Simfit

Tutorials and worked examples for simulation,
curve fitting, statistical analysis, and plotting.
http://www.simfit.org.uk

Simulating a data set for a function of two variables that is exact to computer precision is the first step in simulation, so that random error can be added retrospectively to simulate experimental results.

Choosing the [A/Z] option from the SimFiT main menu is used to open the program **makdat**, which then allows you to create such almost-exact data from a library of models, or from a user-supplied model.

Before attempting to simulate a function of two variables it is essential to understand these issues.

1. Choosing the correct model

2. Choosing sensible parameters

3. Choosing a reasonable range for the independent variable(s)

4. Choosing a sensible number of divisions of the X and Y axes

5. Viewing the current simulation

6. Saving the simulated data to a file

Program **makdat** has access to the particular library that is contained in either `w_models.dll` for 32-bit applications, or `x64_models.dll` for 64-bit applications, and there are numerous versions of these where the standard library has been augmented for special user requirements. The standard version has models for one, two, and three variables as well as for single differential equations. For systems of differential equations program **deqsol** must be used.

Choosing a model

From the interface to the SimFiT library choose a polynomial of degree two which has the following form.

$$f(x, y) = p_1 x + p_2 y + p_3 x^2 + p_4 xy + p_5 y^2 + p_6$$

This can be simulated using the default parameters but the result is not very interesting.

2. Choosing sensible parameters

A more illuminating example would be the function

$$f(x, y) = x^2 - y^2$$

which can easily be simulated by editing the default parameters as follows.

$$p_1 = 0$$
$$p_2 = 0$$
$$p_3 = 1$$
$$p_4 = 0$$
$$p_5 = 1$$
$$p_6 = 0$$

3. Choosing a reasonable range for the independent variable(s)

In this case it is suggested that you use the default values which are

$$-1 \le x \le 1$$
$$-1 \le y \le 1.$$

Choosing the number of divisions should be considered carefully. For instance, accepting the default values of 20 divisions along the axes requires the simulation of 400 data points. With this example the surface can be viewed without problems but, for other models, the range and number of points will depend on the model selected and parameters chosen.

4. Viewing the current simulation

Once a data set has been simulated it is always possible to display a table of the independent variables and simulated values or a plot as follows.
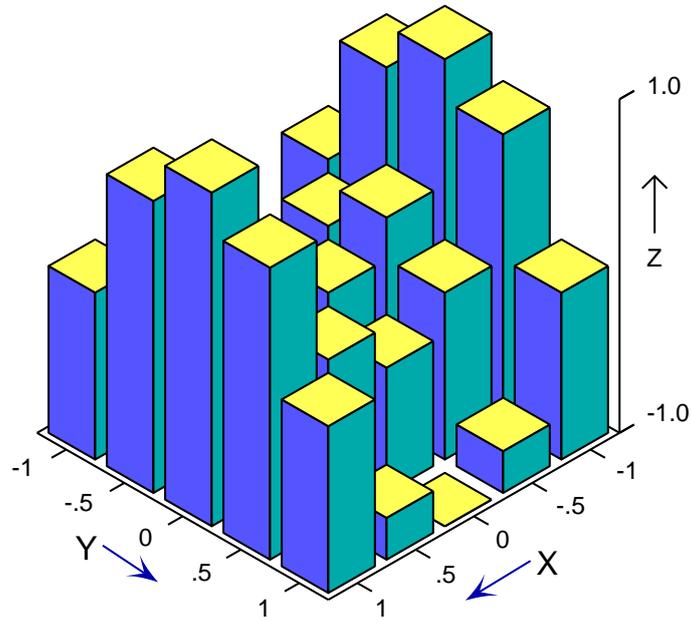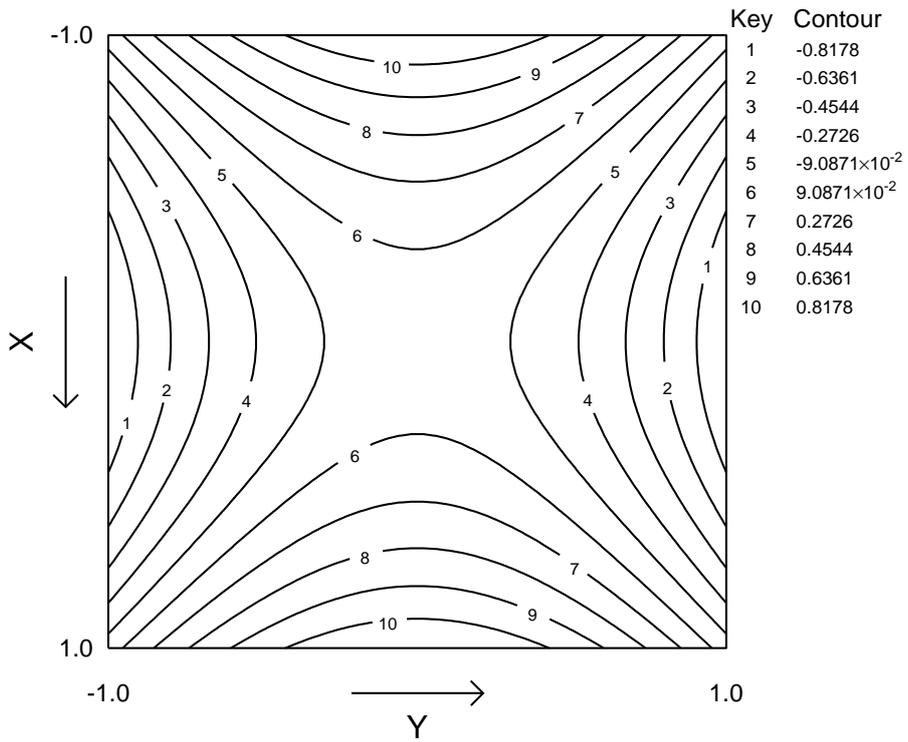
SIMFIT 3D plot for z = f(x,y)
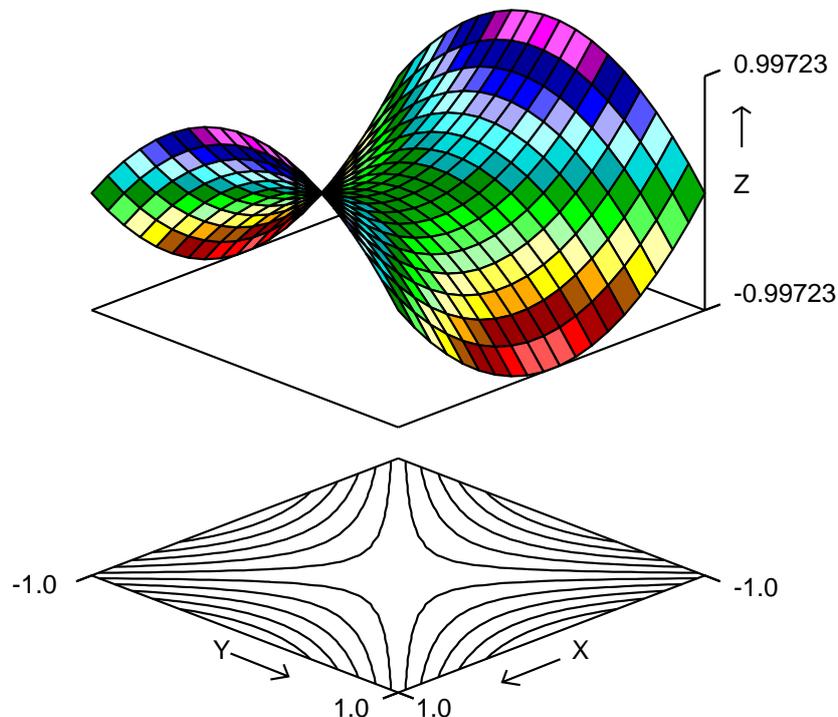
This example highlights the surface with colored rectangles (facets) and includes the corresponding contours, but there are a great many other ways to view the surface, including skyscraper and cylinder plots with or without error bars if the number of divisions is suitably small.

Note that, in order to generate a convincing contour diagram, it may be necessary to increase the number of divisions up to the maximum of 100, which would then involve the simulation of 10000 function values. A skyscraper plot with 5 divisions and a contour plot with 50 divisions for the same function and parameters are illustrated next.

SIMFIT 3D skyscraper plot for z = x^2 - y^2

SIMFIT 2D contour plot for z = x^2 - y^2

Key	Contour
1	-0.8178
2	-0.6361
3	-0.4544
4	-0.2726
5	-9.0871×10^{-2}
6	9.0871×10^{-2}
7	0.2726
8	0.4544
9	0.6361
10	0.8178

5. Saving the simulated data to a file

Curve-fitting files saved will have the independent variables in the first two columns, then the simulated values in the penultimate column, followed by a final column of weights. However such files cannot be used for plotting as will be explained below.

There are three reasons for saving the simulated data to a file.

- **Using the saved file for fitting**

  It is assumed that normally data are simulated in order to be fitted. So a final column of weights (usually 5% of the calculated function value) will be added to the saved file to make it a curve fitting file. The last column must be left in for weighting, or else replaced by a column of 1 if unweighted fitting is required. This is easily done using program **editmt**.

- **Using the saved file for plotting**

  In order to plot a surface SimFIT requires plotting files in a special format. For n_x divisions of the x axis and n_y divisions of the y axis a menu option from the plot concerned allows the creation of a vector plotting file with a column of $N = n_x n_y + 6$ values. After the usual title and row dimension (N) and column dimension (1), the first six values in such a SimFIT vector file are as follows.

  1. Line 1: n_x the number of x divisions
  2. Line 2: n_y the number of y divisions
  3. Line 3: initial x value
  4. Line 4: final x value
  5. Line 5: initial y value
  6. Line 6: final y value

  After these are the N function values in column-major format with priority x.

- **Using the saved file for adding random error**

  As the data written to file will be almost exact, they can be used to confirm that a model simulated with no added error can be fitted to return the correct best-fit parameters. If it is wished to add random error to simulate experimental data, the files can be input into program **adderr**, which can also overwrite the default weights by alternative weighting schemes if required.