

## NPL REPORT MS 11

Software for determining polynomial calibration functions  
by generalised least squares: user manual

Ian Smith

December 2010



# Software for determining polynomial calibration functions by generalised least squares: user manual

Ian Smith  
National Physical Laboratory, UK

December 2010

## ABSTRACT

This report constitutes a user manual for Microsoft Excel-based software developed at the National Physical Laboratory for determining the best-fit low degree (1, 2, 3 or 4) polynomial calibration function to a set of data comprising measured values of stimulus and response variables and uncertainties associated with the values. Additionally, given measured values of the response variable and their associated uncertainties, the software can be used to obtain estimates of the corresponding values of the stimulus variable and their associated uncertainties.

NPL Report MS 11

© Queen's Printer and Controller of HMSO, 2010

ISSN 1754–2960

National Physical Laboratory,  
Hampton Road, Teddington, Middlesex, United Kingdom TW11 0LW

Extracts from this report may be reproduced provided the source is acknowledged and the  
extract is not taken out of context

We gratefully acknowledge the financial support of the UK Department for Business,  
Innovation and Skills (National Measurement Office)

Approved on behalf of NPLML by Bajram Zeqiri,  
Knowledge Leader for the Acoustics and Ionising Radiation Division

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Mathematical background</b>	<b>2</b>
2.1	Polynomial calibration functions . . . . .	2
2.2	Determining estimates of calibration function parameters . . . . .	2
2.3	Inverse evaluation . . . . .	3
<b>3</b>	<b>Installing and uninstalling the software</b>	<b>4</b>
3.1	Software user licence agreement . . . . .	4
3.2	Installing the software . . . . .	4
3.3	Uninstalling the software . . . . .	5
<b>4</b>	<b>Using the software</b>	<b>6</b>
4.1	Entering data and selecting fitting options . . . . .	8
4.1.1	Measurement data . . . . .	8
4.1.2	Data for inverse evaluation . . . . .	9
4.1.3	Selecting polynomial degree and type of fit . . . . .	10
4.2	Evaluating . . . . .	11
4.3	Messages . . . . .	14
4.4	Advanced GLS parameters . . . . .	19
4.4.1	$x_{\min}$ and $x_{\max}$ . . . . .	19
4.4.2	Tolerances on parameter and function values . . . . .	20
<b>5</b>	<b>Additional guidance on calibration functions</b>	<b>21</b>
	<b>Acknowledgements</b>	<b>21</b>
	<b>References</b>	<b>22</b>



# 1 Introduction

Calibration is an important aspect of many measurement procedures and often involves determining a calibration function that best describes the relationship between measured quantities. In many cases, the calibration function describes a response or dependent variable (often denoted by  $Y$ ) as a function of a stimulus or independent variable ( $X$ ) and the calibration process involves obtaining estimates of the parameters of the calibration function. A simple, but commonly encountered, form of calibration function is one in which the response variable is modelled as a polynomial function of the stimulus variable. In general, measurements are subject to measurement uncertainty, which means that the estimates of the calibration function parameters will have associated uncertainties (and covariances).

A common use of a calibration function is to determine an estimate of the stimulus variable corresponding to a measured value of the response variable. This process is termed “inverse evaluation” or “prediction”. Again, the measurement of the response variable will be subject to measurement uncertainty, and this uncertainty, along with the uncertainties associated with the estimates of the calibration function parameters, should be propagated through the calibration function model to provide an uncertainty associated with the estimate of the stimulus variable.

This report constitutes a user manual for Microsoft Excel-based software developed at the National Physical Laboratory for determining the best-fit low degree (1, 2, 3 or 4) polynomial calibration function to a set of data comprising measured values of stimulus and response variables and uncertainties associated with the values. Additionally, given measured values of the response variable and their associated uncertainties, the software can be used to obtain estimates of the corresponding values of the stimulus variable and their associated uncertainties.

The report is organised as follows. Section 2 introduces the basic mathematical terms and notation used to describe polynomial calibration functions and identifies the particular problems that may be solved by the software.<sup>1</sup> Section 3 gives information on installing and uninstalling the software. Section 4 describes how to use the software, detailing how the user provides and processes measurement data, lists the error messages that might appear, and discusses setting non-default values for some fitting parameters for cases involving “difficult” measurement data.

---

<sup>1</sup>It is not the intention of this report to describe in full detail the mathematical approach to solving the problems treated by the software. Such information may be obtained from other sources referenced in this report.

## 2 Mathematical background

### 2.1 Polynomial calibration functions

For reasons of numerical stability, polynomials are not expressed in terms of monomials (i.e., powers of the variable  $X$ ) but are expressed in terms of a Chebyshev basis in a normalised variable  $\tilde{X}$  in the interval  $[-1, 1]$ , viz., a polynomial of order  $n$  (degree  $n - 1$ ) in  $X$  in the interval  $[x_{\min}, x_{\max}]$  is written as

$$p(\mathbf{A}, X) = \sum_{j=1}^n A_j T_j(\tilde{X}),$$

where

$$\mathbf{A} = (A_1, \dots, A_n)^\top, \quad \tilde{X} = \frac{(X - x_{\min}) - (x_{\max} - X)}{x_{\max} - x_{\min}},$$

and  $T_j(\tilde{X})$  is the  $j$ th order Chebyshev polynomial of the first kind [1, 4].

A polynomial calibration function can therefore be expressed as

$$Y = p(\mathbf{A}, X),$$

where  $Y$  represents the response variable,  $X$  the stimulus variable and  $\mathbf{A}$  the (vector of) calibration function parameters.

### 2.2 Determining estimates of calibration function parameters

Let  $(x_i, y_i)$ ,  $i = 1, \dots, m$ , denote a set of  $m$  pairs of measured values of, respectively, stimulus and response variables  $X$  and  $Y$ . Associated with these measured values are standard uncertainties  $u(x_i)$  and  $u(y_i)$ ,  $i = 1, \dots, m$ .

The software classifies the calibration problem as belonging to one of two categories depending on the uncertainty information to hand:

1. If there are non-zero uncertainties  $u(y_i)$  associated with the measured values  $y_i$  and the uncertainties associated with the measured values  $x_i$  are zero or may be considered to be negligible, then the problem is referred to as one of “ordinary least squares” (abbreviated to OLS).<sup>2</sup>

Mathematically, the problem is to determine estimates  $\hat{\mathbf{a}}$  of the calibration function parameters  $\mathbf{A}$  that minimise the sum of squares of weighted residuals

$$F_{\text{OLS}} = \sum_{i=1}^m \left\{ \frac{y_i - p(\mathbf{A}, x_i)}{u(y_i)} \right\}^2, \quad (1)$$

---

<sup>2</sup>The term “weighted least squares” is also often applied to this case.



and the covariance matrix  $\mathbf{V}_a$  associated with  $\hat{a}$ .

2. If there are non-zero uncertainties  $u(x_i)$  and  $u(y_i)$  associated with, respectively, the measured values  $x_i$  and  $y_i$ , then the problem is referred to as one of “generalised least squares” (GLS).<sup>3</sup>

Mathematically, the problem is to determine estimates  $\hat{a}$  of the calibration function parameters  $\mathbf{A}$  and  $\hat{x}_i$  of the stimulus variables  $X_i$ ,  $i = 1, \dots, m$ , that minimise the sum of squares of weighted residuals

$$F_{\text{GLS}} = \sum_{i=1}^m \left[ \left\{ \frac{y_i - p(\mathbf{A}, X_i)}{u(y_i)} \right\}^2 + \left\{ \frac{x_i - X_i}{u(x_i)} \right\}^2 \right], \quad (2)$$

and the covariance matrix  $\mathbf{V}_a$  associated with  $\hat{a}$ .

### 2.3 Inverse evaluation

Given a measured value  $y$  of the response variable, its associated standard uncertainty  $u(y)$ , calibration function parameter estimates  $\hat{a}$  and associated covariance matrix  $\mathbf{V}_a$ , the problem is to determine the value  $x$  of the response variable such that

$$y = p(\hat{a}, x),$$

and its associated standard uncertainty  $u(x)$ .

---

<sup>3</sup>The term “generalised least squares” is usually applied to cases where there are uncertainties and covariances associated with both the measured  $x$ - and  $y$ -values. The case treated by the software, where all covariances are zero, is more usually referred to as “generalised distance regression”.

### 3 Installing and uninstalling the software

#### 3.1 Software user licence agreement

The software is provided with a software user licence agreement (ref: MSC/L/10/003) and the use of the software is subject to the terms laid out in that agreement. By installing the software, the user accepts the terms of the agreement (see step 3 of the installation process in 3.2).

#### 3.2 Installing the software

Prior to installing the software, ensure that any instances of Microsoft Excel are closed.

To install the software, first extract the contents of the ZIP file to a local folder, then:

1. Double-click on **setup.exe** to commence the installation process.
2. Left-click on “Next”.
3. Read the licence agreement, left-click on the radio button corresponding to “I accept the terms in the licence agreement” and left-click on “Next”.
4. Enter user name and organisation and left-click the appropriate radio button to specify if the software is to be made available to all users of the PC or just to the currently logged-on user. Left-click on “Next”.
5. Select the destination folder for the software files. (The default folder is **C:\Program Files\XLGENLINEv1.1** but another folder can be chosen by left-clicking on “Change...” and navigating to that folder and left-clicking on “OK”.) Left-click on “Next”.
6. Installation of the Fortran run-time libraries will commence. Left-click on “Next”. Left-click on “Next”. (A message may be displayed indicating that the read only file **fqwin.hlp** was found when copying files to the destination location. If this message appears, left-click on “Yes” to overwrite the file.) Left-click on “Finish”.
7. Left-click on “Install”.
8. Left-click on “Finish” to complete the installation process.

Note that it is possible to go back and change options chosen in steps 4 and 5 by left-clicking on “Back” and entering new choices.

The installation described above need only be carried out once, before first using the software. The name of the destination folder selected in step 5 above is added to the PATH environment variable during the installation process.

The following software files are placed in the destination folder during the installation process:

- Microsoft Excel workbook **XLGENLINEv1.1.xls**;
- Fortran library file **XLGENLINEv1.dll**;
- **ms11.pdf** (a copy of this document);
- Licence file **NPL\_XLGENLINEv1.1\_MSC.L\_10\_003.pdf**.

The worksheet “Test Sheet” within the workbook **XLGENLINEv1.1.xls** contains example data and can be considered to be a template worksheet. The user may rename this worksheet, make and use a copy (or copies) of the worksheet within the same workbook, and make and use a copy (or copies) of the workbook.

Provided that the name of the destination folder for the software files is not removed from the PATH environment variable and that the file **XLGENLINEv1.dll** is not removed from this folder, the user will be able to use the workbook **XLGENLINEv1.1.xls**, or any copies thereof, placed in this or any other folder.

### 3.3 Uninstalling the software

To uninstall the software:

1. Double-click on **setup.exe**.
2. Left-click on “Next”.
3. Left-click on the radio button corresponding to “Remove” and left-click on “Next”.
4. Left-click on “Remove”.
5. Left-click on “Finish” to exit the wizard.

The uninstallation process deletes the folder where the software files were placed (chosen in step 5 of the installation process) and removes the name of this folder from the PATH environment variable.

The software may also be uninstalled via the Control Panel in Microsoft Windows (e.g., in Windows XP, select “Add or Remove Programs” in the Control Panel, left-click on XLGENLINE then left-click on “Remove” to begin uninstallation).

## 4 Using the software

Upon opening the workbook **XLGENLINEv1.1.xls** (e.g., by double-clicking on it), a message may be displayed indicating that it contains macros and asking if the user wishes to disable or enable macros. The “Enable” option should be chosen. The template worksheet “Test Sheet” should then appear as in figure 1.

**XLGENLINE VERSION 1.1. POLYNOMIAL FITTING SOFTWARE, NOVEMBER 2018**

Full details of how to run the software can be found in the accompanying software documentation.

**Basic user instructions**  
 1. Enter measurement data consisting of x-values, associated u(x)-values (GLS fit only), y-values and associated u(y)-values.  
 2. Enter data for inverse evaluation consisting of y-values and associated u(y)-values.  
 3. Select polynomial degree and type of fit from the dropdown menus.  
 4. Press the EVALUATE button.

**Advanced user instructions**  
 Please refer to software documentation.

MEASUREMENT DATA				DATA FOR INVERSE EVALUATION				DEGREE AND TYPE OF FIT		PROCESS
x-values	Uncertainties associated with x-values	y-values	Uncertainties associated with y-values	y-values	Uncertainties associated with y-values	Polynomial degree:	Type of fit:			
-29.21410751	1.764082504	6.646246519	2.183858181	0.501	0.837	1	OLS	ACCESS ADVANCED GLS PARAMETERS		
-23.543632840	1.177598912	5.685886416	0.136395883	-0.201	0.577			HIDE ADVANCED GLS PARAMETERS		
-42.165975912	1.182405797	2.522175962	0.112932194					CLEAR DATA AND RESULTS		
-11.527694461	0.862303076	0.324507382	1.066768211					CLEAR RESULTS ONLY		
-8.470096034	0.435356218	1.467630514	0.053201461					EVALUATE		
-8.998602581	0.08652210	1.38496320	0.095648405							
-0.03030064	0.637969661	0.957392723	0.932344644							
-7.397088729	0.328163629	0.823303660	0.24410818							
2.063141090	0.127087415	-0.476943019	1.361818956							
2.080495400	0.165048408	-2.278683454	0.714324452							

**RESULTS**

POLYNOMIAL FIT: OLS, DEGREE 1		INVERSE EVALUATION			
		x-values	Uncertainties associated with x-values	y-values	Uncertainties associated with y-values
zmin:	-29.21410751				
zmax:	2.0804954				
Root mean square residual error:	1.257203954	-0.184627529	2.785095194	0.501	0.837
Maximum absolute weighted residual:	2.220177951	-1.852670683	1.936875787	-0.201	0.577
Gradient m:	-0.300962069				
Uncertainty associated with m:	0.005849788				
Intercept with y-axis c:	-1.36103363				
Uncertainty associated with c:	0.107109536				
Covariance associated with m and c:	0.00077091				
Intercept with x-axis x0:	-4.518365329				
Uncertainty associated with x0:	6.257646904				

Figure 1: XLGENLINE template worksheet.

Figure 2 shows a close-up of the five buttons that may be pressed.

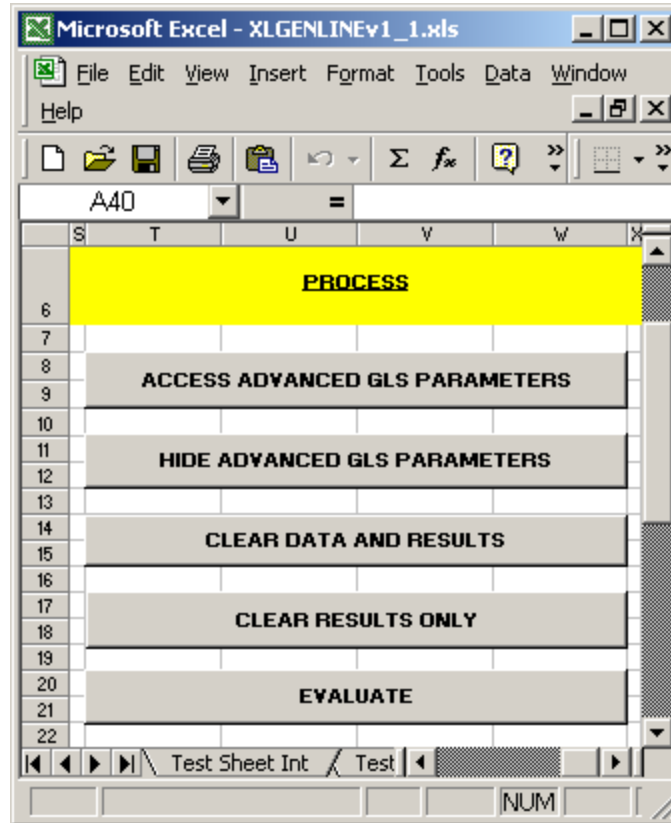


Figure 2: XLGENLINE processing buttons.

Any results on the current worksheet may be cleared by left-clicking on the “CLEAR RESULTS ONLY” button.

Results, measurement data and fitting options on the current worksheet may be cleared by left-clicking on the “CLEAR DATA AND RESULTS” button.

Left-clicking on both the “CLEAR RESULTS ONLY” and “CLEAR DATA AND RESULTS” buttons also has the effect of deleting the two additional worksheets that are present if the data has previously been processed (see 4.2).

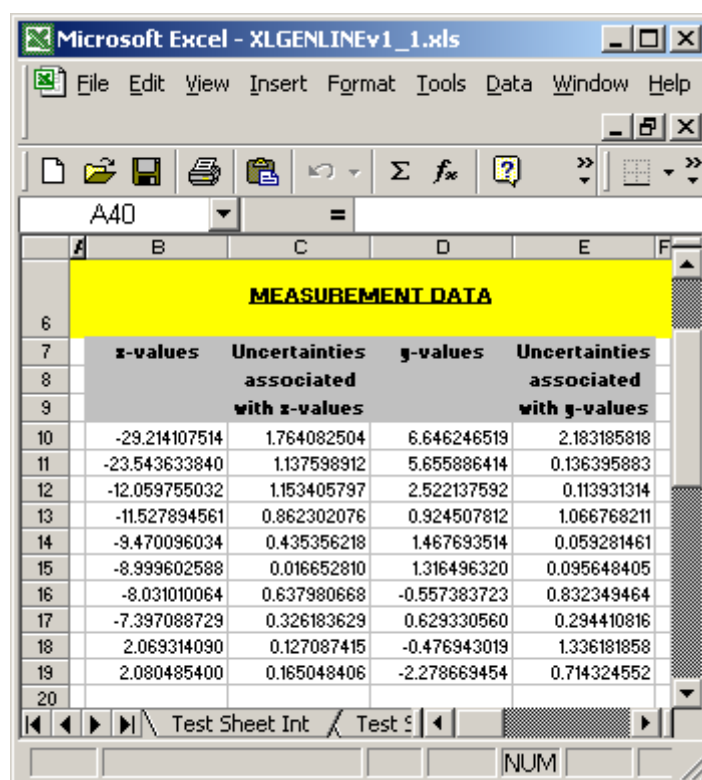
## 4.1 Entering data and selecting fitting options

### 4.1.1 Measurement data

For ordinary least squares (OLS) fitting, measurement data consisting of  $x$ -values,  $y$ -values and standard uncertainties associated with the  $y$ -values should be entered into columns B, D and E, respectively, starting at row 10. Column C may be (but does not have to be) left empty.

For generalised least squares (GLS) fitting, measurement data consisting of  $x$ -values, standard uncertainties associated with the  $x$ -values,  $y$ -values and standard uncertainties associated with the  $y$ -values should be entered into columns B, C, D and E, respectively, starting at row 10.

Figure 3 shows ten measurement data points for GLS fitting.<sup>4</sup>



	B	C	D	E
	<b>x-values</b>	<b>Uncertainties associated with x-values</b>	<b>y-values</b>	<b>Uncertainties associated with y-values</b>
10	-29.214107514	1.764082504	6.646246519	2.183185818
11	-23.543633840	1.137598912	5.655886414	0.136395883
12	-12.059755032	1.153405797	2.522137592	0.113931314
13	-11.527894561	0.862302076	0.924507812	1.066768211
14	-9.470096034	0.435356218	1.467693514	0.059281461
15	-8.999602588	0.016652810	1.316496320	0.095648405
16	-8.031010064	0.637980668	-0.557383723	0.832349464
17	-7.397088729	0.326183629	0.629330560	0.294410816
18	2.069314090	0.127087415	-0.476943019	1.336181858
19	2.080485400	0.165048406	-2.278669454	0.714324552

Figure 3: Example measurement data.

To reduce the risk of the software failing or returning unreliable results (particularly in

<sup>4</sup>While the number of measurement data points that may be processed by the software is limited by the maximum number of rows that may be used in a worksheet in a Microsoft Excel workbook, in practice a lower limit may be imposed by the amount of memory available on the user's PC.

the case of GLS fitting), it is recommended that the measurement data be scaled so that the magnitudes of the  $x$ - and  $y$ -values lie approximately within the interval  $[10^{-2}, 10^2]$ . Subsequent results returned by the software therefore need to be rescaled. See also 4.4.2.

#### 4.1.2 Data for inverse evaluation

Data consisting of  $y$ -values and standard uncertainties associated with the  $y$ -values should be entered into columns I and J, starting at row 10. If no inverse evaluation is to be carried out,<sup>5</sup> these columns should be left blank.

Figure 4 shows two data points for inverse evaluation.

	G	H	I	J	K	L					
6	<b>DATA FOR INVERSE EVALUATION</b>										
7								<b>g-values</b>	<b>Uncertainties</b>		
8									<b>associated</b>		
9									<b>with g-values</b>		
10			0.501	0.837							
11			-0.201	0.577							
12											

Figure 4: Example data for inverse evaluation.

<sup>5</sup>For example, the user may be interested only in the parameters of a straight-line calibration function and have no requirement to implement inverse evaluation.

### 4.1.3 Selecting polynomial degree and type of fit

To select the polynomial degree, first left-click on cell Q7 then left-click on the arrow that appears to the right of the cell. A dropdown menu appears from which the polynomial degree  $d$  (1, 2, 3 or 4) can be chosen by left-clicking on the appropriate number.

To select the fitting type, first left-click on cell Q9 then left-click on the arrow that appears to the right of the cell. A dropdown menu appears from which the fitting type can be chosen by left-clicking on “OLS” (for ordinary least squares fitting) or “GLS” (for generalised least-squares fitting).

Figure 5 shows a close-up of the degree and type of fit choices.

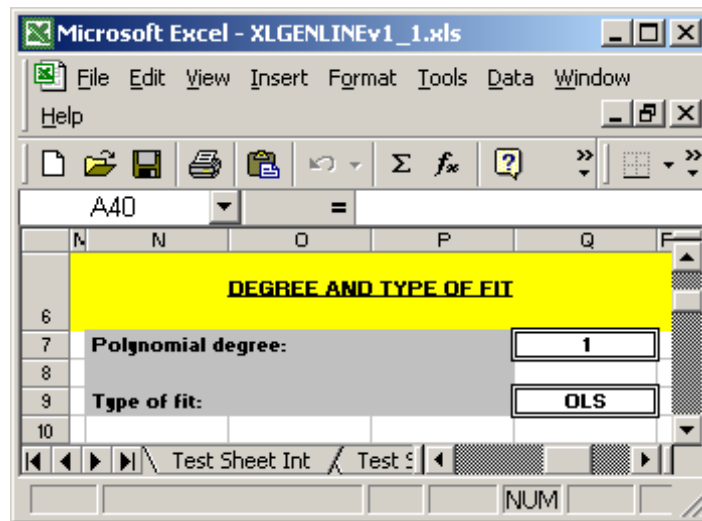


Figure 5: Example fitting parameters.

For GLS fitting, the user may also access additional advanced parameters and enter values for those parameters (see 4.4).



## 4.2 Evaluating

Having entered measurement data, (optionally) data for inverse evaluation and chosen the polynomial degree and type of fit, fitting (and, if required, inverse evaluation) is carried out by left-clicking on the “EVALUATE” button.

If the fitting is carried out successfully, then the following information will be displayed (see figure 6):

- The values taken by  $x_{\min}$  and  $x_{\max}$  (see 4.4.1).
- The root-mean-square (weighted) residual error  $r_0$  given by

$$r_0^2 = \begin{cases} \frac{F_{\text{OLS}}}{m - (d + 1)} & \text{for OLS fitting,} \\ \frac{F_{\text{GLS}}}{m - (d + 1)} & \text{for GLS fitting,} \end{cases}$$

where  $d$  is the polynomial degree and  $F_{\text{OLS}}$  and  $F_{\text{GLS}}$  are the sums of squares of weighted residuals given in expressions (1) and (2), respectively, evaluated at the estimates of the fitted parameters.

- The magnitude  $w_{\text{MAX}}$  of the weighted residual that has greatest magnitude, i.e.,

$$w_{\text{MAX}} = \begin{cases} \max_i \left\{ \left| \frac{y_i - p(\mathbf{A}, x_i)}{u(y_i)} \right| \right\} & \text{for OLS fitting,} \\ \max_i \left\{ \left| \frac{y_i - p(\mathbf{A}, X_i)}{u(y_i)} \right|, \left| \frac{x_i - X_i}{u(x_i)} \right| \right\} & \text{for GLS fitting,} \end{cases}$$

where the weighted residuals (given in expressions (1) and (2) for OLS and GLS fitting, respectively) are evaluated at the estimates of the fitted parameters.

- If the polynomial degree is 1, estimates of the gradient and intercept of the best-fit straight line and their associated standard uncertainties and covariance, and the intercept of the best-fit straight line with the  $x$ -axis and its associated standard uncertainty.

RESULTS							
POLYNOMIAL FIT: OLS, DEGREE 1				INVERSE EVALUATION			
xmin:			-29.21410751	x-values	Uncertainties associated with x-values	y-values	Uncertainties associated with y-values
xmax:			2.0804854				
Root mean square residual error:			1.257209554	-6.184627529	2.789295134	0.501	0.837
Maximum absolute weighted residual:			2.220171551	-3.852107683	1.936875787	-0.201	0.577
Gradient m:			-0.300962069				
Uncertainty associated with m:			0.009848788				
Intercept with y-axis c:			-1.3603383				
Uncertainty associated with c:			0.117015536				
Covariance associated with m and c:			0.001071891				
Intercept with x-axis x0:			-4.519365929				
Uncertainty associated with x0:			0.257040904				

Figure 6: Example results.

If inverse evaluation has been undertaken, the  $y$ -values and standard uncertainties associated with the  $y$ -values entered by the user are displayed along with the corresponding estimates of the  $x$ -values<sup>6</sup> and standard uncertainties associated with the estimates.<sup>7</sup>

<sup>6</sup>For a given  $y$ -value, the number of corresponding  $x$ -value estimates may be between zero and the polynomial degree inclusive.

<sup>7</sup>Uncertainty evaluation is undertaken according to the GUM [3].

In addition, a figure is produced (see figure 7) on a new worksheet, plotting the measurement data with “1 sigma error bars” (i.e., measured value plus and minus one standard uncertainty), the fitted polynomial and (optionally) the inverse evaluation data with “1 sigma error bars” (i.e., value plus and minus one standard uncertainty). The name of the worksheet containing the figure is determined by the name of the worksheet containing the data<sup>8</sup> - if the data worksheet is “Worksheet\_name”, the figure worksheet is “Worksheet\_name Fig”. An additional worksheet named “Worksheet\_name Int” is also generated and is used to store data for the generation of the figure.

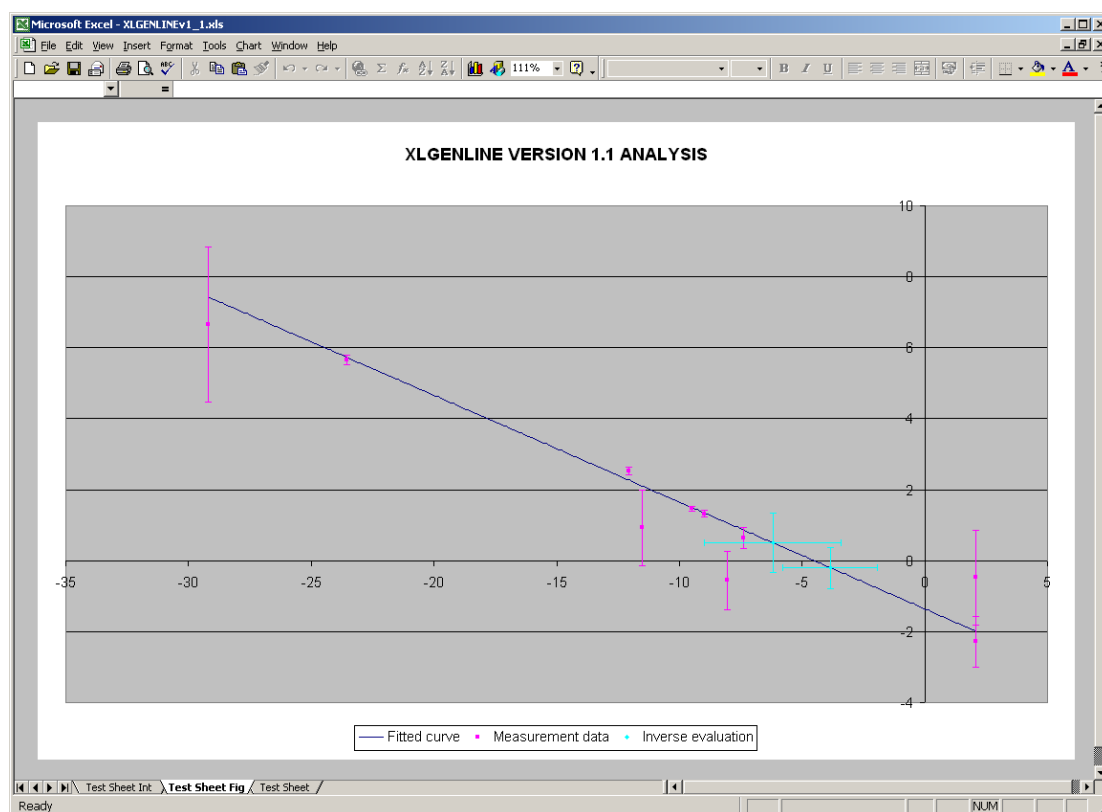


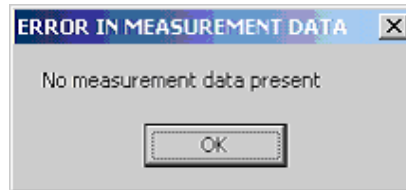
Figure 7: Example figure.

<sup>8</sup>The maximum number of characters in a worksheet name is 31. Since the names of the worksheets containing the figure and the data for the generation of the figure will contain 4 characters more than that of the main worksheet, the name of the main worksheet must consist of 27 or fewer characters.

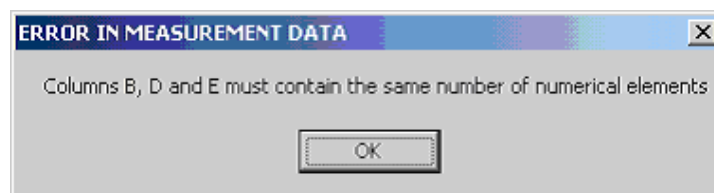
### 4.3 Messages

The following message boxes may appear after pressing the “Evaluate” button:

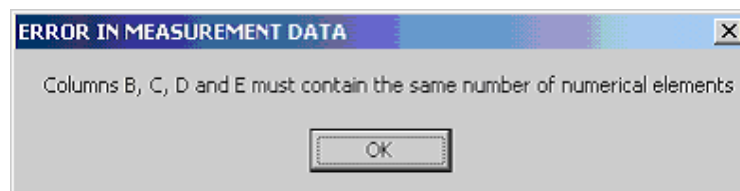
1. There is no measurement data in columns B to E.



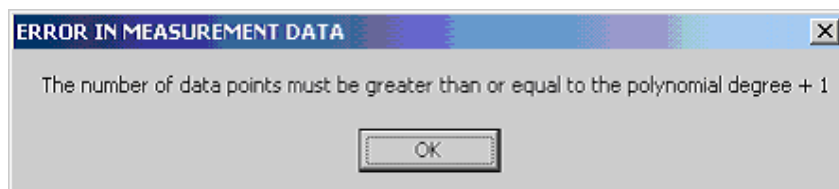
2. In the measurement data for OLS fitting, the numbers of  $x$ -values,  $y$ -values and standard uncertainties associated with the  $y$ -values are not all the same.



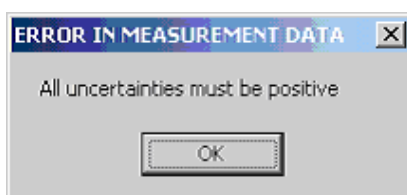
3. In the measurement data for GLS fitting, the numbers of  $x$ -values, standard uncertainties associated with the  $x$ -values,  $y$ -values and standard uncertainties associated with the  $y$ -values are not all the same.



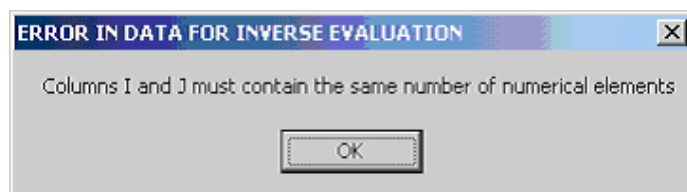
4. There are insufficient measurement data points to fit a polynomial of the chosen degree.



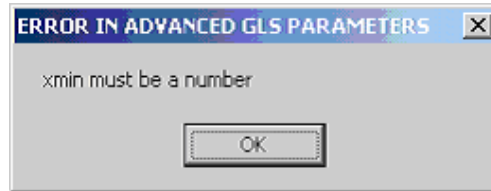
5. At least one standard uncertainty is less than or equal to zero.



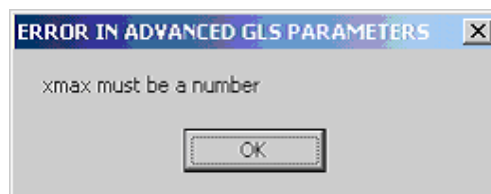
6. In the data for inverse evaluation, the numbers of  $y$ -values and standard uncertainties associated with the  $y$ -values are not the same.



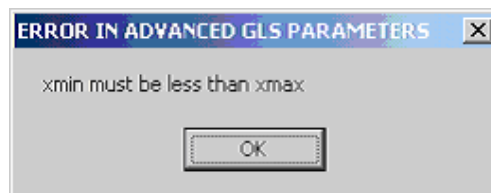
7. A non-numerical value for advanced GLS parameter  $x_{\min}$  has been entered.



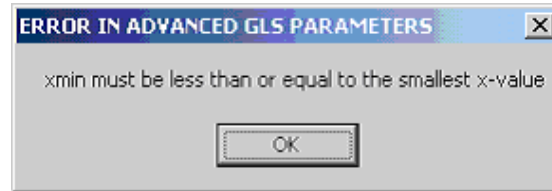
8. A non-numerical value for advanced GLS parameter  $x_{\max}$  has been entered.



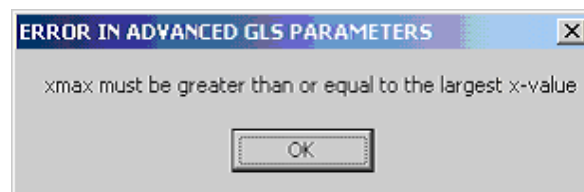
9. Values of  $x_{\min}$  and  $x_{\max}$  have been entered with  $x_{\min} \geq x_{\max}$ .



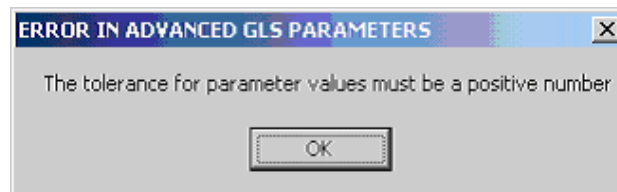
10. A value of  $x_{\min}$  has been entered that is greater than the smallest  $x$ -value.



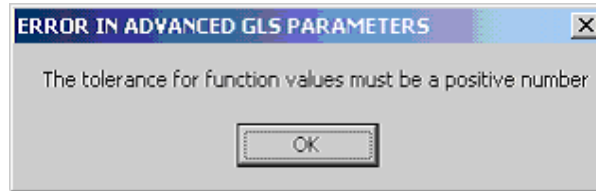
11. A value of  $x_{\max}$  has been entered that is less than the largest  $x$ -value.



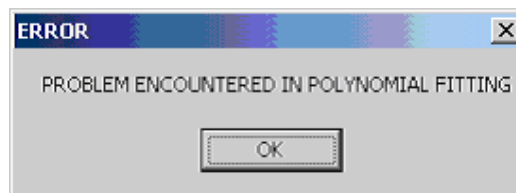
12. A non-positive value of the tolerance  $\rho_p$  on parameter values has been entered.



13. A non-positive value of the tolerance  $\rho_f$  on function values has been entered.



14. The software has encountered a problem when trying to solve the polynomial fitting problem.



The user should be able to take steps to address messages 1 to 13.

Message 14 is more difficult to address. In the first instance, it is recommended that the user check that:

- The measurement data has been entered correctly into the cells.
- The measurement data is appropriate for the choice of polynomial degree.
- The measurement data is not poorly scaled (see 4.4.2).

In the case of GLS fitting, the user may change the values of one or more of the advanced GLS parameters (see 4.4).



#### 4.4 Advanced GLS parameters

The user may access advanced parameters for polynomial fitting as follows:

- Press the “ACCESS ADVANCED GLS PARAMETERS” button to access the advanced parameters for GLS (see figure 8).

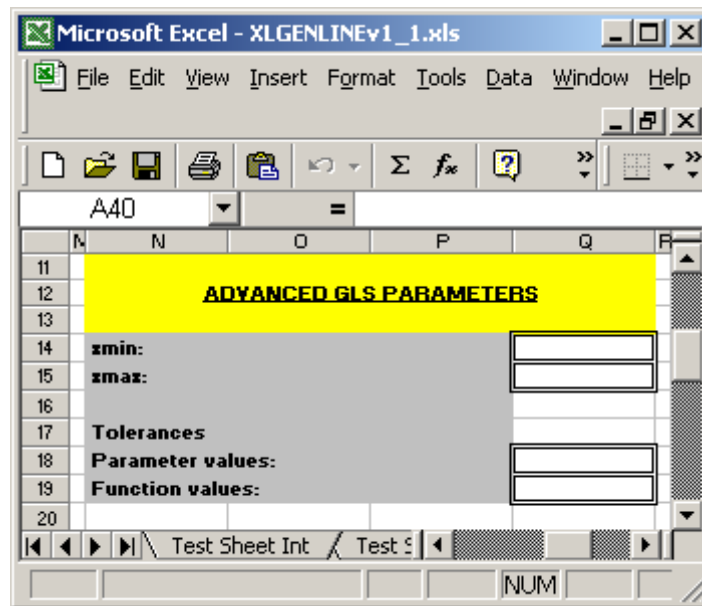


Figure 8: Advanced GLS parameters.

- Enter values for  $x_{\min}$  (cell Q14),  $x_{\max}$  (cell Q15), the tolerance  $\rho_p$  on parameter values (cell Q18) and/or the tolerance  $\rho_f$  on function values (cell Q19) - see 4.4.1 and 4.4.2 for descriptions of these parameters. Leaving any cell (or indeed all four cells) empty means that the default value(s) for the corresponding parameter is used.
- Press the “EVALUATE” button to process the data using the choices of advanced GLS parameters.

The advanced parameters for GLS fitting may be cleared and hidden by pressing the “HIDE ADVANCED GLS PARAMETERS” button.

##### 4.4.1 $x_{\min}$ and $x_{\max}$

For OLS fitting,  $x_{\min}$  and  $x_{\max}$  are set to be the minimum and maximum  $x$ -values, respectively.

For GLS fitting, given initial estimates  $\hat{a}_0$  of the calibration function parameter values and  $\hat{x}_{0,i}$ ,  $i = 1, \dots, m$ , of the  $x$ -values (abscissae), an iterative algorithm (the Gauss-Newton algorithm [5]) is used to find updated estimates. It is possible for one or more of the updated abscissae estimates to be assigned values that are less than the minimum  $x$ -value and/or greater than the maximum  $x$ -value, and values of  $x_{\min}$  and  $x_{\max}$  that allow for this possibility should therefore be assigned.

In the XLGENLINE software,  $x_{\min}$  and  $x_{\max}$  are assigned the default values

$$\begin{aligned} x_{\min} &= \min\{x_i\} - 0.1 \times [\max\{x_i\} - \min\{x_i\}], \\ x_{\max} &= \max\{x_i\} + 0.1 \times [\max\{x_i\} - \min\{x_i\}]. \end{aligned}$$

While for the majority of data sets, the default values of  $x_{\min}$  and  $x_{\max}$  will allow fitting to be undertaken successfully, there may be some data sets for which the iterative process generates abscissae estimates that lie outside the interval  $[x_{\min}, x_{\max}]$ . For such data sets, a message box (see message number 14 in 4.3) will be displayed.<sup>9</sup>

#### 4.4.2 Tolerances on parameter and function values

As described in 4.4.1, for GLS fitting, estimates of the calibration function parameter values and abscissae values are obtained using an iterative scheme. At each iteration, an update step to the current set of estimates is calculated and tests are carried out on both the update step and the differences in the values of the weighted residuals at the current and updated set of estimates to determine if convergence has been achieved.

The tolerance  $\rho_p$  on the parameter values (update step) has a default value of

$$\rho_p = 10^{-6} \times \max\{1, \max\{|y_i|\}\}.$$

If the range of  $y$ -values is much smaller than unity, a smaller tolerance than  $10^{-6}$  may be appropriate. If the data is very noisy, then a larger tolerance may be appropriate.

The tolerance  $\rho_f$  on function values (weighted residuals) has a default value of

$$\rho_f = 10^{-8}.$$

If the uncertainties associated with the  $x$ - and  $y$ -values entered in the spreadsheet are a reasonable reflection of the true uncertainties associated with the  $x$ - and  $y$ -values, then the function values should lie within an interval comparable with  $[-1, 1]$ .

If the uncertainties associated with the  $x$ - and  $y$ -values are overestimated, then a smaller tolerance may be appropriate. Conversely, if they are underestimated, a larger tolerance may be appropriate.

---

<sup>9</sup>Note that problems with the value(s) of  $x_{\min}$  and/or  $x_{\max}$  are not the only reasons this message box may appear.

The default values of the tolerances are set so that the solution determined by the solver is expected to differ from the true mathematical solution by an amount that is much smaller than the uncertainties associated with the fitted parameter values. If the tolerances are too large, the software may return solution parameters that could differ from the true solution by amounts greater than that consistent with the uncertainties associated with the data.

When using the default values of the tolerances, the risk of the software failing is reduced when all the magnitudes of the measured  $y$ -values lie (approximately) within the interval  $[10^{-2}, 10^2]$ . A scaling of all four columns of measurement data may be applied to achieve this.

If the data is multiplied by a factor  $G$  prior to processing, then the effect of the scaling on the results of processing the scaled data may be undone by dividing the  $x$ -values obtained from inverse evaluation and their associated standard uncertainties by  $G$ .

For straight-line fitting, the effect of the scaling may be undone by dividing the estimate of the intercept, its associated standard uncertainty, the covariance associated with the gradient and the intercept, the estimate of the intercept with the  $x$ -axis and its associated standard uncertainty by  $G$  (the estimate of the gradient and its associated standard uncertainty are unaffected by the scaling).

## 5 Additional guidance on calibration functions

Guidance for determining the best-fit calibration function to a set of measurement data having a more general uncertainty structure is available, both for the case of general calibration functions [2] and for straight-line calibration functions [6].

## Acknowledgements

The National Measurement Office of the UK's Department for Business, Innovation and Skills supported the development of this software as part of its Software Support for Metrology (SS/M) programme. The contributions of Frank Guenther and Femi Onakunle are also appreciated.

## References

- [1] Barker R M, Cox M G, Forbes A B and Harris P M 2007 SSfM Best Practice Guide No 4, Discrete Modelling and Experimental Data Analysis *Technical Report DEM-ES 018* (Teddington, UK: National Physical Laboratory)
- [2] Cox M G, Forbes A B, Harris P M and Smith I M 2003 The classification and solution of regression problems for metrology *Technical Report CMSC 24/03* (Teddington, UK: National Physical Laboratory)
- [3] ISO/IEC Guide 98-3:2008 *Uncertainty of measurement - Part 3: Guide to the expression of uncertainty in measurement*
- [4] Fox L and Parker I B 1968 Chebyshev polynomials in numerical analysis Oxford University Press
- [5] Golub G H and Van Loan C F 1983 Matrix computations North Oxford Academic
- [6] ISO/TS 28037:2010(E) Determination and use of straight-line calibration functions