
SQA 2005 Poster (POSTER # 13)

Excel SOLVER for Fitting Bioanalytical Data to Weighted Polynomial Equations

Richard Vanderpool, Ph.D.; RQAP-GLP
Biotechnical Services, Inc.; North Little Rock, AR

ABSTRACT

Bioanalytical laboratory assay results are typically calculated from linear regressions fitting standard peak area ratios in a calibration curve to 1st or 2nd order polynomials. Additionally, many methods such as LC/MS/MS may also incorporate weighting factors into this data analysis process, e.g., Weighted Linear Regression. Laboratory technicians and QC personnel may have access to specialized software (Masslynx, SAS, SigmaPlot) for online or offline analysis and review of assay data. However, this may not be the case for QA personnel. While still specialized, many QA personnel do have access to Microsoft Office Professional. This software package includes Solver, an Add-In to the Excel application, allowing the user to independently determine weighted linear regressions for 1st and 2nd order polynomials used in analytical assays.

This poster presents and discusses templates for determining the coefficients in 1st and 2nd order weighted polynomials ($y = Ax + B$ and $y = Ax^2 + Bx + C$) using Solver. This poster focuses on the 2nd order template to describe the utilization of these templates. The templates are specifically designed for bioanalytical assays based on the FDA Bioanalytical Guidance Document. A specific data set is presented for a $1/x^2$ weighted calibration curve. Additionally, the calculated results for coefficients and back calculated standards are compared to a SAS data analysis. For data with two or more failing standards, the template allows QA to evaluate and comment on the order that standards are dropped from the regression based on an intermediate data analysis, regression with one of two standards rejected. The techniques would also apply to other data analysis curves.

INTRODUCTION

The predominate bioanalytical report reviewed by Biotechnical Services for Quality Assurance in the last 3-5 years has been LC/MS/MS based assays. In these reports, the calibration curves are typically 1st and 2nd order polynomials ($y = Ax + B$ and $y = Ax^2 + Bx + C$ respectively). Additionally, many of the polynomial coefficients are derived through the application of a weighted linear regression with weighting factors of $1/x$ or $1/x^2$. Other weighting factors include 1 (unweighted), $1/x^{0.5}$, $1/y$, $1/y^2$, $1/y^{0.5}$. For the purposes of a data audit, standards, quality control, and unknown samples can be back calculated from peak areas using the reported coefficients and solving the polynomial for

X. However, this procedure does not independently derive the coefficients or allow the auditor to review alternate regression solutions through the exclusion of failing standards from the analysis. The template presented in this poster uses Excel Solver, an Add-In in Microsoft Office Professional, and allows these coefficients to be evaluated by an auditor. The template includes flags to highlight failing standards, QC, and values exceeding the upper or lower calibration curve limits. Currently the flags are set to meet the FDA's bioanalytical guideline [1].

METHODS

Reference data sets were generated for a 1st and 2nd order polynomial. The input data sets are presented in Table 1 for the 2nd order polynomial and Table 2 for the 1st order polynomial. Reference data sets were generated by estimating values from various literature sources [2, 3]. Standard concentrations or peak areas were calculated using the reported coefficients of the published regression model. An Internal Standard (IS) peak area was estimated from a chromatogram or assumed to be 50% of the mid concentration standard. The other IS peak areas were generated taking the initial peak area and adding a random \pm 0-10% of the initial IS peak. A weighted regression was conducted on both data sets using both the appropriate Excel template with Solver and SAS [4]. Results of the 2nd order polynomial SAS/Excel analysis and Table 3 for the 2nd order polynomial and Table 4 for the 1st order polynomial.

Calculations were performed on a Dell Latitude CPi D23ST (233 MHz) and a Dell Latitude D600 (600 MHz) laptops. The PCs utilized MS Windows 2000 (SP-4) and MS XP Professional (SP-2) respectively. Both systems included MS Office Professional with Excel 97 (SR-2) or Excel 2003, both with 15 significant figures and floating point double precision.

Tables 1-2 are data sets used to verify the calculations in the two templates. Tables 2-3 present the comparison of Solver template results with a SAS analysis of the same data. Tables 5-9 present equations and organization used in the 2nd order polynomial template.

DISCUSSION

BSI has used both templates to calculate regression coefficients and back calculate standard, QC and sample concentrations for a variety of clients. The template's calculated results typically match client results to 2-3 significant figures. Note that this is true for very small values. As an example see the %Error for the A coefficient listed in Table 2.

The template includes a weighting factor found under the column "Weighting Factor" and is used under the column labeled " $w * (y_{calc} - y_{obs})^2$ " or weighted residual square. This calculation is a typical residual calculation modified by the weighting factor. The weighted residual follows the example discussed by Miller and Miller [6] and is typically used to weight analytical data.

Solver is an “Add-In” supplied with Microsoft Office Professional [8]. Solver is used to minimize the sum of the weighted residuals (Cell K23) by manipulating the polynomial coefficients. Solver dialog boxes are presented in Figures 1 and 2. For [very] old PCs and old DOS programs, these types of solutions required an initial good estimate (guess) for the coefficient solutions. In nearly all of our applications, an initial estimate of 1.0 has worked.

To date only one client has submitted data in which Solver template results do not match the reported values. In this case the Solver derived coefficients are close to those reported by the client, but have errors of 5-10%. We have found that if the reported coefficients are used as the initial estimate, Solver will solve to the initial estimate. Historically, these errors have been attributed to solutions at an inflection point or a false/local minimum. It is not clear that this is the problem in this instance.

APPLICATIONS

- Template calculates polynomial coefficients and concentration with two to three significant figures.
- Primary application is the independent evaluation of the coefficients in first and second order polynomials and R^2 .
- Secondary application is the ability to independently evaluate the order and effect of excluding standards failing the acceptance criterion, from the calibration curve.
- Standards, quality controls, and sample concentrations can be calculated without the use of solver by entering the reported polynomial coefficients directly into the template. Pass/Fail and <LLOQ/>ULOQ flags are still reported in the template.
- SAS verified test data set (Table 1) used to verify template setup, calculations, and flags.

LIMITATIONS

- Appropriate number of significant figures in input data is essential
- Not a standard data set for testing “goodness of fit”
- R^2 not valid with a “dropped standard/s”
- Rounding issues WILL impact border line passing/failing results. Nearly failing (~15%) standards and QCs in the original data may be calculated by Solver as failing. This may be the result of rounding between Solver data and the original electronic data. Any boundary solution should be evaluated with care.
- We have one client where Solver solutions do not match the reported model parameters (multi analytical runs). We have not determined the cause of this deviation. SAS does reproduce the client parameters as does Solver if the exact solution is used as the initial estimate. Historically, Solver worked for the data until calculations were moved from the instrument to a LIMS system.

TYPICAL TEMPLATE ERRORS

- User data entry errors
- Not editing all appropriate template fields for a new analysis
- Solver default setting is “Max” not “Min”
- Solver retains the parameters from the last solution until 1) Excel is closed, 2) “Reset All” is selected or 3) new parameters are entered.

USING THE TEMPLATE / RUNNING SOLVER

Template Step (Cell Title, Cell Location, or Start of column)

- Enter Study Specific Information (**Tan Fields**):
Analyte Name, Standard Concentrations, Weighting Factor: (H4, E13, I14)
QC and Dilution concentrations, Range or % Error: (E26, E34, R4)
- Enter Run Specific Information (**Blue Fields**):
Batch ID, Calc Date; Reported values R2, A, B, C: (A4, E4, B26, B9, B10, B11)
Analyte Peak Area, Internal Standard Peak Area, Dilution Factor/s: (F13, G13, K3)
- Enter Initial Estimates for A, B, C; Typically 1.0 (**Purple Fields**): (B15, C15, D15)
- Run Excel Solver: (MS Office Professional, Excel Add-in):
In the Excel tool bar select: Tools > Solver...
In the Excel “Solver Parameters” Dialog Box:
Set Target Cell: \$K\$23
Equal To: Min
By Changing Cell: \$B\$15, \$C\$15, \$D\$15
Select: Solve
In the “Solver Results” dialog box select: OK

Interpreting the Template

Template Results (**Yellow Fields**):

- Back calculated standard concentrations, % Errors, Pass/Fail flags: (L13, M13, N15)
- Back calculated QCs, % Errors, Pass/Fail flags: (J26, M26, N26)
- Back calculated dilution QCs, dilution corrected QCs: (J32, L31)
- QCs cont - % error, Pass/Fail flag, out of curve flag (M31, N31, O31)
- Back calculated sample conc., dilution corrected concentration: (J39, L39, 39)

Notes:

- To add a standard(s): Insert blank row(s) between current standards, not after the last standard. Select the previous row and drag down to blank row/s
- To “delete” a standard: Select and delete entire row

-
- To “drop” a standard: Delete the appropriate cell under column “w*(y_{calc}-y_{obs})²”. This retains the “dropped” standards back calculated concentration. However, for this case, the calculated R² is not valid
 - Flags are currently set at the LLOQ and ULOQ so that values at or rounding up to the LLOQ would be passing but flagged as <LLOQ. Values at or rounding down to the ULOQ would be passing, but flagged as <LLOQ.

CRITICAL REGRESSION AND SOLVER EQUATIONS

- 2nd order polynomial used for y_{calc}.
 - $y = Ax^2 + Bx + C$
- Back calculation concentrations used the quadratic equation [5].
 - $$x = \frac{-B \pm \sqrt{B^2 - 4A(C - y)}}{2a}$$
- Weighted regressions are discussed by both Harris (on CD) [5] and Miller [6].
 - Weighting factor is defined as $w = 1/x^2$
- For each concentration, the square of the weighted residual is calculated using y observed and y calculated:
 - $w * (y_{calc} - y_{obs})^2$
- Solver uses the Generalized Reduced Gradient (GRG2) nonlinear optimization algorithm [7].
 - Algorithm varies A, B, C (cells B15, C15, D15)
 - To find a minimum for the sum of the weighted residual (cell k23)

REFERENCES

1. Guidance for Industry. Bioanalytical Methods Validation, United States Food and Drug Administration, Washington, DC, 2001, 25 pages. (See <http://www.fda.gov/cder/guidance/4252fnl.htm>; Last accessed 29 Sept 2004)
2. Z.J. Lin, R. Abbas, L.M. Rusch, and L.J. Shum, *J. Chromatogr. B* 788 (2003) 159-166.
3. A.M. Almeida, M.M. Castel-Branco, A.C. Falcão, *J. Chromatogr. B* 774 (2002) 215-222.
4. SAS Version 9.1, SAS Institute Inc., Cary, NC, 2003.
5. D.C. Harris, Quantitative Chemical Analysis; 5th Edition; Freeman: New York; 1999.
6. J.N. Miller, J.C. Miller. Statistics and Chemometrics for Analytical Chemistry. 4th Edition; Pearson: Essex; 2000.
7. Frontline Systems, Inc.; Incline Village, NV 89450-4288
8. B.E. Joseph. Excel for Chemists. 2nd Edition, Wiley, 2001. (ISBN 0471-39462-9)

REFERENCE DATA AND SAS COMPARISON

Table 1. Reference data used to test Solver template. ¹

2 nd Order Polynomial – Input Data ²		
Standard Concentration	Analyte Peak Area	Internal Standard ³ Peak Area
0.313	1068.14	88434
0.625	1293.20	89301
2.500	2645.22	92769
62.500	45458.38	93636
250.000	173605.75	90168
500.000	331167.17	91035
750.000	473526.62	88434

¹ Values estimated from references 2.

² $Y = \text{Analyte peak area} / \text{Internal Standard peak areas}$.

³ Internal Standard peak area of 86700 with $\pm 10\%$ random error.

Table 2. Reference data used to test Solver template. ¹

1 st Order Polynomial – Input Data		
Standard Concentration	Analyte Peak Area	Internal Standard ² Area
0.10	69194	3259334
0.50	404576	2840704
2.5	2081489	3079921
5.0	4177629	3229432
10.0	8369911	2930411
15.0	12562192	2721096

¹ Values estimated from references 3.

² IS of 2990215 with an added $\pm 10\%$ random error.

Table 3. Comparison of Solver to SAS results for a 2nd Order Polynomial.

	Model Parameters			Back Calculated Concentrations		
	Solver	SAS	% Error ¹	Solver	SAS	% Error ¹
A	-7.017 x 10 ⁻⁷	-6.841 x 10 ⁻⁷	-2.5	0.3130	0.313	-0.006
B	0.007658	0.00766	0.018	0.6268	0.627	0.031
C	0.009681	0.00968	-0.02	2.4596	2.460	0.016
R ²	0.99968	0.9998	0.012	62.485	62.481	-0.0041
				256.145	256.013	-0.052
				496.299	495.728	-0.12
				749.342	747.931	-0.19

¹ Relative to SAS results.

Table 3. Comparison of Solver to SAS results for a 1nd Order Polynomial.

	Model Parameters			Back Calculated Concentrations		
	Solver	SAS	% Error	Solver	SAS	% Error
A	0.29220	0.29210	-0.0010	0.10834	0.1083	-0.037
B	-0.01042	-0.01042	-0.024	0.52325	0.5232	-0.0096
R ²	0.9963	0.9961	-0.017	2.34937	2.3494	0.0013
				4.46437	4.4644	0.00067
				9.81401	9.8140	-0.00010
				15.84067	15.8407	0.00019

¹ Relative to SAS results.

2nd ORDER TEMPLATE DEFINITIONS AND EQUATIONS

Table 5. Primary Template Equations

Cell	Parameter	General Equation	Excel Equation
C7	R ²		=(Q23/(O23*P23)^0.5)^2
H15	Peak Area Ratio	$= \frac{\text{Analyte Peak Area}}{\text{IS Peak Area}}$	=F15/G15
I15	Weighting Factor	$w = \frac{1}{X^2}$	=1/E15^2
J15	y (Calc)	$y = Ax^2 + Bx + C$	=B15*E15^2+C15*E15+D15
K15	Weighted Residual ³	$w * (y_{Calc} - y_{Obs})^2$	=I15*(J15-H15)^2
L15 ¹	BkCalc +x (calc)	$x = \frac{-B \pm \sqrt{B^2 - 4A(C - y)}}{2A}$	=(-C15+(C15^2-4*B15*(D15-H15))^0.5)/(2*B15)
M15	% Error	$\left[\frac{(x_{Calc} - x_{Std})}{x_{Calc}} \right] * 100$	=(L15-E15)/L15*100
N15	P/F ²	Failing if > +20.5 or < -20.5 if > +15.5 or < -15.5 Passing if < +20.499999999... & if > -20.499999999... or if < +15.499999999... & if > - 15.499999999...	=IF(AND(M15>-R15-0.4999999999999999,M15<R15+0.4999999999999999),"-","Fail")
L33	Dilution Correction	Calculated concentration times dilution correction	=K33*J33
O33	Flag ²	< Lower Limit of Quantitation > Upper Limit of Quantitation	=IF(J33<S33,"<LLOQ",IF(J33>T33,">ULOQ","-"))

¹ See reference five.

² Developed for the Bioanalytical Guidance document's acceptance criterion, see reference one.

³ Reference 5 and 6.

Table 6. Data input for a series of related analytical runs (Tan colored fields)

Cell/s	Heading*	Parameter
H4	Analyte	Name of analyte for a series of runs
E15 – E21	Std Conc. x (Std)	Known concentration of standards in calibration curve as defined in the method validation
E27 – E30	QCs	Values for quality control concentrations
E33 – E38	QCs	Values for dilution quality control concentrations
I14	Weighting	Weighting factor equation currently in template
I15 - I21	Weighting	Values for weighting factors
R5 - R10	Specifications	Specifications from the method validation report
R5	Stds Limits – LLOQ	Pass/Fail limit for LLOQ standard
R6	Stds Limits – Other	Pass/Fail limit for all other calibration standards
R7	QC Limits – Standard	Pass/Fail limit for quality control standards
R8	Dilution QC Limit	Pass/Fail limit for dilution quality control standard/s
R9	LLOQ	Method validation report's lower limit of quantitation
R10	ULOQ	Method validation report's upper limit of quantification

* Identifier for column or row.

Table 7. Data input specific to each analytical run (Blue colored fields)

Cell/s	Heading*	Parameter
A4	Batch ID	Identification of batch or analytical run
E4	Calc Date	Date of QA review
B7	R ²	R ² reported in analytical run
B10	Report /A	A coefficient reported in analytical run
C10	Report /B	B coefficient reported in analytical run
D10	Report /C	C coefficient reported in analytical run
F15-F53	Peak Area Analyte	Analyte peak area in each standard/QC/sample in the analytical run
G15-G53	Peak Area IS	Internal Standard peak area in each standard/QC/sample in the analytical run
K33 – K38	Dilution	Dilution factor used with dilution QCs
K41 – K53	Dilution	Dilution factor used with diluted samples

* Identifier for column or row.

Table 8. Solver Parameters (Magenta colored fields)

Cell/s	Heading	Parameter
B15	A	Estimated A Coefficient / Solver's A results
C15	B	Estimated B Coefficient / Solver's B results
D15	C	Estimated C Coefficient / Solver's C results
K25	Sum	=Sum(K15:K21); Weighted Sum of the Square Residuals

* Identifier for column or row.

Table 9. Results of SOLVER weighted regression (Magenta & Yellow colored fields)

Cell/s	Heading	Parameter
C7	R ² / Calculated	Solver calculated R ²
D7	R ² / %Error	% Error reported (B7) and calculated (C7) R ²
B11-D11	% Error / A,B,C	% Error between reported (B-D10) and calculated coefficients (B-D15)
B15-D15	A, B, B	Solver calculated coefficients
L15-L21	BkCalc +x(Calc)	Back calculated standard concentrations (x)
M15-M21	% Error	% Error between back calculated and standards theoretical conc
N15-N21	P/F	Result flagged as fail if % error greater than the standard limit
J27-J30	Calculated	Calculated concentration of quality control standards
M27-M30	% Error	% Error between back calculated and known QC concentration
N27-M30	P/F	Result flagged as fail if % error greater than the QC limit
L33-L38	Corrected	Calculated concentration of dilution quality control standards
M33-M38	% Error	% Error between back calculated and known diluted QC Conc.
N33-N38	P/F	Result flagged as fail if % error greater than dilution QC limit
O33-O38	Flag	Results flagged if calculated conc. < LLOQ or > ULOQ
L41-L53	Corrected	Calculated sample concentration corrected for dilution
O41-O53	Flag	Results flagged if calculated conc. < LLOQ or > ULOQ

* Identifier for column or row.

SOLVER DIALOG BOX

Figure 1. Initial Solver dialog box.

Once values are entered into the template for the calibration curve concentrations and peak areas, the Solver dialog box is obtained through tools menu - Tools > Solver (Figure 1). Initially, all cell areas are blank and the **Equal To** default is **Maximum**. The Target Cell is the purple “sum” cell in column K

Figure 1. Initial Solver dialog box.

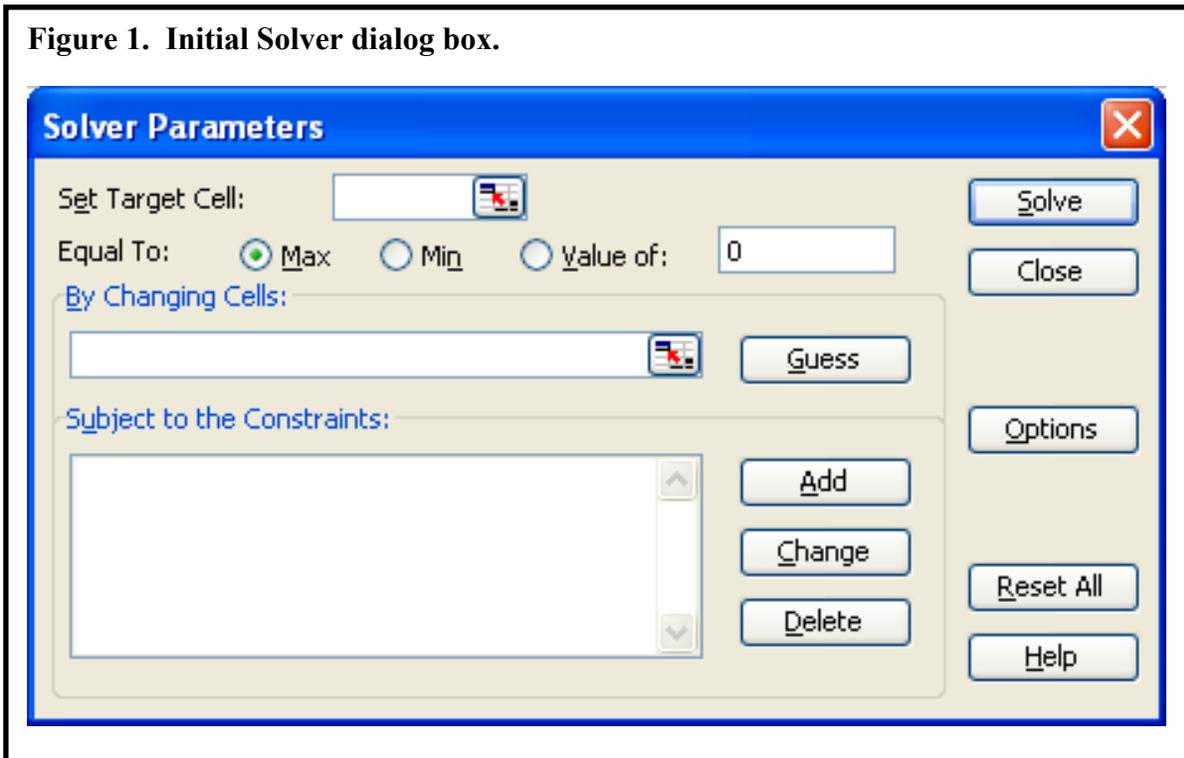


Figure 2. Solver dialog box with parameters for 2nd order polynomial template.

Template cell inputs in Solver to allow program to **Change Cells** values to determine the **Minimum** value for the **Set Target Cell**.

