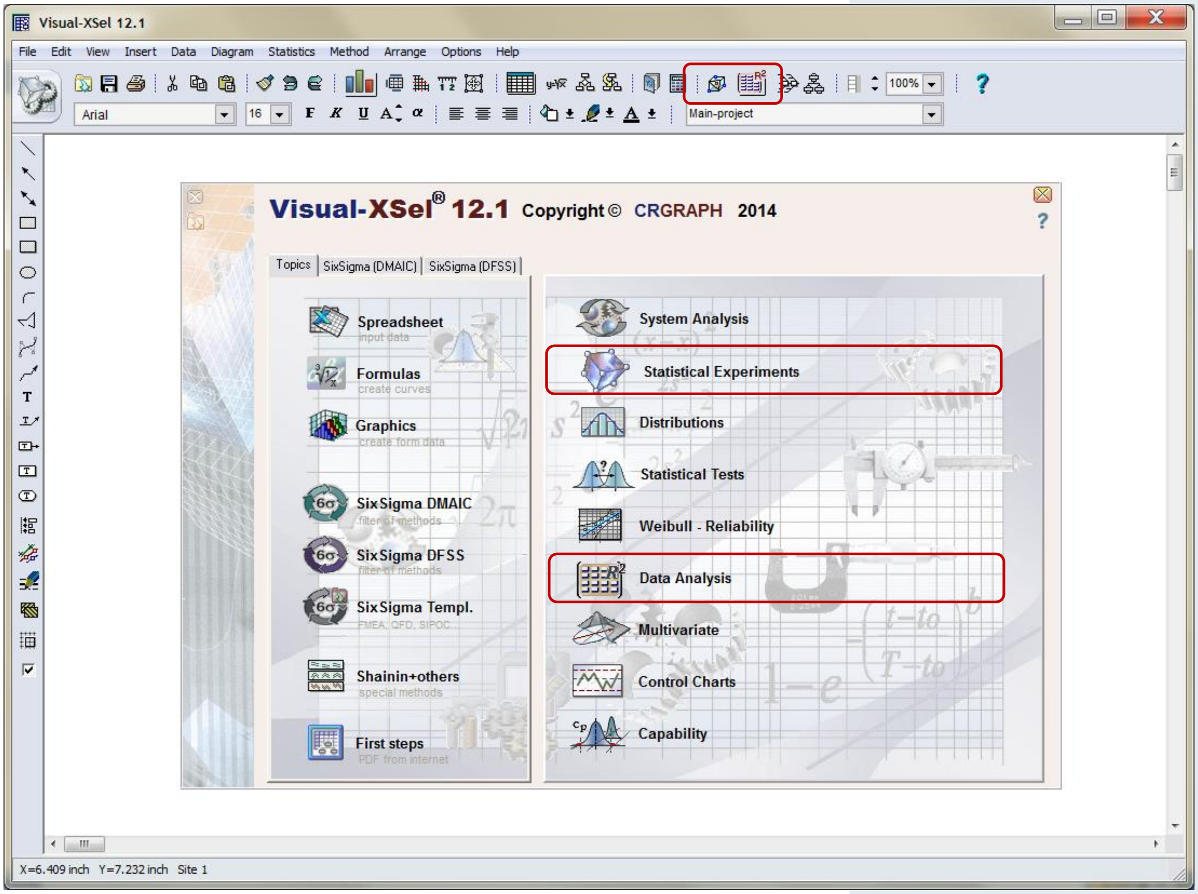
Introduction

Visual-XSel 12.1 is both, a powerful software to create a DoE (Design of Experiment) as well as to evaluate the results, or historical data. After starting the software, the main guide shows the direct access to the important functionality. Above the item Statistical Experiments, there is the System Analysis. It is possible with this method to find out the important factors for a DoE, by using mind maps.



More information to the statistical background one can find under:
http://www.crgraph.com/Statistics.pdf

To use the System Analysis, please have a look to:
http://www.crgraph.com/Systemanalysis.pdf

If you first join the program, it is recommended to use always the main guide (select the menu item ***File / New*** if the guide is not visible). Later one can use also the menu ***Statistics*** or the icons below.

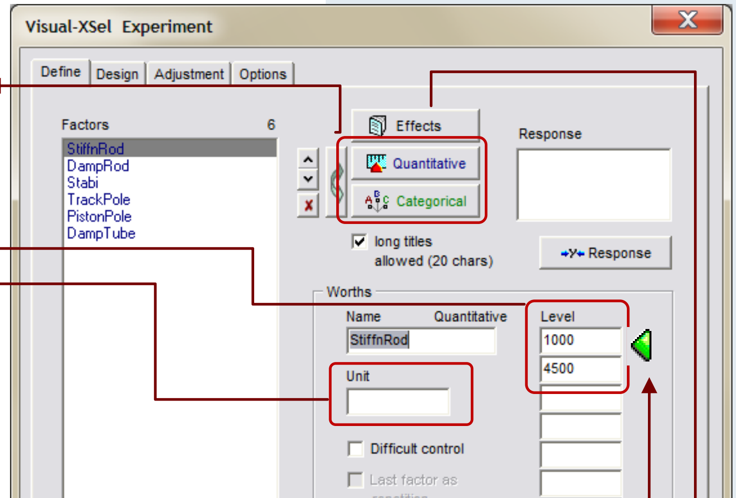
The Visual-XSel setup is available at: www.crgraph.com/Software.htm

On the following pages the most important steps are shown. First use the ***Statistical Experiments*** from the Main-Guide

The first step to create a DoE is to define the Factors (parameter). Push the button ***Quantitative*** for continuing measurable factors, or ***Categorical*** for factors described by text. The names should not be longer than 20 chars.

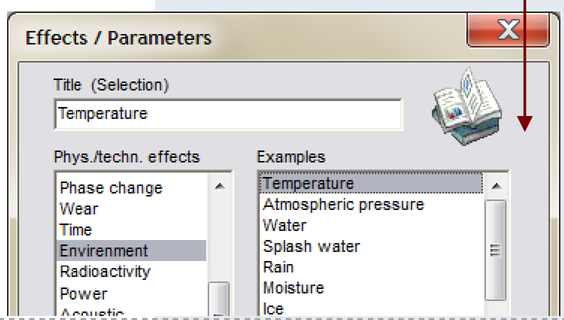
Type for any factor the ***levels*** of the combinations and use optional a ***Unit***.

If you press the button ***Response***, one ore more names of the last columns in the table can be defined, where later on the results of the investigation have to be added. The default name is "Y".



Hint: Please have a look to the green triangles (if visible), where the next input is.

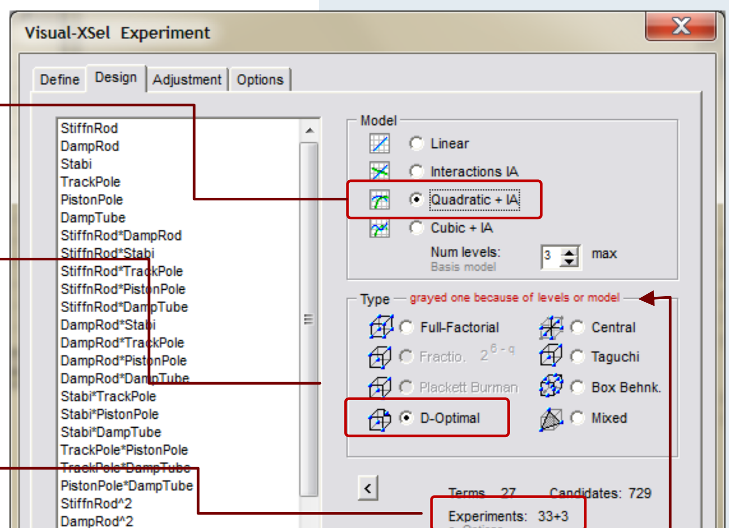
Under the button ***Effects*** a little library of the most important physically effects is available. Own factors can be added here. Use a double click to use a selected item, which is then shown in the Title.
This library is also a good check-list, so that no factor is forgotten in your project.



The next step is to define the ***Model*** and the ***Type*** of the experiment. In this example a quadratic model with interactions is selected (this is sometimes called Response-Service-Model).

The standard-type is ***D-Optimal***, which allow the most options. If the used type is not suitable for the model or the number of factors, you will get a message later, or some options are grayed.

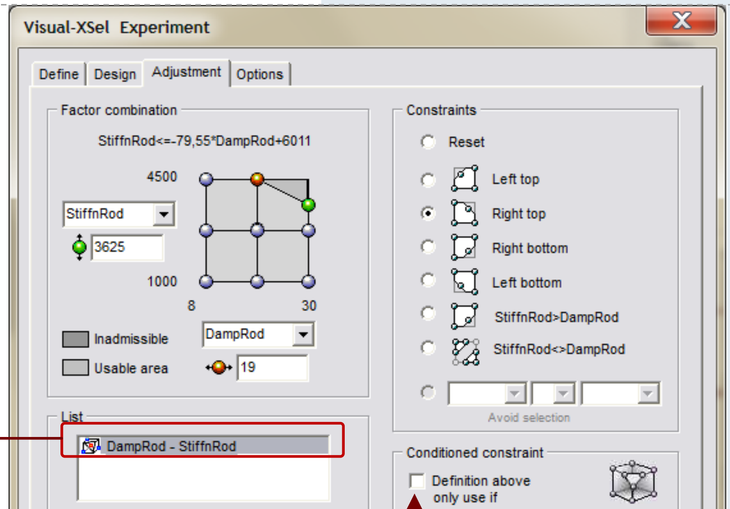
Right below the number of experiments is shown + 3 for the number repetitions.



Hint: Please have a look to the red comments in the top of the group-boxes. Here you have an explanation why some options are grayed.

On the next page it is possible to define ***Constrains***. Maybe there is a technical restriction, which is not possible. In the shown example the StiffnRod=4500 cannot be tested in combination with DampRod=30. But DampRod=15 is possible. To fix this constrain, push the button New, which is below of the ***List***.

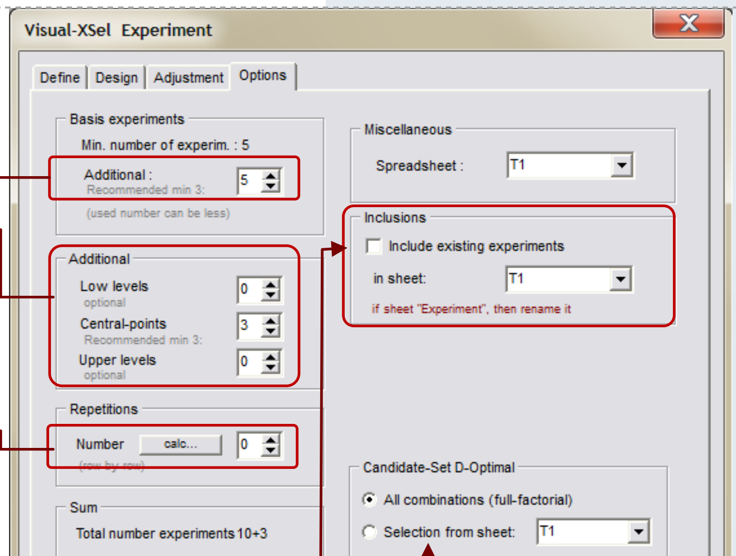
Note:
The view of the Factor combinations is only possible for quantitative factors and for D-Optimal design.



Hint: Constrains can be combined with other conditions, so that only an edge will be excluded from the DoE.

Under the rubric ***Options*** you can define additional experiments for D-Optimal design to ensure that the p-values can be calculated in the evaluation later on. The minimum is 1.

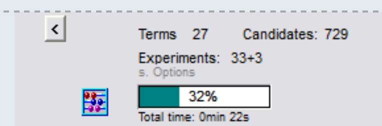
Under ***Additional*** you can define repetitions with the same factor values to determine the so called “pure error”. This is needed to get the information of the inaccuracy of the measurement-procedure (equipment).
Alternatively ***Repetitions*** for each trial can be set. Under the button ***calc*** it is possible to calculate how much trials are needed to detect the effect sure.



Especially for D-Optimal designs, a very important feature is the possibility to use already existing measurements. Use ***Inclusions*** and define the table where are these results. The column-names must be in the same order like in the list before (first col. is always no.)

Hint: For D-Optimal designs a pre-defined table can be used, from which the algorithm will try to get the best determinant. This is an alternative to constrains, may be if complex restrictions with categorical factors are excluded from the sheet.

Now start to create the plan with button ***Create***. If D-Optimal is selected the iterations begins.

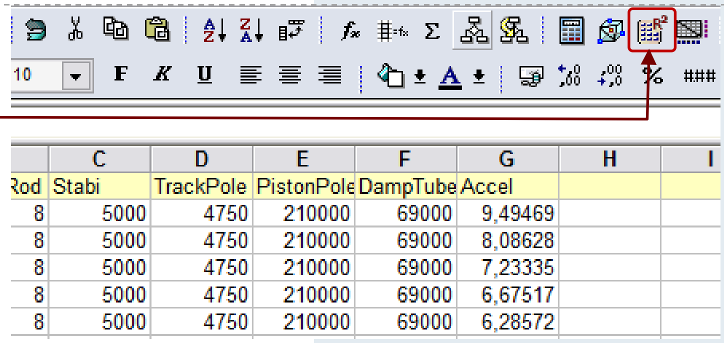


Finally the table with the DoE matrix is shown, where the empty column for the “response” have to be filled.

A	B	C	D	E	F	G	H
No	StiffnRod	DampRod	Stabi	TrackPole	PistonPole	DampTube	Y
1	2750	19	12500	5750	225000	210000	
2	1000	8	20000	1500	225000	210000	
3	4500	8	5000	5750	50000	400000	
4	2750	30	5000	10000	400000	210000	

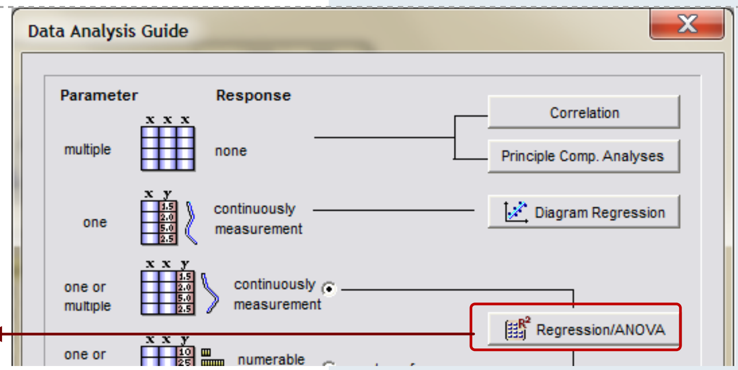
The results of the experiment can be loaded from the file *Example_MulReg.vxt* in the XSel path. The response name is here “Accel”

Use then the icon ***Data-analysis*** to open the dialog for the multiple regression. Alternatively click to ***Data-analysis*** in the ***Main-Guide***, shown at the first page and follow the speech bubbles.



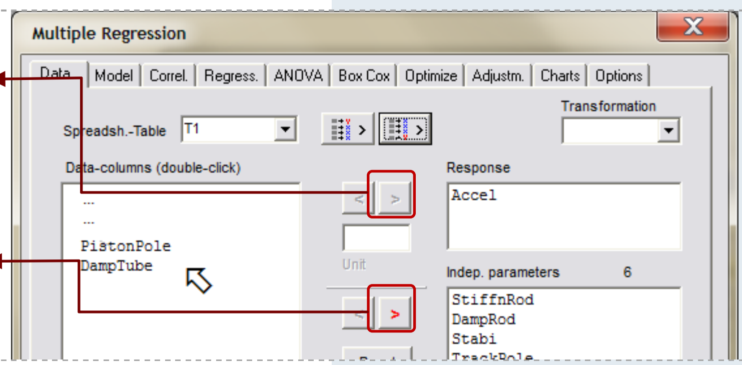
Depending from the entrance, normally the ***Data Analysis Guide*** will help you find out the right method for the used data. In this case, a normal Regression without transformation is suitable (continuously measurements).

For this click to the button ***Regression/ ANOVA***



In the dialog ***Multiple Regression*** the response and the factors (here independent parameters) must be selected with the respective buttons.

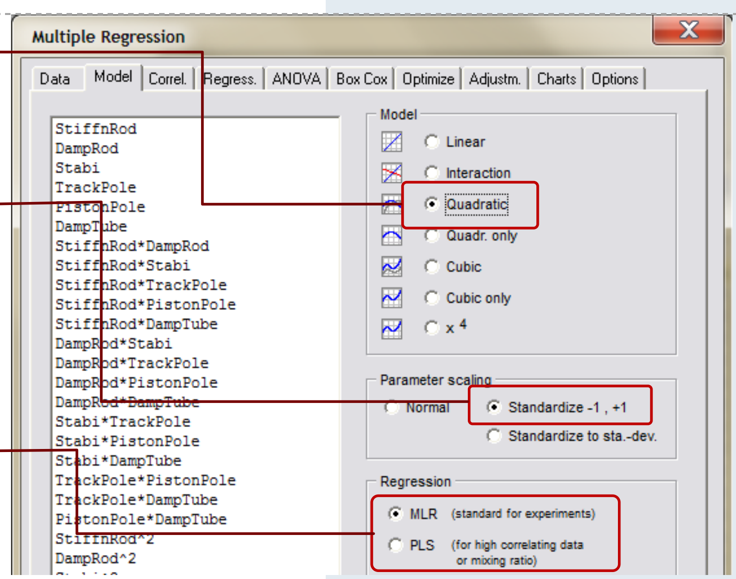
Note: If in the list of Data-columns a double click is used, the names will be moved in this field, where the button is red marked.



In the rubric ***Model*** the ***Quadratic*** Model has to be used analog to the experiment-definition.

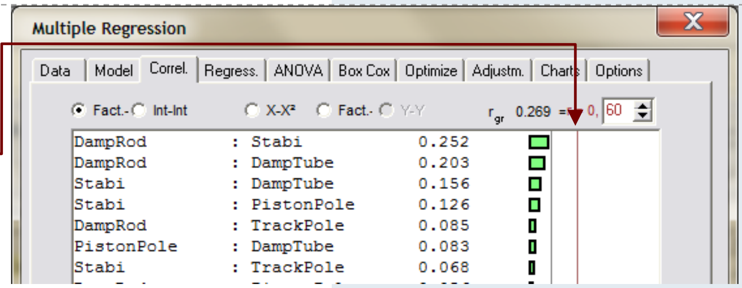
The ***Parameter-scaling*** should be always ***Standardize***. For the multiple regression, this is the best way for the least-square-method. Note that the respective coefficients of the model concerns to the standardization. Other software use often ***Normal***.

For the evaluation there are two methods: MLR (lease-square) and PLS (Partial-Least-Square). PLS is to use if the date doesn’t come from a DoE and have a great correlation. Use at first only the default MLR, Visual-XSel tests



the data if PLS will be better.

The next step is to check the correlation. Because of the DoE the data is not critical here. There is a limit with the a red line, to decide if the MLR is suitable. This limit comes more from experience and is not a statistical factor.
If this limit will be overstepped automatically a dialog-box appears to give some alternatives, how to proceed.

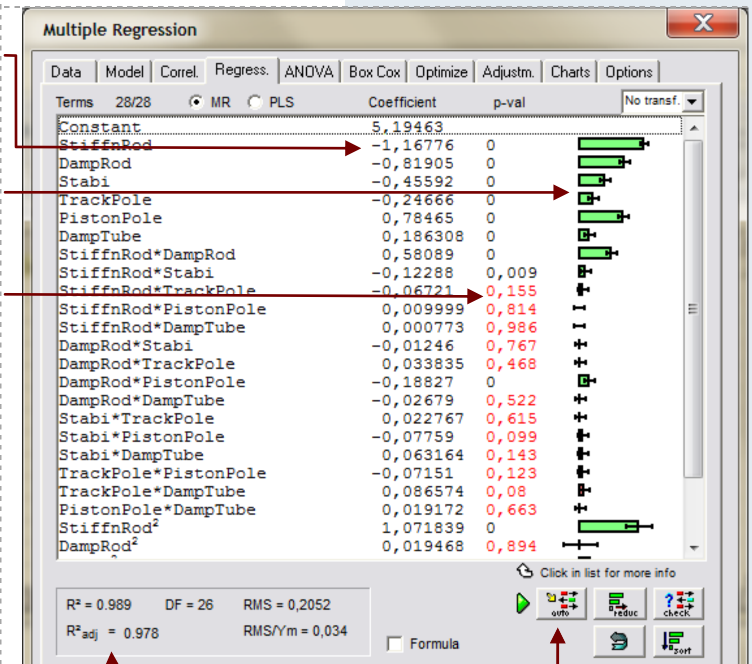


limit for hypothesis of independent data

The result of the regression is shown on the next rubric. The ***Coefficients*** are the weight of the influence of each term.
The visualisation of this coef. are the green horizontal bars on the right with additional confidence ranges.

The *p-val* (value) is the significance for the coefficients. If the defined limit of 0.05 is exceeded the recommendation is to exclude this term from the model. This will be done for all terms by step-wise regression (see button below).

After excluding the non signigicant term those will be grayed, but can brought back manually (sometimes it is better to decide by technical understanding than by statistical issues).

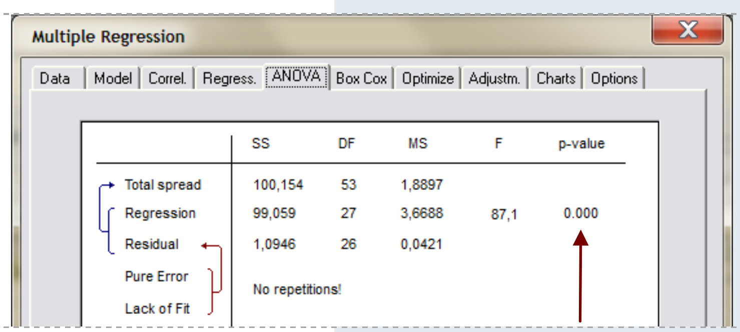


the coefficient of determination shows how much the model can explain the data

use the auto button to start the stepwise regression

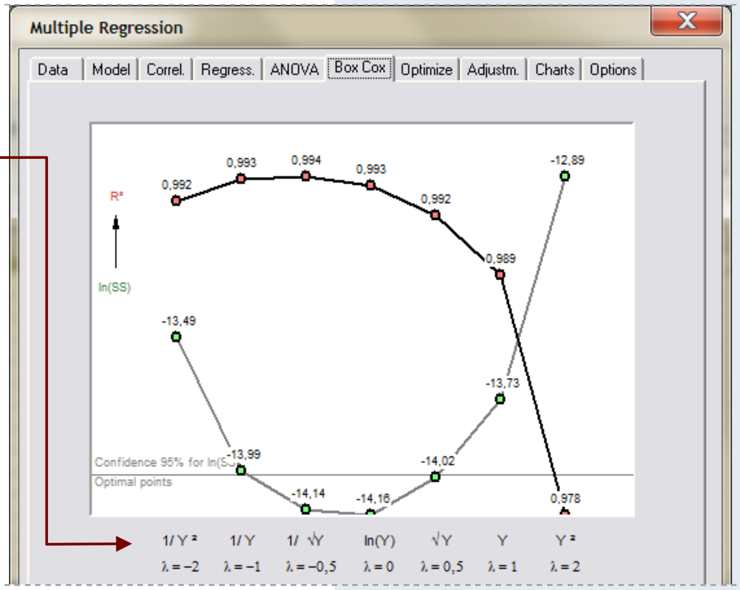
The (Model)-ANOVA gives enhanced information of how much trust one can have to the model.

For more information about the statistical values see the statistics-doc at the beginning.



this p-value shows the significance of the whole model

The so called Box-Cox-Transformation checks, wether the Y-data (response) should be used better by converting with mathematically standard formulas. The curve with the green points shows the special Box-Cox transformation with the goal to have the best normallity of the data (must be as small as possible). The curve with the red points shows the best coefficient of determination R² (must be as large as possible).
Note: Sometimes the best transformation between the two arguments is not the same.

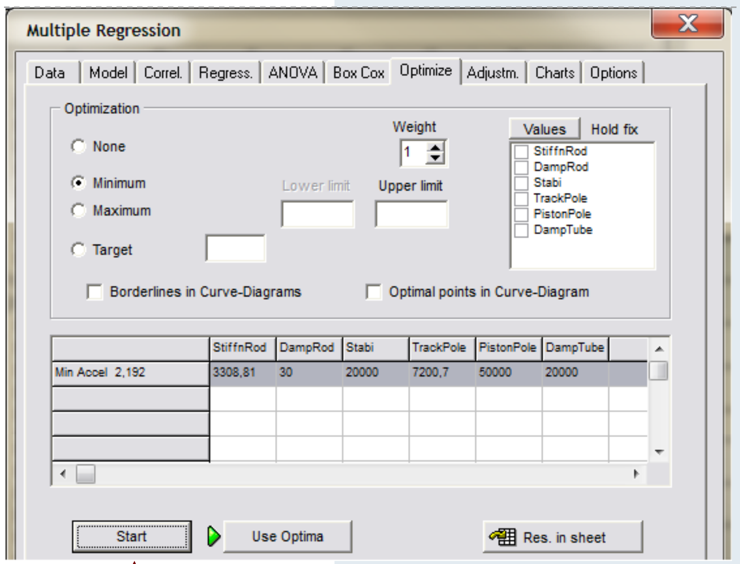


Recommendation: Use this transformation only, if there is a greate advantage by R², for example by lifetime data.

The optimizer calculate on the basis of the model the best-point, what you have defined (here for example the minimum of the response). This calculation finds mostly better parameter adjustments, than the best observation of the data in the table.

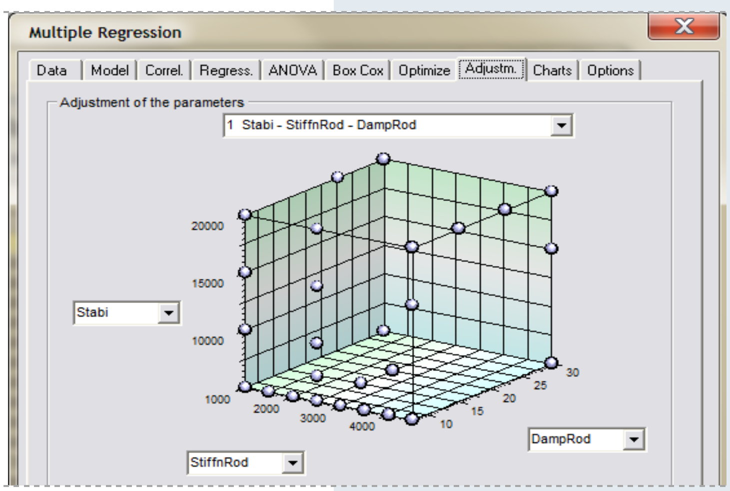
If there more responses the optimizer try to find the best compromize.
If there are restrictions of some parameters, one can fix this. So only the non fixed parameter will be adjusted.

Select the “Optimal points…” to get a mark in the “Curve-diagram” later.



Click to start after selecting the options

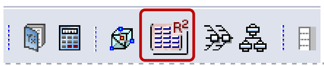
In the view of Adjustments one can judge, wether it is usefull to complete experiments, espacially if there are doubt about interactions. In this case it is recommended to add experiments of missing edges-points.

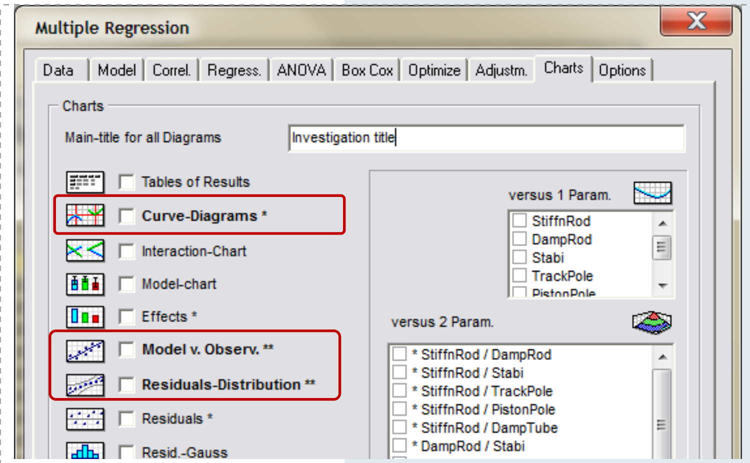


At the end one can select charts for the represantion on the main window. The most important charts are bold marked.

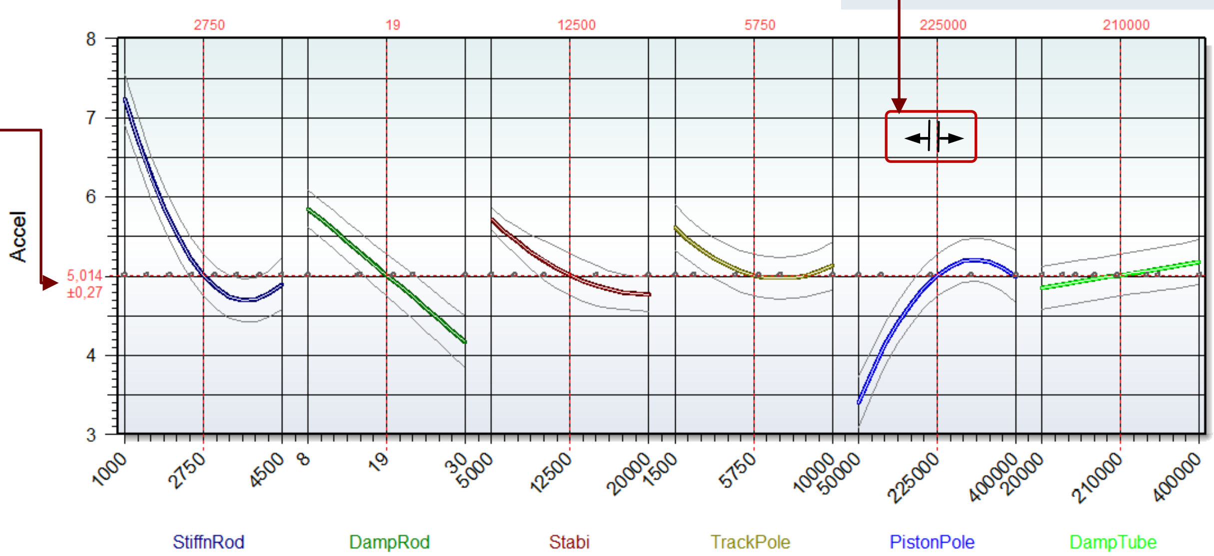
Note: do not forget to describe the project by a meaningful title in the top.

Click to the OK-button to create the charts in the main window. You can go back to the regression dialog at any time with the Data-analysis button.





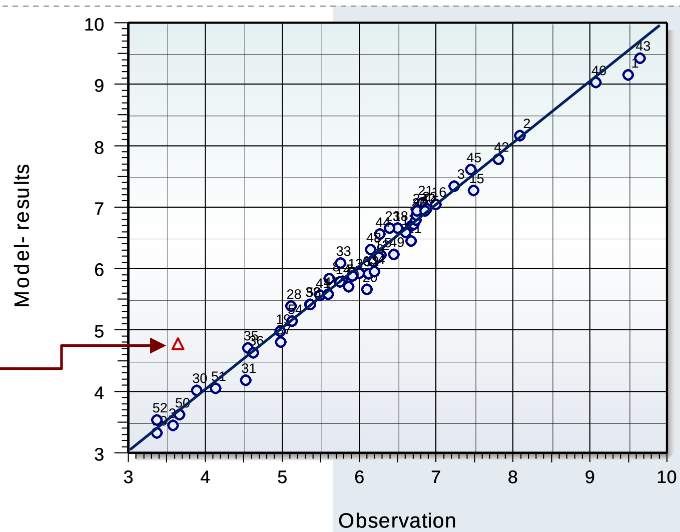
The curve-diagramm shows the function of the model grafically. The steeper the slope of the curve is, the higher is the influence of the parameters.
The vertical red lines represent the actual parameter adjustments with their values on the top.
The horizontal red line shows the result of in the response axis together with a confidence range.
Move the vertical red lines to change the parameter sets and to calculate the new result of the model.



The next diagram "Model versus Observation" gives one an overview where are the deviations between the regression model (function) and the measured values .

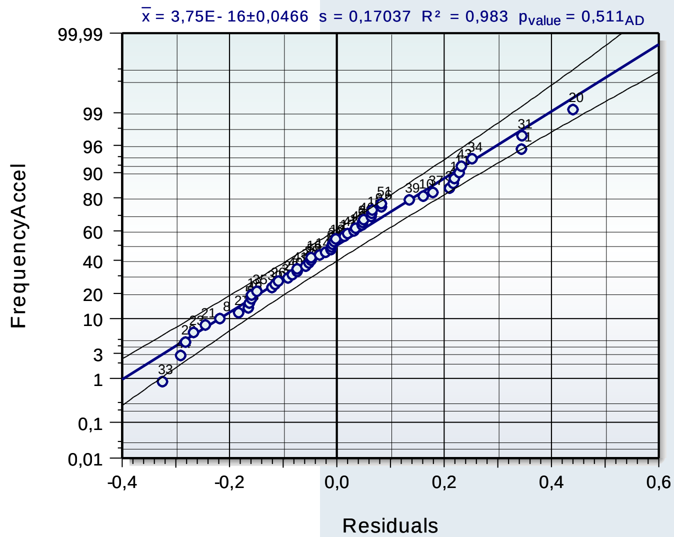
The best model is, if all points lie on the straight line. In this case the coef-ficient of determination R^2 would be 1.

If outliers exists they are marked in red (not including in the example data set)



Also a very important information about the condition of the model gives the “Residuals-distribution”.
The methode of least square requires that the residuals are normal distributed. If there are a deviation of a group of points this is a strong indicator that there are unknown disturbing factors or too much scatter.
To decide whether a deviation is critical, there is a p-value for the hypothesis of normality (if p-val<0.05 there is no normality.

Outliers will be marked in red also here.



In addition there are a lot of further charts, which are not described here.

If there are any suggestions or hints about this short introdution, please give us a feedback to

info@crgraph.de