Response surface

Menu:	QC.Expert	Optimization

This module helps to find optimum values of technological or other (independent) variables, using observations of some output (dependent or response) variable obtained experimentally. Assumptions are that a) the optimum independent variables settings (e.g. temperature, pressure, drying time) correspond to minimum or maximum of the dependent variable mean, b) dependent variable values were observed for various settings of independent variables, c) minimum or maximum of the dependent variable mean exists and is not very far from experimental settings tried. The *Response surface* module fits a model through the experimental data – complete Taylor polynomial of second order and tries to find its extreme by looking for a stationary point (a point with zero first partial derivatives). When extreme (minimum or maximum) exists, its estimate and confidence interval are given in the protocol. When the optimized model has no extreme, stationary point corresponds to a saddle point and no optimum setting for independent variables can be found. Optimum independent variables setting can be located outside of experimental region, the estimate is less reliable in such case however.

Data and parameters

Data are organized into columns. Each column has to have the same number of data points. Dependent and independent variables are selected in the Response surface methodology dialog panel, see Figure 23.

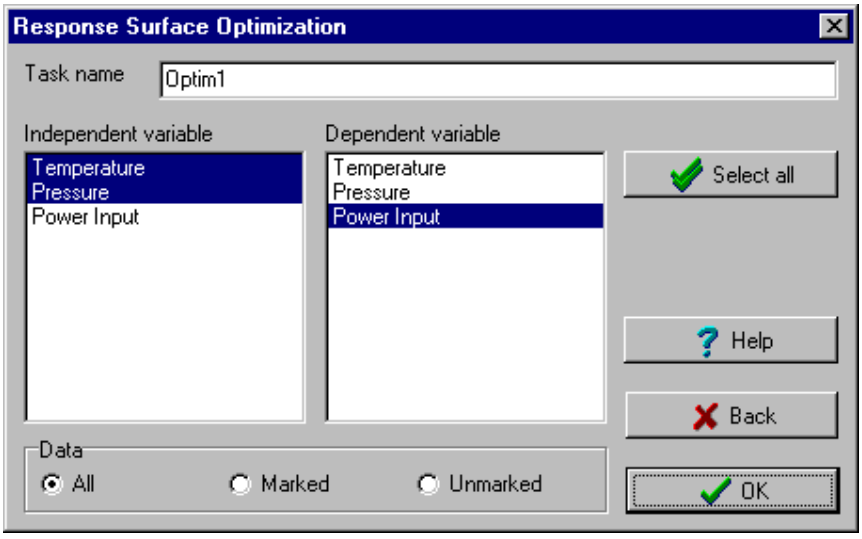


Fig. 1 Response surface methodology dialog panel

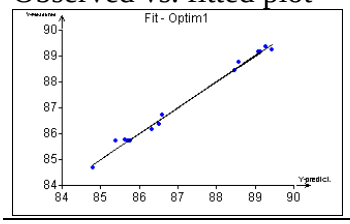
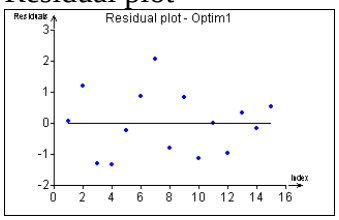
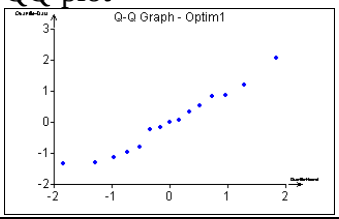
Protocol

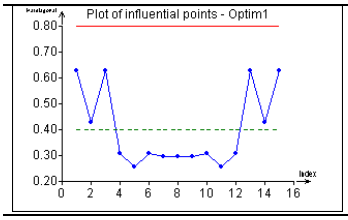
Number of variables	Number of independent variables, i.e. variables for which optimum setting is sought.
Number of observations	Number of observations, i.e. number of data rows.
Degrees of freedom	Difference between the number of observations and the number of parameters in the quadratic optimization model. Precision and reliability of the optimum setting estimate depends on the degrees of freedom to the number of data points ratio. The number of data points should be roughly comparable to degrees of freedom.
Stationary point type	Minimum, maximum, or saddle point. Saddle point does not correspond to an extreme, so that no optimum setting can be found. A saddle point can occur ① due to

	narrow or otherwise improper range of independent variables in the experimental dataset, ② due to large experimental error, ③ because no optimum setting exists around the experimental points.
Stationary point X0	If minimum or maximum was found, the optimum setting of independent variables is given.
Lower limit	Lower 95% confidence limit for the dependent variable optimum value.
Upper limit	Upper 95% confidence limit for the dependent variable optimum value.
Estimate	Estimate of the dependent variable optimum value.
Confidence interval	Confidence interval for the dependent variable value at the stationary point.
Mean absolute error	Mean absolute difference between dependent variable observations and the quadratic optimization model.
Residual sum of squares	Sum of squared differences between dependent variable observations and the quadratic optimization model.
Residual variance	Variance of the residuals after the quadratic optimization model.
Design conditioning number	A diagnostic value useful for checking the experimental settings layout (experimental design). If it is too large, the "Ill conditioned plan" warning is issued. It means that the independent variables are almost linearly dependent and some new experimental points should be added to decrease collinearity. Collinearity can be checked in the Correlation module.
Correlation coefficient	Multiple correlation, summarizing how well the quadratic model fits the experimental data.
Determinant	Determinant of the X'X matrix.

Graphs

Plots are generated only if the quadratic model is chosen.

Observed vs. fitted plot 	Shows how good is the fit of quadratic model. Observed values of the response are plotted on y-axis, while fitted values are plotted on x-axis. The closer points are to the y=x line, the better is the fit. If the points plot as a random cloud, it is likely that optimum will not be found.
Residual plot 	Standardized residuals, i.e. standardized differences between observations and the model fit. Residuals divided by the residual standard deviation are plotted on y-axis. Points, falling above 3 or below –3 horizontal lines correspond to data which are not fitted well by the quadratic optimization model. Sequential observation number is plotted on x-axis. When the points plot as a random cloud, the fit is rather good. If a trend is visible, the quadratic model does not fit data well.
QQ-plot 	This plot is useful for residual diagnostics. It is a useful addition to the Residual plot. The meaning is similar to the general QQ-plot meaning, described in the Basic data analysis module, 5.1.3.
Hat matrix diagonal plot	Shows points, which influence the fitted model heavily – much more than the rest of the data. Any influential points found should be checked

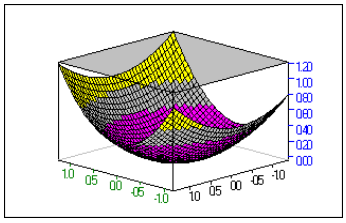


for validity and accuracy. When the influential points cannot be checked independently, it might be advisable to exclude them.

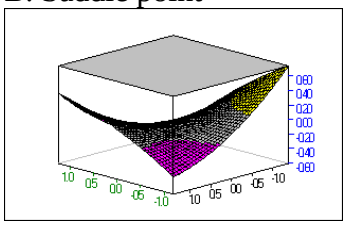
3D quadratic response surface plot

In case of two independent predictors, this plot shows the shape of the estimated quadratic response surface with visible minimum, maximum (A) or a saddle point (B). Hint: Use the *2D View* option to get the contour plot (C) where the position of the extreme may be more clearly visible.

A. Minimum



B. Saddle point



C. 2D-view