Calibration

Menu:	QCExpert	Calibration

The Calibration module is most useful for analytical laboratories and metrological departments. It contains linear and nonlinear calibration models. Automatic detection of departures from linearity can be requested. Due to the fact that the module implements weighted regression, it can be successfully used for models with heteroscedastic errors. This feature can be useful namely for analyzing low-level measurement data (e.g. trace analysis).

Simple calibration models work with two variables. First is the measured variable of ultimate interest *X*, (e.g. concentration, viscosity, temperature). Second is the measurement device response *Y*, e.g. absorbance, voltage, resistance, number of particles. Generally, the calibration problem consists of two parts: (a) calibration model construction, and (b) application of the model constructed previously. When constructing the model, one uses measurement device responses: $y_1, y_2, y_3, \ldots$ to known values of the measured variable of ultimate interest, *X* (usually administered in the form of certified standards): $x_1, x_2, x_3, \ldots$. Dependence of *Y* on *X* is expressed by a regression model, which describes the relationship between known x_i and experimentally evaluated device responses y_i in the best way. QC.Expert™ uses either linear or quadratic regression model for calibration. The computations are based on direct application of either weighted or un-weighted regression *Y* on *X*. Compared to inverse regression (*X* on *Y*), the direct regression is more appropriate both statistically and logically. The regression fit encompasses step (a). Fitted model is applied when one looks for "the best" estimate of the unknown value of interest, *X*, based on one or more device response records, *Y*. This estimate, based on inversion of the regression relationship should be always accompanied by some form of uncertainty assessment, e.g. in the form of $(1 - \alpha)\%$ confidence interval ($\alpha = 0.05$ is selected very often in practice). Width of the confidence interval for *X* is related to the precision with which regression parameters are estimated (and hence to the confidence band width in the regression step (a)). Further, calibration limits related to noise variability and minimum reliably measurable value are computed (critical value, detection limit and quantification limit).

Data and parameters

The calibration module expects data in the form of a two-column-table. In the regression terminology, one column contains *X* , the explanatory variable, while the second column contains corresponding device responses *Y* , the dependent variable. These two columns have to be specified and used subsequently when entering analysis requests. If one wants to use the calibration model fitted on these two columns in order to estimate *X* for some additional unknown samples, their recorded responses *Y* have to be entered in additional columns. Recorded responses *Y* of individual unknown samples are entered as individual rows of one column. If there is more than one recorded response per one sample, additional replications should be listed as additional cells on the same row. The following table serves as an example of the situation where we had 5 calibration standards with values *X*=1.281, 2.558, 5.430, 7.373 and 11.59. The first four standards were measured repeatedly (twice each). After the regression model was fitted, we used the resulting calibration to analyze 4 samples from Czech rivers Upa and Labe denoted as *Upa A*, *Upa B*, *Labe AE*, *Labe AR*. The first two of these samples were analyzed repeatedly (three times each, replication denoted by *Replic1* to *Replic3*), next two samples were analyzed only once. The column denoted by *Sample* is intended only to hold comments only; it cannot be selected for further operations in the dialog panel. If calibration relationship estimation is all what is needed, entered data table will consist of the first two columns only. Additional data are not required then.

Table 1 Calibration data example

X	Y	Sample	Replic 1	Replic 2	Replic 3
1.281	25.53	*Úpa A*	33.69	33.74	33.73
1.281	25.58	*Úpa B*	39.25	39.25	39.27
2.558	51.37	*Labe AE*	50.6		

2.558	51.23	*Labe AR*	57.3		
5.430	106.4				
5.430	108.7				
7.373	148.4				
7.373	146.6				
11.59	233				

Dialog panel selections for the example just described is in the Fig. 1.

Task name is a text string, originally taken from the name of the data containing spreadsheet. It can be edited. The finally selected Task name will appear as a header in the resulting protocol. The *Calibration dependence* part specifies calibration model type. Here, the user must specify explanatory variable (*X*, values of the certified standards) and the dependent variable (*Y* , the measurement response recordings). In addition, a calibration model type has to be specified as either linear (calibration relationship is linear) or quadratic. Quadratic model is the simplest model allowing for curvature (nonlinearity) in the calibration relationship. When the *auto* choice in the *Calibration model* selection is invoked, automatic linearity check will be performed. This is done as follows: quadratic model is fitted first. A statistical test is used to test whether the quadratic term is significantly different form zero. If it is statistically significant, quadratic model is used. If it is not significant, all subsequent calculations are based on linear model. We strongly recommend using the *auto* choice, if the user is not sure about calibration model type. *Heteroscedastic errors* selection should be checked when one suspects that error variability of the *Y* reading depends on *X*. Heteroscedastic errors are quite common for instance when the calibration model is fitted across more than one order of magnitude (e.g. in trace analysis). When a heteroscedastic model is invoked, the calibration model is fitted by iteratively weighted regression procedure (IRWLS). The weights are then given as reciprocal values of the predicted residual variance, computed via nonparametric regression. The predicted variability can be inspected visually in the absolute residuals plot. Heteroscedastic model tends to give narrower confidence band in the intervals where the measuring device readings are more precise. If this increased precision occurs for *X* close to zero, heteroscedastic model based detection limits tends to be smaller.

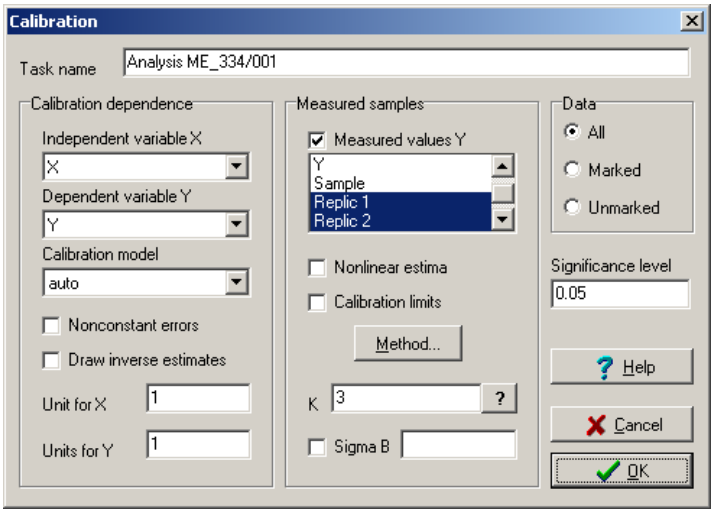


Fig. 1 Calibration module dialog panel

One should mark the *Plot the inverse estimates* selection, if inverse estimates are requested in the calibration plot. You should not use this choice if there are many calibration points and/or when confidence band of the calibration model tends to be wide, since then the resulting plot is hard to read. Text appearing in the *X units* and *Y units* fields is self-explanatory, denoting measurement units of these variables. It does not influence any calculations. It appears only in the final report of results. If they are not needed, the two fields can be left blank.

In the *Measurement device reading* part, the item *New samples reading* should be selected if one wants to estimate unknown *X* from additionally recorded *Y*. After selecting this item, roll-down menu can be opened, offering names of all columns that can hold the new device response readings. Indirect (sometimes imprecisely denoted as nonlinear) estimates are specially constructed estimates of unknown *X*, derived for the situation of the so-called statistical calibration, when both *X* and *Y* are considered to be random variables and the data are viewed as a two-dimensional cloud. The *Indirect estimates* selection invokes their calculation. This method can be used both in linear and nonlinear situation. In any case, the results are rather imprecise and confidence intervals for the *X* are not computed. These estimates should not be used when the calibration relationship is strong. When you mark the *Calibration limits* selection, resulting protocol will contain the critical value, detection limit and limit of quantification. Taken together, these values are referred to as calibration limits. QC.Expert™ offers five methods to calculate these limits, based on various literature sources. When the *Method…* button is pressed, a menu appears (Fig. 2), where one can select which of the methods should be computed. At least one method has to be selected.

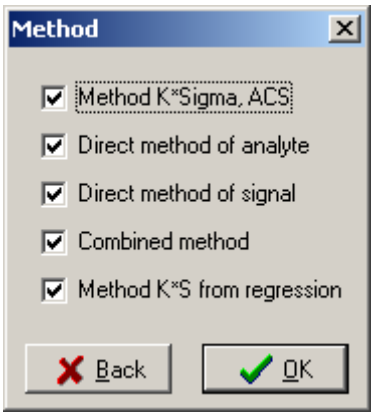


Fig. 2 Calibration limit calculation method

In the *K* field, the *K* coefficient can be entered, which will be subsequently used for the trivial calibration limits calculation method: K*Sigma. Typically, *K*=3 is selected often in practice. Since this coefficient has the meaning of the normal distribution quantile, its value should correspond to a chosen significance level α in order to keep the results comparable with the other methods. The „?“ button can be used to compute the value of *K* which corresponds to a selected significance level (e.g. for $\alpha = 0.05$, we have $K = 1.96$). On the other hand, if we insist on using $K = 3$, we should change significance level to $\alpha = 0.0027$ in order to be consistent across various methods of calculation and to get comparable results. When the standard deviation of the blank σ_{blank} is known (that is the standard deviation of the measurement device signal obtained without adding any sample, i.e. when $X=0$), the *Sigma B* selection is marked and σ_{blank} value entered. In the *Data* part, one can (as in other modules of the software) choose whether all data, or marked row data, or unmarked data will be used for computations. Data rows can be marked, using the button in the upper bar. *Significance level* must be a value smaller than 0.5 and larger than 0. It is used for all tests, and for calculation of (1-α)% confidence limits and calibration limits.

Further, we describe briefly various methods of calibration limits calculation and list their definitions. Since this is an analytical chemistry material, we use a common chemical terminology.

Y_C … critical level of *Y*. The smallest value of *Y* that can be reliably distinguished from noise. (a *Y* value, which is exceeded by noise with probability smaller than α). Values smaller than Y_C are considered to consist of noise only, respectively to be the blank readings.

Y_D … detection limit of *Y*. Analyzed substance can be safely proved (with probability 1−α) when it gives measuring device reading above this value. Probability of obtaining the reading $y > Y_D$ under the blank measurement condition is smaller than 1−α.

Y_Q ... quantification limit of Y. The value, above which true Y value can be estimated with the relative error, smaller than α. Quantitative analysis should not be conducted for samples giving measurement device readings under this limit.

X_C … critical value of X. It is tied to Y_C through the calibration model.

X_D … detection limit of X. Minimum value of X (e.g. concentration, weight) detectable by the given method.

X_Q … limit of quantification for X. Minimum value of X, which can be estimated with the relative error smaller than α. Only the X values above X_Q should be estimated quantitatively.

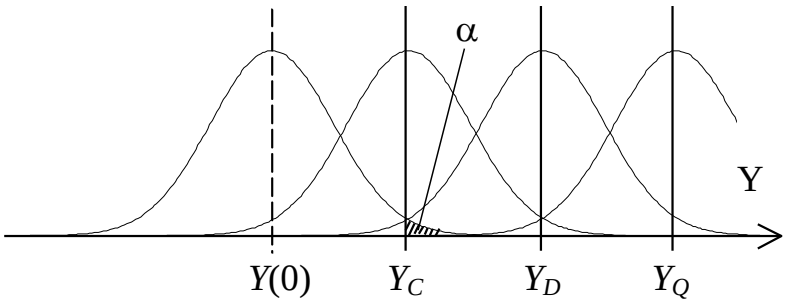


Fig. 3 Schematic draws of Y_C, Y_D and Y_Q

For K*Sigma method, see Figures Fig. 3 and Fig. 4.
$Y_C = K.\sigma$, $Y_D = 2K.\sigma$, $Y_Q = 3K.\sigma$

Sometimes, the $Y_Q = 10/3K.\sigma$ is used. For K=3, it corresponds to the 10 σ units. We suggest to choose K as the $(1-\alpha)$ – quantile of the standard normal distribution in order to keep the results comparable to the results of other methods. This method does not provide critical values of X. One can obtain them informally from the calibration plot, however. Estimate of σ is obtained either as the blank's standard deviation, or as the square root of residual variance from calibration model fitting.

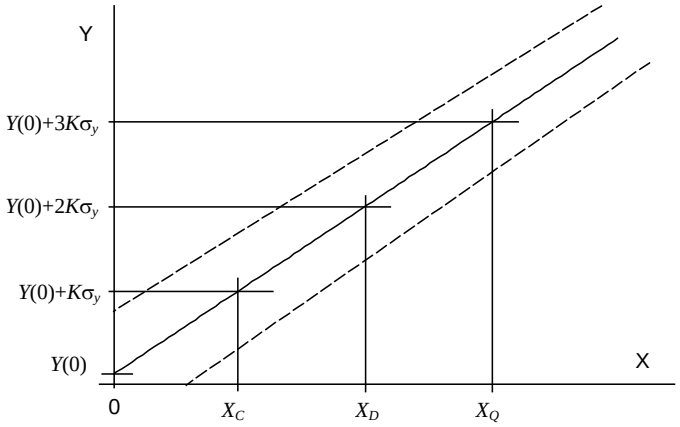


Fig. 4 Method K*Sigma

The following three methods use fully statistical properties of the calibration model and hence, they can be used to compute also critical values of X correctly. They are typically smaller (and hence more desirable) than those obtained from the K*Sigma method. The direct analyte method (Fig. 5) uses confidence intervals of X estimates.

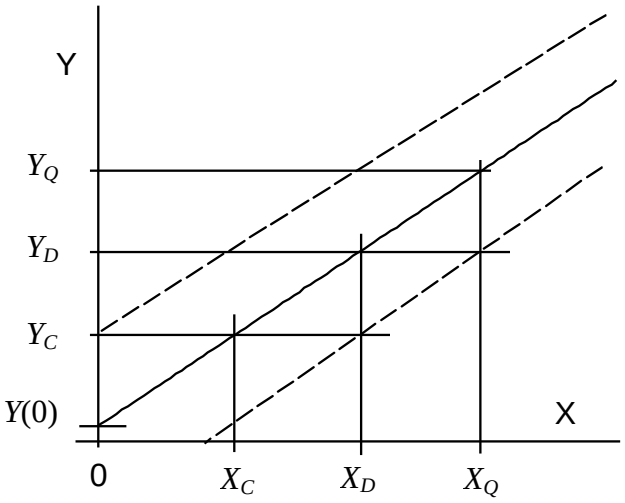


Fig. 5 Direct analyte method

The direct signal method uses confidence intervals for *Y*, Fig. 6.

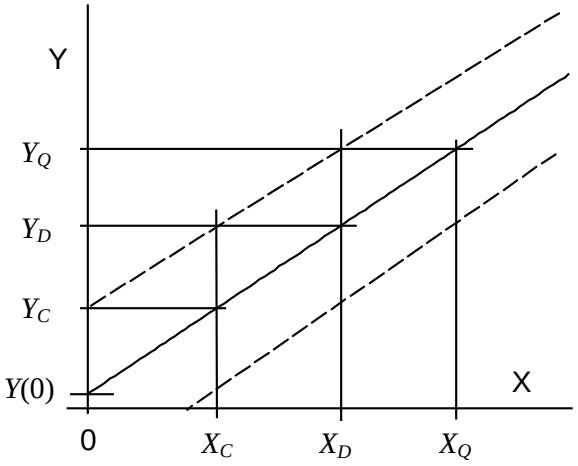


Fig. 6 Direct signal method

The Ebel and Kamm combined method combines elements of the two previous methods, Fig. 7.

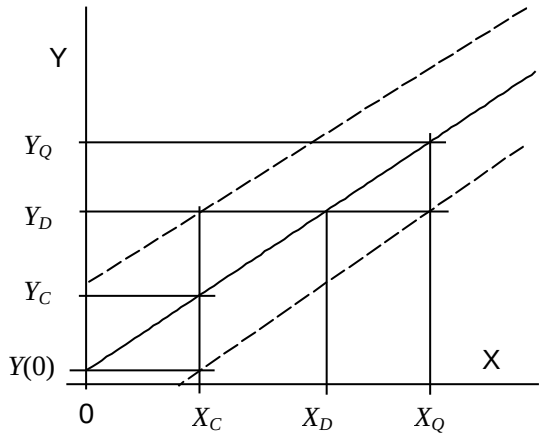


Fig. 7 Combined method, Ebel and Kamm

The last method implemented, *K*Sigma from regression* is similar to the first one (*K*Sigma*) only with the distinction, that for $K\sigma$, half-width of the confidence interval for situation with x=0 is used. That is the half-width of the regression confidence band at x=0, for a given significance level.

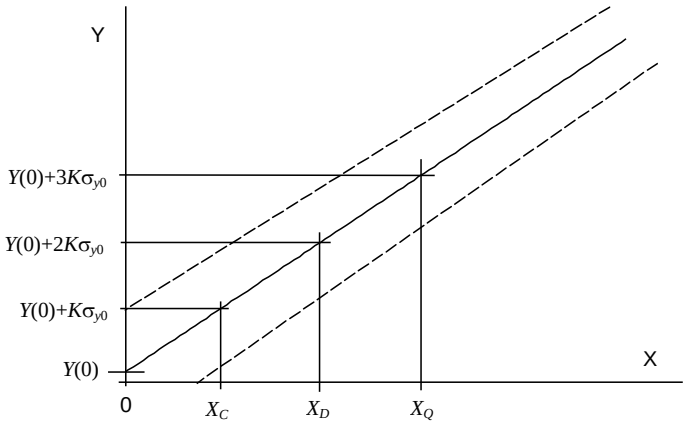


Fig. 8 Method K*sigma from regression

Inverse estimates. The main practical purpose of the calibration procedure is to be able to estimate unknown X from the recorded measurement device output Y. The required estimate is obtained by calibration relationship inversion (given by the previously estimated calibration model). One important thing to be kept in mind is the fact that, since the X is obtained from the random variable Y, it is random variable as well. Hence, it is not enough to report the point estimate itself, some measure of uncertainty should be attached. One possibility is to use the confidence interval (say the 95% confidence interval). Fig. 9 shows an example of the X estimate construction by the inversion of the calibration relationship. If we have some information about variability of the Y_i reading for the particular sample (obtained for instance from the repeated readings Y_{ij}, j=1,…n_i, then we are able to get a more realistic (even though sometimes wider) confidence interval for Y_i, and hence a more realistic estimate confidence interval for X_i. When the interval is constructed in this, more elaborate way, it reflects both variability related to the uncertainty of calibration relationship estimation and the current Y measurement variability, which is connected to a particular sample, see Fig. 10. This is one of the reasons why it is so important to replicate calibration measurements if it is possible. Whenever possible, the measurement outcome should be given in the form of interval ($x_{0.025}$, $x_{0.975}$), possibly listing the point estimate x_i as well. ***Remark:*** Confidence interval for the X estimate (obtained by the inversion of the calibration relationship) is not symmetric around x_i in general.

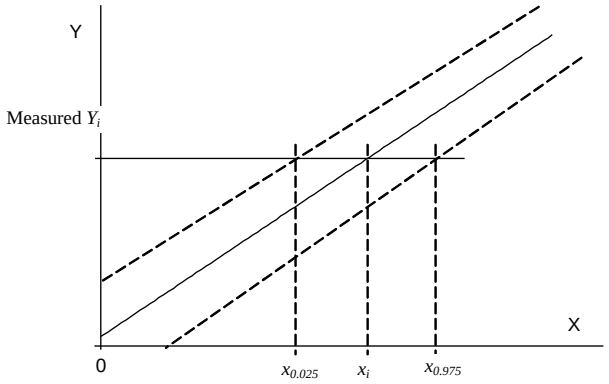


Fig. 9 Estimation by inversion for one measuring device reading

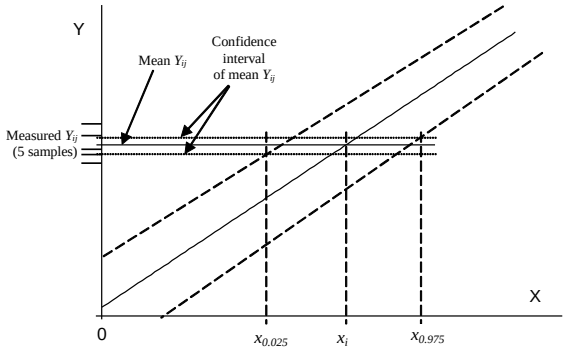


Fig. 10 Estimation by inversion for repeated measuring device reading

New method validation. The Calibration module can be successfully used to validate a new method through comparison with another method, established and validated previously. For this purpose, the established (or validated, or certified) method's results are entered as *X*, while the new method's results are entered as *Y* (entered so that *X* and *Y* values corresponding to the same sample appear on the same line). The samples, on which the *X* and *Y* pairs are measured, should cover densely whole range in which the new method is to be validated. When computation request is specified in the Dialog panel, automatic calibration model type selection (or *"auto"*) should be marked. The new method is validated, if the *linear* model is selected; its intercept is not significantly different from *zero*, while it's slope is not significantly different from *one*. Appropriate slope and intercept test results can be found in the *Intercept significance* and *Slope validation* paragraphs of the Protocol. Alternatively, one can use the module *Two-sample comparison – Paired comparison*, see the respective paragraph.

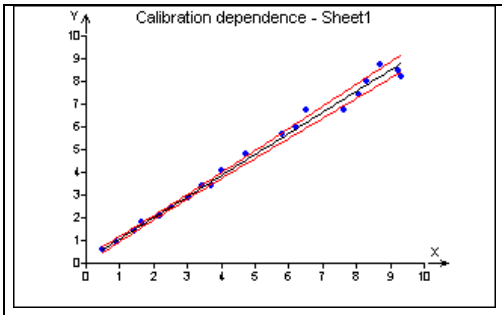
Protocol

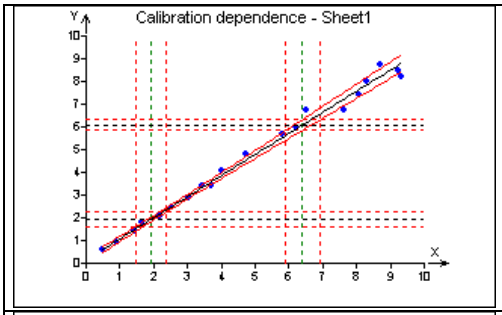
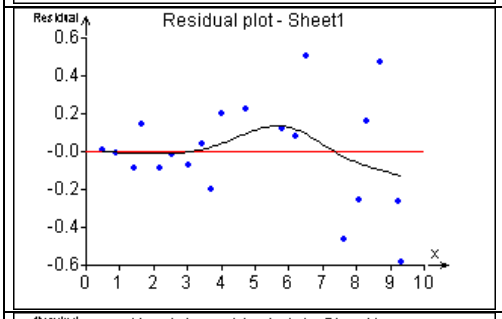
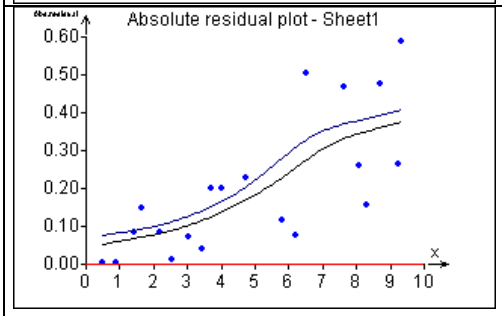
Task name	Name of the project, taken from the Dialog panel
Sample size	*n*, number of valid *X*,*Y* pairs used for calibration model estimation.
Significance level	Significance level α.required and entered by the user.
Calibration model selection	The requested way of calibration model selection (manual or automatic).
Calibration model type	Calibration model type used (linear or quadratic). Number of degrees of freedom ν for a given calibration model is equal to the sample size minus number of estimated parameters. For linear model, $\nu = n - 2$, for the quadratic model $\nu = n - 3$.
Is the model used feasible?	Linearity test. If we use (in the manual mode) linear model for data showing a substantial nonlinearity; or on the contrary, if we use quadratic model for linear data, we make a mistake that manifests by the measurement error increase, at least. If such a problem is detected by the statistical test implemented in the software, Protocol contains the word "*not feasible*"; otherwise it contains the word "*feasible*".
Weighted regression used?	Indicates, whether the weighted regression was used (*Yes* or *No*), that is whether a heteroscedastic errors model was invoked effectively.
Calibration model parameters	Information about calibration model parameters. If the model is linear, intercept "Abs" and slope (linear trend coefficient) will be reported. If the model is quadratic, the quadratic term coefficient is reported as well. Linear term does not have the first derivative meaning in that case.
Parameter	Parameter name: "Abs" = intercept, X=linear term coefficient, X^2=quadratic term coefficient

Estimate	Parameter estimate.
Std. deviation	Standard deviation of the parameter estimate (its standard error).
Lower limit	Lower limit of the 1-α confidence interval.
Upper limit	Upper limit of the 1-α confidence interval.
Intercept significance	Test of the null hypothesis that intercept is equal to zero. Result of this test is interesting when validating a new method via comparison with an established method, among other situations.
Value	Intercept value, copied from the *Calibration model parameters* paragraph
Conclusion	Conclusion of the significance test. If the intercept is *not significant*, than we have no particular reason to believe that the calibration curve (or line) does not go through the coordinate origin (more precisely: we cannot reject the zero intercept hypothesis at the significance level α). Even if not significant, the intercept should remain in the model (the *Calibration* module, unlike the *Linear* regression module does not allow models without intercepts anyway). This is on purpose, since models without intercepts can lead to unwanted confidence interval distortion at x=0 and to the situation, where it is impossible to compute calibration limits. When the intercept test is *significant*, it is possible to reject the null hypothesis about its' population value being zero, accepting the alternative that the calibration curve/line does not go through origin, i.e. that $Y(X=0) \neq 0$.
Slope validation	This test is very useful when validating a new method (respectively, when comparing two methods).
Value	Linear term coefficient value (when the model is linear, this is the derivative of the calibration relationship). It is copied from the *Calibration model parameters* paragraph.
Linear term coefficient=1	Conclusion of the unit slope hypothesis test (*Yes* or *No*). If the model is linear, the tested coefficients have meaning of the calibration relationship derivative.
Method sensitivity	Sensitivity of a particular method is defined as the measuring device response (Y) change, when X is changed by one unit. When model is linear, the sensitivity is equal to the slope. When model is nonlinear (quadratic), the sensitivity is given by the derivative of the calibration relationship and changes with the X value. Therefore, the software gives the sensitivity at four important points: at $x = 0$, at the lowest data value **min**(x), in the middle of measured data range – i.e. at (**min**(x) – **max**(x))/2, and at the highest data value **max**(x).
Selected K	Selected K for use in the K*sigma method.
Blank signal standard deviation	The value entered as Sigma B in the Dialog panel (if that was entered). That is the user-inputted value of σ_{blank}.
Computed blank signal standard deviation	If the blank signal's standard deviation is not entered in the Dialog panel, residual standard deviation is used instead. This value is used for calibration limits calculation by the K*Sigma method. Residual standard deviation is usually higher than σ_{blank}, the K*sigma does not give reliable results then.
Calibration limits	Critical value, detection limit and limit of quantification for Y and X. They are computed by the following methods: K*Sigma, direct analyte method, direct signal method, combined Ebel-Kamm method and K*Sigma from regression method.
Yc, Yd, Yq, Xc, Xd, Xq, Yq(10sigma), Xq(10sigma)	Yc = Critical value for Y, Yd = detection limit for Y, Yq = limit of quantification for Y. Xc = Critical value for X, Xd = detection limit for X, Xq = limit of quantification for X. Yq(10sigma) and Xq(10sigma) are alternatives to the quantification limit. For K=3, they correspond to 10σ, while Xq and Yq correspond to 9σ, for K=3.

Calibration table	This paragraph collects the results, whose computation gives main motivation to the calibration procedure. *X* estimates for a new, unknown sample are obtained by the estimated calibration relationship inversion from the measuring device response reading *Y* . If the data spreadsheet does not contain any *Y* readings, or when *New samples reading* was not selected, this paragraph does not appear in the Protocol.
Sample number	Integer indicating the sample number.
Estimate of *X*	Estimate of *X* by inversion.
Lower limit	Lower limit of the 100(1−α)% confidence interval of *X*, computed by inversion.
Upper limit	Upper limit of the 100(1−α)% confidence interval of *X*, computed by inversion.
Indirect estimate	Indirect estimate from the so-called statistical calibration. This estimate is computed only when the *Indirect estimate* was selected and when the variability of Y is large enough. Otherwise, column of zeros is printed here.
New samples readings	New samples readings of the measuring device, if they were entered. The NA acronym (not available) denotes missing data.
Residual analysis	Analysis of the residuals after the model.
Residual sum of squares	Residual sum of squares.
Mean absolute residual	Average of absolute values of residuals.
Correlation coefficient	Correlation coefficient estimate. ***Remark:*** Correlation coefficient *cannot* be used to judge whether linear or quadratic model is appropriate for given data!
Measurement number	Integer denoting numbering the *X*,*Y* pairs.
Measured X	Value of the known standard, taken from the inputted data.
Measured Y	Value of the measurement device reading, taken from the inputted data.
Computed Y	Value of *Y* , computed for a given *X* from the calibration model.
Residual	Difference: (Measured Y – Computed Y)
Weight	Weight attached to a particular measurement. If the *Heteroscedastic errors* choice was not selected, this column contains ones only.

Graphs

	Calibration plot, showing estimated calibration relationship, together with the confidence band (solid red line). If there are measuring device readings *Y* for an unknown sample, and the *Plot inverse estimates* selection is marked, corresponding *X* estimates are plotted as well. Horizontal dashed lines correspond to Y values (in case of repeated readings, they correspond to average and the related confidence interval. Vertical lines correspond to *X* estimates and appropriate confidence intervals, obtained by inversion of the calibration relationship. Upon double clicking on the plot, a new dynamic window is opened. In this window, further plot operations can be performed, see below.

	
	Residual plot. Plot of residuals e_i, fitted by a nonparametric, kernel estimate K(e_i) (black line). Substantial curvature can serve as a warning that the calibration relationship is nonlinear and is not described by the selected model satisfactorily. Observed curvature often relates to the presence of an outlier.
	Absolute residual, $\|e_i\|$ plot. Nonparametric, kernel regression estimate is superimposed (two dashed curves). The two curves depict estimates of standard deviation as a function of X, i.e. $\sigma(x)$. The upper (blue) curve is given by the square root of the residual squares fit, $\sqrt{K(e_i^2)}$. The lower (black) curve is given by the kernel fit of the absolute residuals, K($\|e_i\|$). The upper curve tends to be a better estimate of $\sigma(x)$ (when the residuals behave normally). This is because $\sigma = \sqrt{\frac{1}{n}\sum \sigma_i^2}$, a quantity which is estimated rather directly in this case. The lower curve is more robust.

Interactive calibration plot

Double clicking on calibration plot invokes a new window with an interactive plot. This window can be used for inspection or even reading the X estimates, while keeping all other interactive plot features. While the mouse is moving *above* the calibration curve, *Y* coordinate is shown together with the corresponding *X* estimate and confidence interval obtained by inversion, as seen on the Fig. 9. When the mouse is moving *below* the calibration curve, *x* coordinate is shown, together with the corresponding *Y* estimate, together with the confidence interval. When a particular detail is magnified (by zooming-in), plotted values can be read with a substantial precision. Nevertheless, when interested in *Y* estimates for a given *X*, or *X* estimates for a given *Y*, the "calibration calculator" should be used. The calculator can be invoked by clicking on the *Interactive estimates*, button.

The *Interactive estimates* window has 8 fields. *X* and *Y* cursor coordinates, relative to the interactive plot, appear on the uppermost line, originally. These values can be edited, however. Beneath the *X* field, there is the corresponding *Y* estimate, accompanied by its confidence interval. On the other hand, beneath the *Y* field, there is the corresponding *X* estimate, accompanied by the appropriate confidence interval. By clicking on the X or Y field, *X* or *Y* value can be entered from keyboard. After pressing <Enter>, appropriate estimates are computed. For instance, if we want to compute *X* estimate for the measurement device reading $Y = 25.7$, we click on the *Y* field in the *Interactive estimates* window. Next, we erase the content of the *Y* field and enter the value 25.7. The request is submitted by pressing <Enter>. Subsequently, *X* estimate, lower ($X-$) and upper ($X+$) limits or the $100(1-\alpha)\%$ confidence interval obtained by the inversion of the previously estimated calibration relationship appear beneath the *Y* field subsequently.

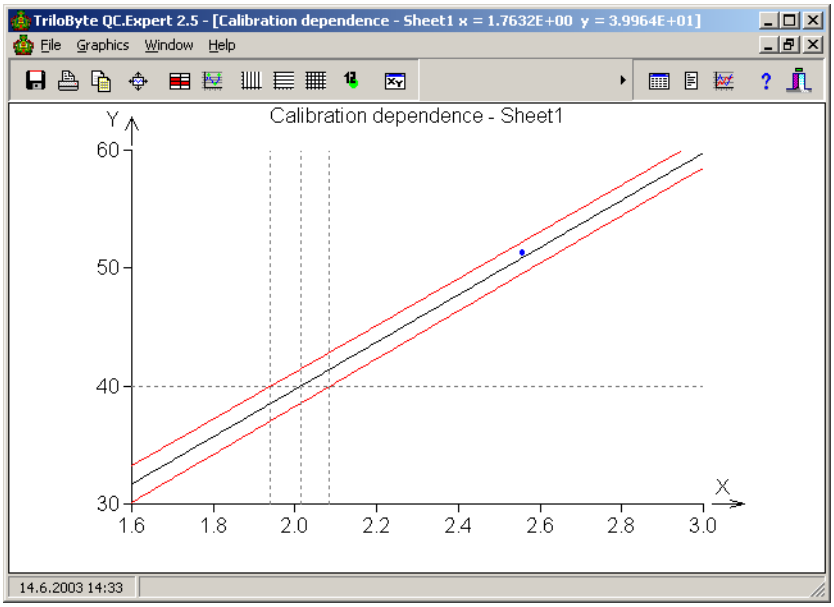


Fig. 11 Interactive estimates window with the calibration plot

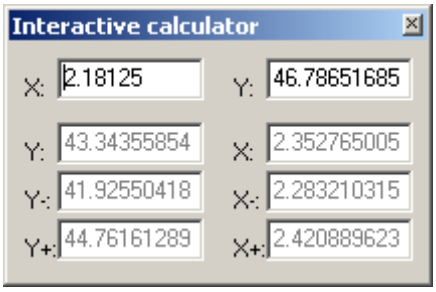


Fig. 12 Interactive estimates