LNE-RegPoly: A USER-FRIENDLY SOFTWARE FOR ESTIMATING THE CALIBRATION FUNCTION AND THE MEASUREMENT RESULT

C. Yardin[1] , I. Caré [2], J. Barbe[1]

[1]LNE, 1 rue Gaston Boissier, 75724 Paris cedex 15, France
E-mail (corresponding author): catherine.yardin@lne.fr
[2]CETIAT, 25 av des arts BP 52042,69603 Villeurbanne, France

Abstract

Firstly, LNE-RegPoly estimates a polynomial calibration function y = f(x) with weighted regression methods (generalized least squares). They take differently into account heteroscedasticity and correlation in variables. These methods stem from the ISO/TS 28037 standard and include ordinary least squares, weighted least squares, generalized least squares (with uncertainty only on y variable) and Gauss-Markov generalized least squares (with uncertainty both on x and y variables). In the latter case, LNE-RegPoly gives an estimate for the x unknown variables. The degree of the polynomial is up to 6.

Secondly, LNE-RegPoly computes predictions with the estimated function. The process is direct evaluation to calculate y_0 value (like $y_0 = f(x_0)$ where x_0 is a known predictor) and inverse evaluation to calculate x_0 value (like $x_0 = f^{-1}(y_0)$ where y_0 is a known predictor). The software evaluates the uncertainty of a prediction in propagating both estimated function and predictor uncertainties.

In LNE-RegPoly, input data are loaded from an Excel file. The data, and especially the structure of uncertainties, are scrupulously checked by the software. The sense of the regression like response value explained by the standard value or the contrary is indicated by the user.
To analyze correctly the results, LNE-RegPoly incorporates statistical tools like Fisher and chi-squared tests and also it gives some explanation when the input uncertainties are not validate. It shows a table of simple and weighted residuals and a graphical representation.

LNE-RegPoly is a graphical interface simple to use which leads you through the different steps and proposes a lot of help windows. The software has been validated with the ISO/TS 28037 examples and it is distributed with a collection real cases examples analyzed in LNE. LNE-RegPoly is freely available on line (www.lne.fr/fr/logiciels/Regpoly/logiciel-etalonnage-regpoly.asp).

1. Introduction

The VIM defines calibration as an "operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication".
Thus, calibration consists in evaluating the measurement model that transforms an indication into a measurement result. In the case of a measuring device used in a measuring interval, the calibration model includes a mathematical function like y = f(x) where the f function is different from the identity function. The function depends on both the chosen model (measurement result model, correction model) and the chosen x and y variables (for example, x is an indication and y a measurement result or the contrary).
The concept of calibration function is mentioned in the VIM but no definition is given; in the published articles, calibration function often associates an indication y to the x variable which represents both the standard value and the measurement result.
In this article, the concept of calibration function is extended to general function f associated to a calibration model whatever it is.
The function is estimated through data (values of the standards and indications) and the relation mentioned in the first step of the calibration operation. Then, this function is used to evaluate the target result, directly or with an inversion.

The calibration laboratory or the metrologist need to make two types of calculations :

- a function estimate calculation
- a predicted value calculation with the estimated function

These computations have been made with a software.

2. A software for calibration

2.1 – The need

From a statistical viewpoint, calibration is characterized by the following features:

- the calibration function is estimated through data with uncertainties both to x and y variables;
- the estimated calibration function can be used directly or with an inversion like in inverse regression problem;
- uncertainty associated to a prediction includes two components: one due to the estimated function and another due to the predictor.

Mathematical and statistical software products rarely bear functionalities adapted to this special case.
In LNE-RegPoly, the calibration function is estimated with weighted least squares methods which take into account the variation of the uncertainties in the calibration range (heteroscedasticity). The estimated function is used, directly or in inverse mode, to predict a new value. LNE-RegPoly propagates the uncertainties of input variables to its new value.
The software incorporates also tools for validating and interpreting the results as statistical tests, additional calculus or graphical representations. These tools are necessary in order to alleviate the lack of experiment on the subject.
The estimation of the function is treated in the paragraph 3 and its use in the paragraph 4. Before that, the structure of the software is described.

2.2 - Structure of LNE-RegPoly

LNE-RegPoly was developed in MATLAB® language and compiled to obtain an executable version which can be used without having to purchase MATLAB® software. This executable software consists of two graphical user interfaces: a home window (Figure 1) and a computation window (Figure 2). You can use the pushbuttons and drop-down lists to browse the software, . No skills in programming are necessary.

The computation window groups all functions necessary to estimate and use the calibration function. It is divided in five main menus: *I – Données étalonnage* (Input data), *II – Degré du polynôme* (Degree of the polynomial), *III – Méthode d’estimation* (Estimation method), *IV – Estimations* (Estimates), *V – Prévisions* (Predictions).

These menus are activated sequentially during the computation process. Each menu features a help icon which describes its content. The menus are presented in paragraphs 3 and 4.

Figure 1: Home window of the software

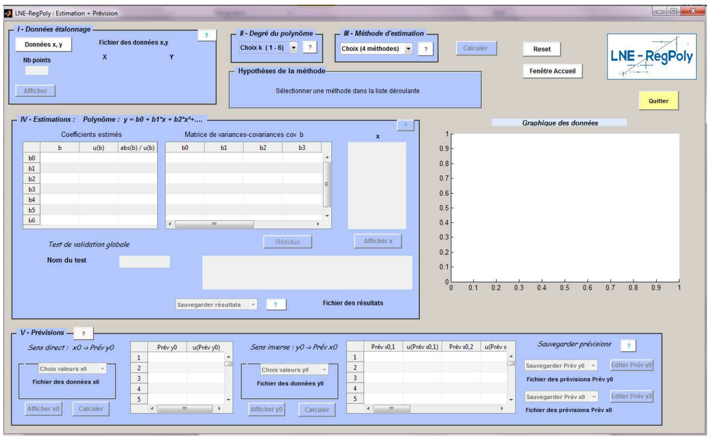

Figure 2: Computation window of the software

Input and Output data are saved in Excel files (Excel 2000). At each step, LNE-RegPoly tests the validity of input data and user's choices. In case of inadequacy, an explicit error message advises the user.

LNE-RegPoly is distributed with a user's manual and a set of examples. These examples are come from ISO standard [2] and practical cases analyzed in our laboratory. They belong to different areas of metrology science, show calibration functions linking quantities of the same or different kind, predicted values computed directly or with an inversion of the function, some analyses on the influence of the uncertainties and the covariances which are introduced progressively in the input data.

3. Estimation of the calibration function

LNE-RegPoly estimates the parameters of the calibration function y = f(x) without considering how the estimated function will be used at the next step to compute predictions. The estimation (process) is made with the menus (I) to (IV) of the computation window.

3.1 input data (I)

These are the values with the associated uncertainties of the x and y variables. They are obtained during the first step of the calibration operation. For example, they consist of the standard values, the instrument indications, the associated uncertainties and when the variables are correlated, the covariance matrix.
In general, this step required a preliminary analysis of the uncertainties to determine clearly what component is associated to each value. Moreover, the evaluation of the covariances can be a difficult task.
This step is not supported by the software. Nevertheless, one can test different correlation situations.

As the variables may be of two types (with or without uncertainty), input uncertainties need to be checked. The verifications are comprised of the following:

- all values of a variable have non zero uncertainty unless the variable is without uncertainty
- the symmetry and the invertibility of the variance-covariance matrix
- the level of the covariances and when appropriate, a message which informs the user that the correlations are too high (near the value 1 or -1).

LNE-RegPoly does not work with measurement repetitions.

3.2 Degree of the polynomial (II)
In LNE-RegPoly, the calibration function is a polynomial curve of degree k. It is written as:

$$y = b_0 + b_1 x + b_2 x^2 + ... + b_k x^k \qquad (1)$$

The user chooses the value of k and LNE-RegPoly computes the terms x^2, …, x^k . The polynomial function is defined with a constant term.
To compute residuals (deviation between estimated fuction and (x,y) couples) and realize statistical tests for the validation, the value of k should not exceed (n-2) where n is the number of (x,y) input couples. The maximum value of k is fixed to k = 6.This value is considered as sufficient for modeling a lot of instrumental calibration relationships.
It is not possible to switch the variables x and y directly in the software, the user can only specify the values of the two variables in the input data file.

3.3 Estimation method (III)
LNE-RegPoly proposes three methods, such as weighted least-squares where the weights depend on uncertainties on the variables x and y.
The WLS (weighted least-squares) are concerned by the case where x is known without uncertainty and y is known with an uncertainty which varies in the calibration range. The y values are not correlated.
The GLS_GGMR (generalized least-squares Generalized Gauss Markov regression) are concerned by the case where both x and y variables are known with an uncertainty which varies in the calibration range. Both the x and y values may be correlated.
The GLS_simples (generalized least-squares) are concerned by the case where y variables are correlated. In the case of calibration, the GLS_GGMR method is more appropriate. It is symmetrical since it estimates the same polynomial: y = f(x) and y = g-1(x) are identical. WLS method can be used when the uncertainty on one of the two variables is considered as neglected. When correlations are present, they are replaced by the GLS_simples.
LNE-RegPoly also contains the OLS (ordinary least-squares) method where only y has a constant and unknown uncertainty . OLS don't correspond to calibration context but they are the fundamental of the other methods and they give a good support to the interpretation of the results. OLS estimations are often compared to these of the other methods.
LNE-RegPoly evaluates the polynomial coefficients with their associated uncertainties. This estimated polynomial is used to evaluate the y_fitted values with an uncertainty. In the case of GLS_GGMR, LNE_Regpoly evaluates also the x values. Apparently, only this software does this.
The calculations discussed above are made with the matrix tools. When x has no associated uncertainty, the estimator has an analytical expression. When the two variables have uncertainty, the estimator of the unknowns (x values and polynomial coefficients) is defined by a non linear equations system. It has no analytical expression and it is computed with a numerical method. LNE-RegPoly uses the Gauss-Newton method and the algorithm proposed in [2,4] to compute the GLS_GGMR estimator.

3.4 Estimates (IV)
This menu deals with the estimates of the unknown parameters of the model and the results associated to the validation of the function.
LNE-RegPoly returns most of the results associated to regression methods: parameters estimates and associated covariance matrix, estimates for x variable and associated uncertainties (GLS_GGMR method), observed Fisher value (OLS), observed chi-squared and Birge ratio values (WLS, GLS_simples, GLS_GGMR), observed Student value...The estimated polynomial and its confidence interval are graphically represented. Moreover, it is possible to show classical and weighted residuals.

LNE-RegPoly contains tools for global and partial validation. The estimated polynomial is overall validated with a Fisher test (OLS method) or a chi-squared test (other methods). The statistic for the chi-squared test is the ratio between the data deviations to the function and the uncertainties associated to the variables. In the case of no validation, LNE-RegPoly indicates some factors which can explain this situation. For example, in the case of the chi-squared test, the uncertainty associated to the input data may be too small compared to the data deviations. With the chi-squared, LNE-RegPoly computes the Birge ratio which has a target value equal to one whatever the degree of the polynomial and the number of input couples (x,y).

Also, the user can verify that the residuals are not too large and no outlier is present.

The partial validation is realized with the Student statistic computed for each coefficient of the model:

$$\frac{|b_k|}{u(b_k)} \qquad (2)$$

When this ratio is too low (below its critical value), the coefficient b_k is not significant and the associated term x_k have to be removed from the equation (1).
Sometimes, a too high k value leads to a situation where all the coefficients b_k are not significant. Reducing the k value and removing the last term x_k allows to correct this problem. When the shape of the calibration polynomial is not well-known, it is recommended to estimate the function by increasing the k value gradually.

4. Use of the calibration function

The estimated function is used to predict a new value both for x or y variable. This operation is realized in menu “V – previsions” (predictions) with two options: "Prév y_0" and "Prév x_0".

4.1 Predictors (x_o known and y_o known)

The predictor is indicated with an associated uncertainty which can be null. For simplicity, LNE-RegPoly computes many predictions in the same time. It verifies that the predictor values belong to the calibration range with a maximum deviation of ± 10%. Thus, the possible deviation between specified and practical calibration range is considered.

4.2 Predictions (V)

A y_0 prediction is evaluated in using equation (1) at x_0 :

$$y_0 = b_0 + b_1 x_0 + b_2 x_0^2 + \ldots + b_k x_0^k \qquad (3)$$

LNE-RegPoly verifies that y_0 belongs to the calibration range with a maximum deviation of ± 30%. If it is not the case, the user is advised.
The uncertainty associated to the estimated parameters is propagated to the y_o value. The calculation is realized by using the law of propagation of uncertainties :

$$u_f^2(y_0) = u^2(b_0 + b_1 x_0 + b_2 x_0^2 + \ldots + b_k x_0^k) \qquad (4)$$

A x_0 prediction is evaluated in reverse calculation. This calculation often named "calibration" consists of evaluating the roots of the following equation:

$$y_0 - (b_0 + b_1 x_0 + b_2 x_0^2 + \ldots + b_k x_0^k) = 0 \qquad (5)$$

The number of the roots depends on the degree of the polynomial. LNE-RegPoly checks the obtained roots to eliminate :

- complex roots;
- roots which are outside the calibration range with a maximum deviation of ± 10%.

Evaluating the associated uncertainty is also more difficult than direct prediction case. The computation is realized in using the jacobian matrices of the polynomial derived from the parameters (J_b) and the x terms (J_x) :

$$u_f^2(x_0) = J_x^{-1} J_b U_b J'_b J'^{-1}_x \qquad (6)$$

LNE-RegPoly displays all validated roots and associated uncertainties.

4.3 Overall uncertainty

Equation (4) or equation (6) gives the uncertainty component due to the estimated function. Statistically, this uncertainty is associated to a mean value located on the function.
Another possible component is due to the predictor. For example, the uncertainty associated to an indication given by the device user. We give in the following the predictor uncertainty propagation formula used to obtain the uncertainty associated to the predicted value.
In the case of a y_0 prediction (forward evaluation), the uncertainty associated to x_0 is established using the law of propagation of uncertainty and equation (3):

$$u_f^2(y_0) = u_{x0}^2(b_0 + b_1 x_0 + b_2 x_0^2 + \ldots + b_k x_0^k) \qquad (7)$$

where the x_0 terms are supposed uncorrelated.
In the case of a x_0 prediction, the uncertainty associated to y_0 is established using the jacobian matrix of the polynomial in x_0:

$$u_{y0}(x_0) = \frac{u(y_0)}{J_{x0}} \qquad (8)$$

The two uncertainty components are supposed independent and summed quadratically.

5. Example

A flowmeter with laminar flow of type "Molbloc" is calibrated in the range 0 – 10 000 ml/min. The calibration method is dynamic gravimetric method which consists in determining by real time the mass of gas lost by a reservoir under pressure during an interval od fixed time. 5 points of flow was generated with the gravimetric bench and measured by the flowmeter placed at the end of the bench.

Données X-Y

	Débit lu	u(Débit lu)	Correction	u(Correction)
1	1.0000e+03	0.0029	-0.0928	0.9220
2	2.5012e+03	0.0029	3.1727	2.4420
3	5.0019e+03	0.0029	6.9365	4.5990
4	7.5029e+03	0.0029	6.9470	6.9530
5	1.0011e+04	0.0029	-0.1773	9.3000

Figure 3: calibration data

The calibration function is a polynomial of degree 2 with corrections and indications of the flowmeter : Correction = f(Indication).
The corrections are highly correlated with a correlation coefficient of 0.9. This is due to the values of the standards. The indications are supposed to be uncorrelated.

The polynomial was estimated with the GLSS_GGMR method and it was validated. These resultats are shown in Figure 4 below.

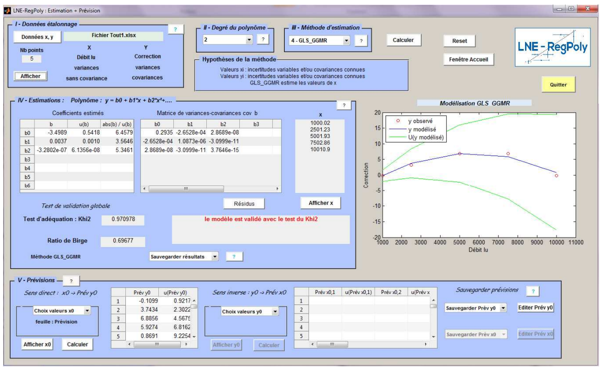

Figure 4: GLS_GGMR estimates

Then the polynomial is used to deduce correction values in the calibration range. As we can see in Figure 5, GLS_GGMR method gives uncertainties which are not under-estimated. This method propagates the standard uncertainties on the predictions.

Prédicteurs x0 et Prévisions y0

	Débit lu- x0	u(x0)	CorrectionPrév y0	u(Prév y0)
1	1000	0	-0.1099	0.9217
2	2500	0	3.7434	2.3022
3	5000	0	6.8856	4.5675
4	7500	0	5.9274	6.8162
5	10000	0	0.8691	9.2254

Figure 5: predictions

6 – Conclusion

LNE-RegPoly is suitable for a lot of measurement devices since it is able to estimate a polynomial up to the degree six. Hence, it can be used in different metrological areas (ex: gas flowmetry, chemical analysis, thermometry, ..).
By using the estimated function in direct or inverse sense, propagating the predictor and function uncertainties, LNE-RegPoly is intended both for calibration laboratories and their customers, which are the end users of the devices.
LNE-RegPoly helps the user to interpret and validate the estimated function. The set of practical laboratory examples ascertains this quality.
Moreover, LNE-RegPoly can be used out of the calibration framework. It allows to estimate a polynomial function with uncertainty associated both to x and y variables.

Technical and ergonomic improvements will be considered in the future.
Technical improvements will concern principally help tools to interpret the results with the presentation of the fitted y and new graphs. Some special moduli which regroup some functionalities can be created.
Ergonomic improvements will concern:

- the entry of the covariances and correlations;
- the possibility of data entry or data correction directly in the software instead of using exclusively files;
- the possibility of using only a part of the uncertainties;
- the switch of the x and y variables in the software.

Finally, LNE-RegPoly can incorporate a functionality allowing to compare the results obtained by different methods like Monte Carlo simulations or Bayesian techniques.

Acknowledgments
The author wishes to thank Hajar Cherradi, Anandan Lamartine and Sacko Kaba, trainees in the mathematical and statistical department, for their participation to the work on weighted regression methods and the LNE-RegPoly development.

References

[1] International vocabulary of metrology, NF ISO/CEI GUIDE 99, 2011.
[2] ISO/TS 28037, Determination and use of straight-line calibration functions, 2010.
[3] Yardin, C., “Estimer la droite d’étalonnage avec les moindres carrés généralisés et évaluer le résultat de mesure”, Revue française de métrologie, Volume 12-3, n°31, pp. 21 – 39, 2013.
[4] M. Milton, P. Harris, I. Smith, A. Brown and B. Goody, Implementation of a generalized least-squares method for determining calibration curves from data with general uncertainty structures, Metrologia, 43, S291-S298, (2006)