choose Run. Use the Browse command to look in the CurveExpert directory and find the executable, which is CURXPT.EXE. Double click on this file. You should then see the Run... dialog again with the full path to CURXPT.EXE in the Command Line field. Add a `-d` to the string already in the dialog, and press OK. CurveExpert will now record information both onscreen and in a file called CXPTxxxx.REC, where xxxx is a unique number. If multiple debug files are saved, they will be numbered consecutively. If you are having a problem with CurveExpert, the best way to solve the problem is to contact me, describe the problem, and send the debug file mentioned above. Obviously, the problem should be reproduced while the debug file is being made.

14.13 Why are grid lines drawn as solid lines (instead of dotted) on some computers?

I wish that it was a bug in CurveExpert; if it were, I could fix it. However, this happens because of a bug in the display driver installed on your computer. Contact your video card vendor or Microsoft for the latest driver. If your driver is the latest version, then you may report the bug to the author of the driver (usually the company that manufactured the video card).

14.14 The Internet links pointed to by the menu items under Help|Internet are now broken. Is there any way that I can change the addresses to the correct home page?

As long as you know the correct home pages, the links can be easily updated. Run NOTEPAD.EXE, and open the file CURVES.INI in your Windows directory. Find the section marked [URL], and simply type in the updated CurveExpert home pages next to the equal signs. For example, if the CurveExpert home page has moved to www.curveexpert.net, you might change the following:

```
[URL]
Home=http://curveexpert.webhop.net
```

to:

```
[URL]
Home=http://www.curveexpert.net
```

14.15 The file import dialog never appears, now that I need it. How do I get it back?

Select “Tools|Options” and check the “Always show File Import Dialog” checkbox.

REFERENCES

1. Bevington, Philip R., *Data Reduction and Error Analysis for the Physical Sciences*, 1969, McGraw-Hill, Inc., Library of Congress QA278. B48.
2. Daniel and Wood, *Fitting Equations to Data*, 1971. Wiley & Sons, New York. Library of Congress QA297.D35.
3. Neter, Wasserman, and Kutner, *Applied Linear Regression Models*, 2nd ed. 1989, Irwin Press, Boston, MA. Library of Congress QA278.2 .N46.
4. Press, Flannery, Teukolsky, and Vetterling, *Numerical Recipes in Pascal*, 1989. Cambridge University Press, New York. Library of Congress QA76.73 .P2N87.
5. Seber and Wild, *Nonlinear Regression*, 1989. Wiley & Sons, New York. Library of Congress
6. QA278.2 .S425.
7. Ratkowsky, David A., *Handbook of Nonlinear Regression Models*. 1990. Marcel Dekker, Inc., New York. Library of Congress QA278.2 .R369.
8. Ratkowsky, David A., *Nonlinear Regression Modeling: A Unified Practical Approach*. 1983. Marcel Dekker, Inc., New York. Library of Congress QA278.2 .R37.

INDICES AND TABLES

- *genindex*
- *modindex*
- *search*