Computer Methods and Programs in Biomedicine, Volume 107, Issue 2, August 2012, Pages 155-163

Monte Carlo simulation of parameter confidence intervals for non-linear regression analysis of biological data using Microsoft Excel

Ronald J.W. Lambert[1], Ioannis Mytilinaios[1], Luke Maitland[1] and Angus M. Brown[2]

[1]Applied Microbiology Group, Cranfield Health Cranfield University Cranfield MK43 0AL, U.K.
[2]School of Biomedical Sciences, Queens Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK.

Corresponding author:
Ronald J.W. Lambert
[1]Applied Microbiology Group, Cranfield Health,Cranfield University
Cranfield MK43 0A, U.K.

Tel +44-1234 758314; rjwlambert@cranfield.ac.uk

Keywords: Monte Carlo simulation, SOLVER, Excel

Abstract

This study describes a method to obtain parameter confidence intervals from the fitting of non-linear functions to experimental data, using the SOLVER and Analysis ToolPaK Add-In of the Microsoft Excel spreadsheet. Previously we have shown that Excel can fit complex multiple functions to biological data, obtaining values equivalent to those returned by more specialized statistical or mathematical software. However, a disadvantage of using the Excel method was the inability to return confidence intervals for the computed parameters or the correlations between them. Using a simple Monte-Carlo procedure within the Excel spreadsheet (without recourse to programming), SOLVER can provide parameter estimates (up to 200 at a time) for multiple ‘virtual’ data sets, from which the required confidence intervals and correlation coefficients can be obtained. The general utility of the method is exemplified by applying it to the analysis of the growth of *Listeria monocytogenes*, the growth inhibition of *Pseudomonas aeruginosa* by chlorhexidine and the further analysis of the electrophysiological data from the compound action potential of the rodent optic nerve.

1 Introduction

We have previously described the use of the Microsoft Excel spreadsheet to conduct non-linear regression (NLR) analysis of biological data [1]. Direct fitting of the dose response curve, for example, through the use of NLR techniques has been widely advocated, but access to, and comprehension of, commercial software are often at odds with the direct needs of the researcher. We recognised the ease of access and understanding that most researchers have of Microsoft Excel [2], which could allow even those with an elementary understanding of the spreadsheet to conduct relatively sophisticated data analyses, without the expense of purchasing and learning a new statistical or advanced mathematical package.

The SOLVER Add-In package of Excel allows the user to conduct investigations of non-linear (NL) functions using the minimization of the sum of squares of the errors between the observed and modelled values [1]. We further described the use of the technique for the modelling of multiple Gaussian functions, which described the observed electrophysiological data from the compound action potential of the rodent optic nerve [3]. One particular failing of the SOLVER package was the inability to return parameter confidence intervals. It was noted that this requires the use of the Hessian matrix, whose calculation and use would invalidate the aim of making NLR open to anyone [1]. Hence the error analysis of the modelled and observed data was terminated at the calculation of the standard error of the fit.

Confidence intervals can be calculated from knowledge of the Hessian but can also be estimated using either the Bootstrap technique of re-sampling errors between the modelled fit and the observed, or from Monte-Carlo (MC) simulation [6]. The MC technique uses the standard error of the fit of the non-linear model to the observed data to produce sets of ‘virtual’ data. These data are modelled using the same non-linear model and a new group of parameters obtained for each virtual set. From the statistical distribution of these parameters, confidence intervals as well as correlation coefficients can be obtained.

We describe here the use of NLR within the Excel environment and augment our original method with a simple MC analysis, and show its general utility by applying it to analyse the growth of *Listeria monocytogenes*, the growth inhibition of *Pseudomonas aeruginosa* by chlorhexidine and the further analysis of the electrophysiological data from the compound action potential of the rodent optic nerve.

2 Computational Methods and Theory

For a given data set and a particular model (y_{fit}), the sum of squares of the errors is given by

$$SSE = \sum_{i=1}^{n}(y_i - y_{fit})^2$$

In a regression analysis the value of SSE is minimised by changing the parameter values of the model y_{fit}, resulting in the best estimates of these parameters. In linear regression this is solved analytically, but if using non-linear regression this is carried out numerically, based on the input of initial parameter estimates. The square root of the mean of the square of the error (RMSE) is the standard error of the fit. For a given set of conditions the model will return the expected value of y_i , $E(y_i) = \hat{y}_i$. With linear regression, if all the prerequisite conditions are met, then the reported 95% confidence intervals will contain the true value of the regression parameters 95% of the time. With non-linear regression confidence intervals are found using linear approximations and the labelled 95% confidence intervals may not contain the true interval as often.

If the conditions required for regression are met, e.g. constant variance of error (homoscedasticity), normal distribution of errors, then the RMSE is an unbiased estimator of the standard deviation of the fit. A virtual data set can be calculated by adding random error to the expected value of y;

$$Y'_i = \hat{y}_i + N(0,RMSE)$$

This virtual data set can be analysed by NLR to give another set of parameters (the best fit estimates for this virtual data set).

$$SSE = \sum_{i=1}^{n}(Y'_i - y_{fit})^2$$

Another set of virtual data can be generated and the NLR fitting repeated. In the procedure outlined here the sum of squares of the errors from multiple data sets are summed and the fitting of *m*-sets of data are conducted simultaneously

$$SSE_{total} = \sum_{j=1}^{m}\sum_{i=1}^{n}(Y'_i - y_{fit})^2$$

From the *m*-sets of parameters obtained, frequency analyses of the parameter values are performed and the 95% confidence intervals obtained from the normal quantiles; covariance between parameter pairs can be found by calculating the parameters' correlation coefficient.

2.1 FITTING THE MODIFIED GOMPERTZ EQUATION TO MICROBIAL GROWTH DATA

The modified Gompertz equation is a standard empirical model for the fitting of microbial growth data [5].

$$logN(t) = A + Cexp\{-exp(B(M-t)\} \qquad (1)$$

Where A is the asymptotic number (log cfu ml^{-1}) as t tends to negative infinity, A+C = maximum population density as t tends to positive infinity, B is a measure of the slope and M is the time of maximum slope. From these fitted parameters the growth rate and lag are calculated, respectively, as

$$CB/exp(1) \qquad (2)$$

$$M-1/B \qquad (3)$$

Data for the growth of *Listeria monocytogenes* at 30^{o}C in growth media (Tryptone Soya broth) containing 9% salt in terms of log cfu ml^{-1} over a period of 100 hours were obtained.

2.2 FITTING THE LAMBERT-PEARSON MODEL TO MICROBIAL GROWTH INHIBITION DATA

The time taken for a microbial culture to reach a specific optical density (also known as the time to detection, TTD) in the presence of an inhibitor is dependent on the concentration and dose response of that inhibitor. The Lambert-Pearson model (LPM, equation 4) [4] describes the visual growth of a culture as an exponential decay function of the concentration of the applied inhibitor. A plot of the log concentration against the relative rate to detection (RRTD, the ratio of the time to detection of the uninhibited culture, or positive control, to the time to detection of the test culture) gives a characteristic sinusoid, with inflexion at RRTD = 1/exp(1). A linear extrapolation from this point to the log concentration axis allows the estimation of the minimum inhibitory concentration (MIC, equation 5), and a linear extrapolation to the RRTD = 1 axis allows the estimation of the non-inhibitory concentration (NIC, equation 6), the concentration below which normal visual growth is observed even in the presence of the inhibitor.

$$RRTD = \begin{cases} \mathit{if} \quad [x]=0, \quad 1 \\ \mathit{else}\ \mathrm{if} \\ \quad [x]<[P_1] \\ \mathit{then} \\ \quad \exp\left(-\left(\frac{[x]}{P_1}\right)^{P_2}\right) \\ \mathit{else}\ \mathrm{if} \\ \quad \frac{1}{e}(1-P_2(\ln[x]-\ln P_1))<0, \quad 0 \\ \mathit{else} \quad \frac{1}{e}(1-P_2(\ln[x]-\ln P_1)) \end{cases} \tag{4}$$

Where RRTD = relative rate to detection, [x] is the concentration of the given inhibitor, P_1 is the concentration of inhibitor giving a relative inhibition of 1/e, where e is the exponential of 1, and P_2 is a slope parameter which has been defined as the dose response due its similarity with the Hill model.

Two biologically important parameters can be obtained from the LPM; the minimum inhibitory concentration (MIC) and the non-inhibitory concentration (NIC) and are defined, respectively, as

$$MIC = P_1 \exp\left(\frac{1}{P_2}\right) \tag{5}$$

$$NIC = P_1 \exp\left(\frac{1-e}{P_2}\right) \tag{6}$$

Data from the growth inhibition of *Pseudomonas aeruginosa* (ATCC 15442) in the presence of chlorhexidine at 37^{o}C was obtained using standard, published, methods (Lambert and Pearson 2000).

2.3 Further analysis of compound action potential of the rodent Optic Nerve

The compound action potential from the rodent optic nerve typically has three peaks (indicating the presence of three populations of axons with different conduction velocities) with a rapidly decaying transient or artefact from the initial stimulus. Originally this phenomenon was modelled using the sum of four Gaussian functions, one for each feature of the CAP (Brown 2006).

$$CAP = \sum_{i=1}^{4} \frac{A_i}{w_i\sqrt{\pi/2}} Exp\left\{-2\left(\frac{t-c_i}{w_i}\right)^2\right\} \quad (7)$$

where A_i is the area under the curve, w_i the width at half the maximum amplitude and c_i is the latency to the maximum amplitude of peak *i*. This equation models the artefact as a Gaussian; a secondary model which has some desirable features as described in this report, models the artefact as a simple decay, modelling from the initial recording time of 1.04 ms.

$$CAP = D_1 Exp(-D_2(t-1.04)) \sum_{i=1}^{3} \frac{A_i}{w_i\sqrt{\pi/2}} Exp\left\{-2\left(\frac{t-c_i}{w_i}\right)^2\right\} \quad (8)$$

Where D_1 and D_2 are parameters describing the decay of the initial stimulus.

3 Program description and sample runs

The method can be split into 3-stages: Stage 1 fits the given NL model to the observed data using SOLVER. This generates the initial best fit parameters and the RMSE value of the fit. Stage 2 uses the RMSE to generate a set of random numbers based on N(0,RMSE), which are added to the predicted data from Stage 1. The NL model is then applied to multiple virtual data sets simultaneously (using SOLVER) to generate multiple values of best-fit parameters. In stage 3 these values are statistically analysed to provide the mean of the best-fit parameters, their standard errors, 95% confidence intervals and parameter correlations. The data can also be used to provide confidence intervals for parameters calculated from the regressed parameters, which are often dependent on the magnitude of the correlation between those parameters.

3.1 NLR: EXCEL ANALYSIS OF THE GROWTH OF LISTERIA MONOCYTOGENES AT 30^{o}C IN HIGH SALT MEDIA

3.1.1 Monte-Carlo: Excel Analysis

Stage 1. The initial part of the procedure generally follows that given previously [3] except in this case the modified Gompertz model (1) was used. The data used and the initial NLR are shown in Figure 1. The sum of the squares of the errors (SSE, Cell E39) was calculated using the inbuilt “SUMXMY2(data range 1,data range 2)”, where the first data range was the modelled values and the second data range was the observed values. Cell E41 divides this by the degrees of freedom (Cell E40) and takes the square root to give the root mean square error (RMSE). This is the standard error of the curve fit. The

SOLVER Package was used to minimise this value, by changing the values of the four parameters A, C, B and M. Figure 2 shows a plot of the observed data and the modelled function.

Stage 2. *Generation of random numbers with a distribution of N(0,RMSE)*: (It is assumed the user has installed the Analysis ToolPak Add-In). On a separate worksheet, the Random Number Generator was used to generate an array of 50 columns of 33 random numbers using the RMSE of the fitted model as the standard deviation (Figure 3). The number of columns used is set by the maximum number of values that SOLVER can handle (200) divided by the number of parameters in the model. A random seed number of 2 was used for illustrative purposes as the use of this seed number will allow any reader to recreate the exact procedure carried out here.

Generation of virtual data: The random data was added to the modelled log cfu ml^{-1} data to produce a set of 50 virtual observed data. On the spreadsheet these were conveniently located below the random number array.

Stage 3. Fitting multiple models simultaneously using SOLVER: Below each set of virtual data the modified Gompertz model was entered (Figure 4). The regression parameters were placed below this: the initial parameters for each set used the parameters from the initial model fit, although it is advised to check that the parameters do not represent a local rather than the global minimum). The SSE between the virtual data and the modelled data was calculated per data set (Cells C115 to AZ115). The SSE from each data set was summed and this total value placed in cell C117; SOLVER was used to minimise this total SSE by changing all the 200 parameters concurrently. This procedure gave 50 sets of modelled parameters per run. A target SSE of 50 times the SSE obtained from the initial fit (cell B118) was used to monitor the progress of the fitting.

After the minimization procedure, from the fitted parameters the mean and standard deviations were found. The 95% confidence intervals were calculated from the 95% percentiles using the syntax '=PERCENTILE(array, 0.025)', and '=PERCENTILE(array, 0.975). The correlation coefficient was found using the "CORREL(data range 1, data range 2)" function. Tables 1a and 1b give the results of this MC analysis.

3.1.2 Calculation of Biological Parameters

The growth rate, maximum population density and the lag before the onset of growth are important biological parameters and have to be calculated from the parameters obtained from the fitting of the modified Gompertz. However, a singular problem is the calculation of the confidence interval of the calculated parameter. For example the MPD = A + C, but the variance is given by

$$Var(A + C) = Var(A) + Var(C) + 2Cov(AC)$$

Knowledge of the correlation between parameters allows the covariance to be calculated. However in more complex cases such as the calculation of the growth rate (given by BC/e), the calculation of the confidence interval becomes complex. The confidence intervals for these biological parameters, however, can be estimated from the parameter data of the MC analysis. For each parameter set the particular biological parameter was calculated, giving 101 values (including the original fitting parameters).From these values the required percentiles were calculated. In this particular case a re-parameterized version of the modified Gompertz [5] was used to show that the ranges obtained by running a non-linear analysis using JMP were equivalent to those obtained directly from the MC analysis (Table 2).

4 Samples of Program Runs

4.1 INHIBITION OF PSEUDOMONAS AERUGINOSA BY CHLORHEXIDINE

The LPM (Eqn. 4) can be written in Excel as a nested series of IF-statements (e.g. see Figure 6). Initial estimates for the regression parameters can be obtained from an analysis of a plot of the chlorhexidine concentration against the RRTD (Figure 5). Using SOLVER estimates for parameters P1 and P2 and the RMSE were obtained. The fitted parameters and the RMSE were used to prime the MC analysis. The results of the Excel MC analysis are given in Table 3 and Figure 6 shows the spreadsheet used.

A table (81x100) of random numbers based on N(0, RMSE) was produced (cells C4 to CX84). These random numbers were added to the modelled values (cells C88 to CX168), and the non-linear fitting repeated (C172 to CX252) by regressing all 200 parameters at once (Cells C255 to CX256). Figure 6 shows a portion of the calculation performed. Cell C259 sums the SSE for each regression performed (cells C258 to CX258). Cells B262 to CX262 and C263 to CX263 calculate the MIC and NIC values from the regressed parameters respectively.

4.2 FURTHER ANALYSIS OF COMPOUND ACTION POTENTIAL OF THE RODENT OPTIC NERVE

The stimulus-evoked compound action potential from the mouse optic nerve was successfully modelled using multiple Gaussian functions (Brown 2006). The original work used four Gaussian functions to simulate the three peaks of the CAP and the brief stimulus artefact (Eqn.7). This model was set up in Excel and the 16 parameters regressed. The sum of squares obtained was 0.09987; the parameter values for Peaks 1,2 and 3 were essentially identical to those published (nb., a typographical error in the publication gave the area of peak 1 as 7.663, whereas it should have read 0.633). The estimated parameter values for the artefact were, however, different from those published. The peak area found in this analysis was 5.577 vs. 7.180 found previously and a calculated amplitude of 30.8 vs. 38.042.

The standard error of the fit (0.02986) was used to produce an array (96 x124) of normally distributed random numbers. These values were added to the modelled data (on a separate Excel sheet) to produce 96 virtual data sets (Cells DH5 to GY128), Figure 7. (nb columns and rows have been 'hidden' to show the full sheet). The model was added to each cell M5 to DD128, the parameters were placed below each set of modelled data: cells M132 to DD143. The calculated SSE between the modelled data and their respective virtual data set was placed in cells M145 to DD145. The sum of these SSE was calculated in Cell L146. Due to the SOLVER limit of 200 parameters, the MC analysis had to be done in batches of 16. The results of the Monte-Carlo analysis (96 runs; 6 runs of 16) are given in Table 4a. The values obtained for the parameters are very similar to that previously published, apart for the values for the artefact. The confidence interval for the area of the artefact ranged from 4 to 10.7 and all correlations between parameters A4, w4 and c4 had magnitudes greater than 0.994, suggesting that the model was over-parameterized.

A NLR analysis using the JMP statistical package gave an estimate for A4 of 5.5788 (95% CI of 2.595 – 21.835), and correlations greater than 0.994 between A4, w4 and c4.

In a second Excel MC study, the artefact was modelled by a simple exponential decay, replacing the Gaussian for the artefact by the function $D_1Exp(-D_2(t-1.04))$. A similar MC analysis was undertaken; the parameter estimates found for the three principal peaks were relatively unchanged, but the parameter estimates for the artefact now had narrower confidence intervals and the correlation between D1 and D2 = 0.246 (Table 4b)

4.3 Caveats to using MC analysis within Excel

The MC analysis within Excel is initially primed using the parameter estimates from the initial SOLVER minimisation procedure. It is possible that the estimates relate not to a global minimum but to a local minimum, especially if there is a high degree of correlation between given parameters (as was observed in the first analysis of the CAP data). One method to overcome any such possibility is to use different initial parameter estimates for each virtual dataset. This can be done using Excel's "RANDBETWEEN(a,b)" function, where the user generates a scaled value of a parameter between two integer extrema (a and b), and rescales to accommodate the desired magnitude of the initial parameter.

4.4 Comparison of the Excel NLR and MC analysis with *Mathematica*

The non-linear regression capability of *JMP* was used to fit the three sets of sample data; Tables 1, 2, 3, and 4c give the parameters, and their confidence intervals obtained using this sophisticated software; Table 1b also compares the parameter correlations obtained between the Excel MC and JMP NLR analysis. A small program was written to conduct a Monte-Carlo analysis within *Mathematica* (Version 8) with the subsequent analysis of 10,000 virtual data sets (approx 2 to 5 minutes per 10,000 runs) for each of the fitted models. The results of these MC analyses are also given in Tables 1a, 3, and 4d. The NLR analysis using *JMP* essentially gave the same parameter estimates as that from the NLR Excel analysis; the confidence intervals calculated using the Excel MC technique closely agree with those calculated using the Hessian method within *JMP*. A comparison of the results of the MC analysis of *Mathematica* also compares well with the Excel output. It should be noted that the confidence intervals obtained from the MC analysis and the direct NLR analysis using *JMP* do not completely agree with each other and this simply reflects the differences in the techniques used. Each is equally correct.

The comparison between the parameter estimates and their confidence intervals as obtained from the Excel MC analysis with either the direct NLR or the MC analyses using JMP or *Mathematica* demonstrates the capacity of Excel to produce results equivalent to those from more sophisticated packages.

5 Hardware and software specification

The method was carried out on a basic desktop computer with an AMD Phenom 9750 Quad core processor (2.4GHz), using Microsoft Excel 2007 under Windows 7. Non linear regression comparisons were carried out using the JMP (4.0.4) Statistical Software (SAS Institute Inc, Cary NC) and Monte-Carlo comparisons were carried out using *Mathematica* Version 8.0.0.0 (Wolfram Research Inc, Champaign IL, USA).

6 Program availability

Spreadsheets with the worked examples are available from the author on request. The *Mathematica* coding used to produce the NLR fits and Monte-Carlo simulations are also available from the corresponding author.

Acknowledgments: we would like to acknowledge the financial support given by the Food Standards Agency (UK) to Mr Ioannis Mytilinaios (post graduate scholarship PG 1025).

7 References

[1] A.M. Brown, A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. Comp. Meth. Prog. Biomed. 65 (2001) 191-200.

[2] w.w.w. Microsoft.com, Microsoft, Redmond, WA, USA.

[3] A.M. Brown, A non-linear regression analysis program for describing electrophysiological data with multiple functions using Microsoft Excel, Comp. Meth. Prog. Biomed. 83 (2006) 51 – 57.

[4] R.J.W. Lambert, and J. Pearson, Susceptibility testing: accurate and reproducible minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. J. Appl. Microbiol.(2000) **88**, 784-790.

[5] M.H. Zwietering, I. Jongenburger, F.M. Rombouts and K. Van't Riet, Modeling the Bacterial growth curve. Appl. Environ. Microbiol. (1990) **56**, 1875 – 1881.

[6] C. Robert and G. Casella, Monte-Carlo Statistical Methods. (2004) Springer texts in Statistics, Springer (New York, NY USA), 2nd Edition, ISBN 0-387-21239-6.

Tables

Table 1a. Excel MC (100 iterations), JMP (non-linear regression) and *Mathematica* MC (10,000 iterations) analyses for the fitting of the modified Gompertz equation to *Listeria monocytogenes* growth data

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	**A**	**3.948**	**0.0278**	**3.906**	**3.999**
	C	**4.896**	**0.0520**	**4.801**	**4.977**
	B	**0.095**	**0.0033**	**0.090**	**0.101**
	M	**30.318**	**0.2617**	**29.766**	**30.792**
JMP NLR	**A**	**3.951**	**0.0279**	**3.892**	**4.008**
	C	**4.900**	**0.0486**	**4.802**	**5.000**
	B	**0.095**	**0.0033**	**0.089**	**0.102**
	M	**30.341**	**0.2858**	**29.738**	**30.937**
***Mathematica* MC**	**A**	**3.95**	**0.0279**	**3.895**	**4.0048**
	C	**4.9003**	**0.0481**	**4.807**	**4.995**
	B	**0.095**	**0.0033**	**0.089**	**0.102**
	M	**30.340**	**0.2833**	**29.776**	**30.882**

Units: A and C: $\log_{10}$cfu ml^{-1}; B: 1/hr M: hr

Table 1b. Parameter Correlation Table Obtained using Excel MC and JMP NLR analysis (brackets)

	A	C	B	M
A	**1**			
C	**-0.697 (-0.675)**	**1**		
B	**0.347 (0.357)**	**-0.658 (-0.650)**	**1**	
M	**0.546 (0.499)**	**-0.135 (-0.088)**	**0.0122 (0.161)**	**1**

Table 2. Calculation of Biological Parameters from the modified Gompertz using Excel MC, compared to the JMP NLR fitting of the re-parameterised modified Gompertz equation for the growth of *Listeria monocytogenes* at 30^{o}C in high salt (9%).

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	**Growth rate**	**0.172**	**0.005**	**0.162**	**0.182**
	Lag	**19.806**	**0.470**	**18.733**	**20.589**
	MPD	**8.845**	**0.039**	**8.765**	**8.918**
JMP NLR	**Growth rate**	**0.171**	**0.005**	**0.162**	**0.182**
	Lag	**19.829**	**0.499**	**18.765**	**20.878**
	MPD	**8.850**	**0.036**	**8.778**	**8.924**

Units: Growth rate: $\log_{10}$ cfu ml^{-1} hr^{-1}; lag: hrs; MPD $\log_{10}$ cfu ml^{-1}

Table 3. Excel MC (100 iterations), JMP (non-linear regression) and *Mathematica* MC(10,000 iterations) analyses for the fitting of the LPM to the inhibition of *Pseudomonas aeruginosa* in the presence of chlorhexidine.

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	P1	7.265	0.093	7.109	7.455
	P2	1.150	0.022	1.107	1.187
JMP NLR	P1	7.257	0.085	7.092	7.430
	P2	1.149	0.019	1.114	1.185
Mathematica MC	P1	7.257	0.085	7.095	7.427
	P2	1.149	0.019	1.113	1.186
Excel MC	MIC	17.34	0.428	16.61	18.18
	NIC	1.630	0.044	1.550	1.709
Mathematica MC	MIC	17.333	0.372	16.62	18.09
	NIC	1.626	0.037	1.554	1.698

Units; P1, MIC and NIC (mg/l); P2 dimensionless.

Table 4a. Excel MC (96 iterations) analysis of CAP data (Eqn. 7)

Peak	Parameter	Estimate	Stdev	LCL	UCL
Pk1	A1	0.633	0.007	0.620	0.648
	w1	0.243	0.002	0.238	0.247
	c1	1.397	0.001	1.394	1.398
Pk2	A2	1.552	0.017	1.517	1.579
	w2	0.418	0.004	0.410	0.423
	c2	1.875	0.001	1.872	1.877
Pk3	A3	1.288	0.013	1.262	1.313
	w3	0.609	0.006	0.598	0.621
	c3	2.566	0.004	2.559	2.574
Artefact	A4	8.214	2.675	4.002	10.697
	w4	0.152	0.009	0.135	0.162
	c4	0.883	0.021	0.863	0.920

Table 4b: Excel MC analysis (96 iterations) of CAP data (Eqn. 8)

Peak	Parameter	Estimate	Stdev	LCL	UCL
Pk1	A1	0.621	0.008	0.606	0.636
	w1	0.238	0.002	0.234	0.243
	c1	1.397	0.001	1.394	1.399
Pk2	A2	1.564	0.019	1.522	1.595
	w2	0.421	0.005	0.411	0.427
	c2	1.874	0.002	1.871	1.877
Pk3	A3	1.283	0.015	1.254	1.313
	w3	0.607	0.007	0.596	0.621
	c3	2.568	0.004	2.559	2.577
Artefact	D1	4.996	0.034	4.928	5.058
	D2	30.906	0.194	30.447	31.455

Table 4c: JMP NLR analysis of CAP data (Eqn. 8)

Peak	Parameter	Estimate	ApproxStdErr	LCL	UCL
Pk1	A1	0.620	0.009	0.603	0.637
	w1	0.238	0.003	0.233	0.243
	c1	1.396	0.001	1.394	1.399
Pk2	A2	1.566	0.019	1.529	1.603
	w2	0.421	0.004	0.413	0.430
	c2	1.874	0.002	1.871	1.877
Pk3	A3	1.281	0.015	1.252	1.312
	w3	0.607	0.007	0.593	0.621
	c3	2.568	0.004	2.560	2.576
Artefact	D1	4.997	0.034	4.931	5.063
	D2	30.907	0.414	30.124	31.712

Table 4d: *Mathematica* MC (10,000 iterations) analysis of CAP data (Eqn. 8)

Peak	Parameter	Estimate	ApproxStdErr	LCL	UCL
Pk1	A1	0.621	0.008	0.603	0.636
	w1	0.238	0.002	0.233	0.242
	c1	1.397	0.001	1.394	1.399
Pk2	A2	1.564	0.018	1.530	1.601
	w2	0.421	0.004	0.413	0.430
	c2	1.874	0.016	1.871	1.877
Pk3	A3	1.282	0.015	1.253	1.312
	w3	0.607	0.007	0.594	0.621
	c3	2.568	0.004	2.560	2.576
Artefact	D1	4.997	0.034	4.931	5.063
	D2	30.925	0.413	30.141	31.752

Legends to Figures

Figure 1. Non-linear regression analysis of the observed growth data for *Listeria monocytogenes* at 30^oC in 9% salt from an initial cellular density of $7.586x10^3$ cfu ml^{-1}. The growth was monitored over a 90 hour period. The observed numbers as their decimal log were modelled using the standard modified Gompertz equation [5].

Figure 2. Plot of the growth data of *L. monocytogenes* at 30^oC in 9% salt (symbols) with the fitted modified-Gompertz model (line).

Figure 3. A portion of the (50 x 33) random number array generated using the random number feature of Excel's Analysis Addin with a distribution of N(0, 0.09047), using a random seed number of 2.

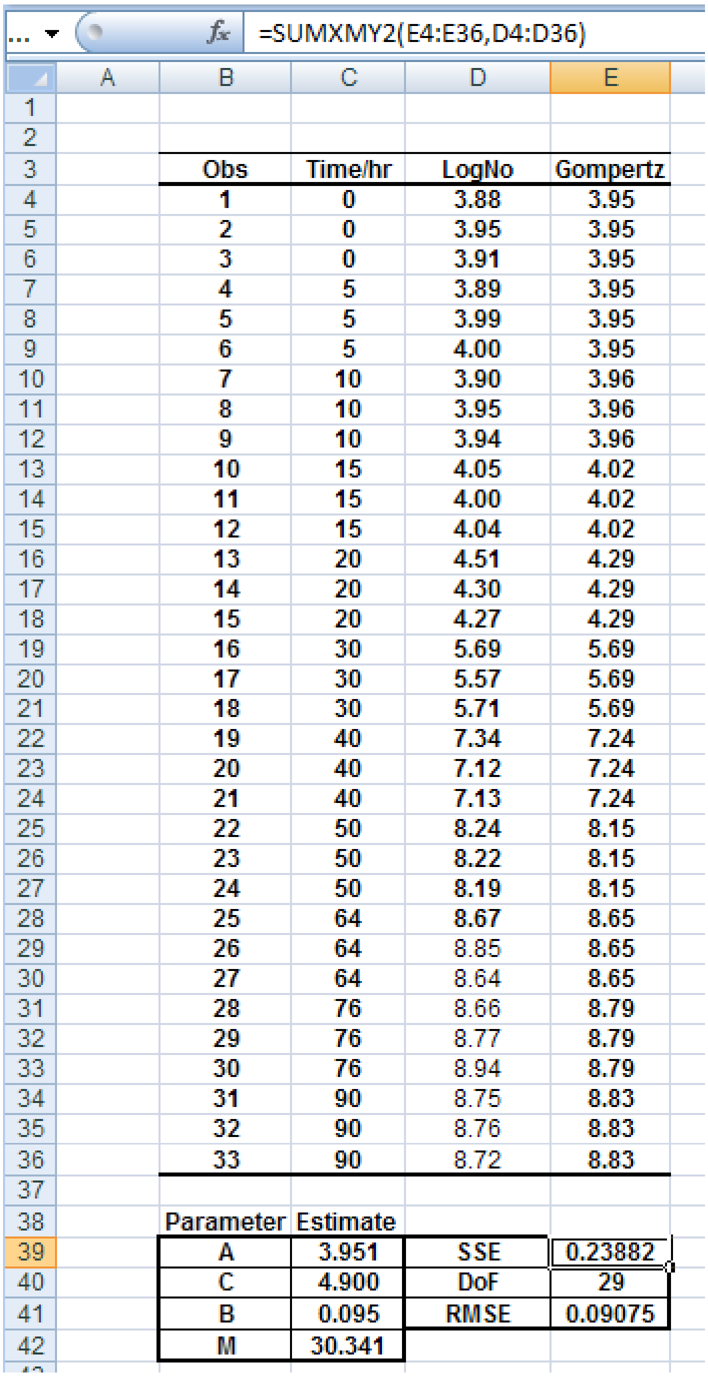
Figure 4. A portion of the (50 x 33) NLR array; column B reproduces the NLR fitting of the original data, with cells B111 to B114 reproducing the regressed parameters. Well C77 shows the syntax used for the formula, which is reproduced over the array. Cells C115 to AZ115 calculate the SSE between the modelled data and the respective virtual data set. Cell B117 sums all the 50 individual SSE values. This value is then minimised using the SOLVER utility.

Figure 5. A plot of chlorhexidine concentration (mg l^{-1}) against the observed relative rate to detection for the growth of Pseudomonas aeruginosa (ATCC 15442) in TSB at 37^oC (symbols) and the fitted NLR model (line).

Figure 6. Spreadsheet used for the Excel MC analysis of the fitting of the LPM to the data for the growth inhibition of *Ps.aeruginosa* by chlorhexidine.

Figure 7. Spreadsheet used for the Excel MC analysis of the fitting of a multiple Gaussian function (Eqn. 7) to CAP data.

Figure 1.



=SUMXMY2(E4:E36,D4:D36)

Obs	Time/hr	LogNo	Gompertz
1	0	3.88	3.95
2	0	3.95	3.95
3	0	3.91	3.95
4	5	3.89	3.95
5	5	3.99	3.95
6	5	4.00	3.95
7	10	3.90	3.96
8	10	3.95	3.96
9	10	3.94	3.96
10	15	4.05	4.02
11	15	4.00	4.02
12	15	4.04	4.02
13	20	4.51	4.29
14	20	4.30	4.29
15	20	4.27	4.29
16	30	5.69	5.69
17	30	5.57	5.69
18	30	5.71	5.69
19	40	7.34	7.24
20	40	7.12	7.24
21	40	7.13	7.24
22	50	8.24	8.15
23	50	8.22	8.15
24	50	8.19	8.15
25	64	8.67	8.65
26	64	8.85	8.65
27	64	8.64	8.65
28	76	8.66	8.79
29	76	8.77	8.79
30	76	8.94	8.79
31	90	8.75	8.83
32	90	8.76	8.83
33	90	8.72	8.83

Parameter	Estimate		
A	3.951	SSE	0.23882
C	4.900	DoF	29
B	0.095	RMSE	0.09075
M	30.341		

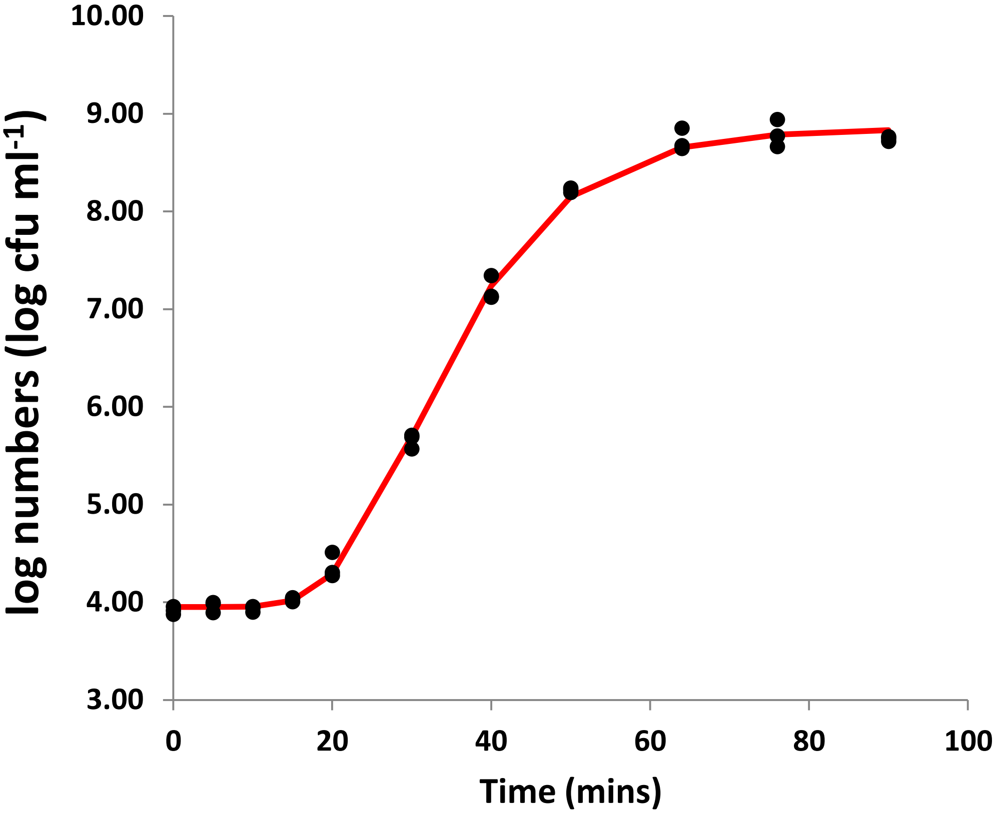

10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
log numbers (log cfu ml-1)
0
20
40
60
80
100
Time (mins)

Figure 2

	A	B	C	D	E	F	G	H
1								
2	RMSE	0.090747						
3								
4								
5		**no**	1	2	3	4	5	6
6		**1**	-0.27177	0.112094	0.057959	0.009818	0.116198	0.022917399
7		**2**	-0.0408	0.073162	0.027441	0.01999	-0.00484	-0.016152197
8		**3**	0.013121	0.005853	-0.04004	-0.12229	0.121177	0.08340568
9		**4**	-0.13812	-0.14519	-0.03006	0.063873	-0.0022	-0.025016443
10		**5**	0.074846	0.033108	0.040228	-0.01283	-0.03762	-0.09728479
11		**6**	-0.11262	-0.0679	0.155139	0.038026	0.108661	0.031230421
12		**7**	-0.03568	-0.14927	0.053359	-0.12548	-0.07223	0.0336726
13		**8**	0.154248	0.00085	-0.0747	-0.14346	-0.09552	0.118928897
14		**9**	0.084781	-0.06085	-0.0096	-0.10539	0.011188	0.04305846
15		**10**	-0.05888	0.084385	-0.10152	0.116863	0.113325	-0.030171924
16		**11**	-0.06138	0.070147	0.207505	0.022917	0.031732	0.042074346
17		**12**	-0.12809	0.049439	0.01095	-0.09689	0.029607	0.245178551
18		**13**	-0.00068	-0.00305	0.086133	-0.05195	0.056966	0.036319875
19		**14**	-0.07968	0.219385	0.176514	0.026766	-0.08762	0.099635975
20		**15**	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Figure 3

=C$111+C$112*EXP(-1*EXP(C$113*(C$114-$A77)))

	A	B	C	D	E	F	G	
76	Time/hr	MOD	1	2	3	4	5	
77	0	3.951	=C$111+C	3.94	3.97	3.92	3.97	
78	0	3.951	3.91	3.94	3.97	3.92	3.97	
79	0	3.951	3.91	3.94	3.97	3.92	3.97	
80	5	3.951	3.91	3.94	3.97	3.92	3.97	
81	5	3.951	3.91	3.94	3.97	3.92	3.97	
82	5	3.951	3.91	3.94	3.97	3.92	3.97	
83	10	3.955	3.91	3.95	3.98	3.92	3.98	
84	10	3.955	3.91	3.95	3.98	3.92	3.98	
85	10	3.955	3.91	3.95	3.98	3.92	3.98	
86	15	4.017	3.98	4.03	4.06	4.00	4.04	
87	15	4.017	3.98	4.03	4.06	4.00	4.04	
88	15	4.017	3.98	4.03	4.06	4.00	4.04	
89	20	4.288	4.29	4.33	4.35	4.30	4.32	
90	20	4.288	4.29	4.33	4.35	4.30	4.32	
91	20	4.288	4.29	4.33	4.35	4.30	4.32	
92	30	5.695	5.77	5.69	5.71	5.69	5.74	
93	30	5.695	5.77	5.69	5.71	5.69	5.74	
94	30	5.695	5.77	5.69	5.71	5.69	5.74	
95	40	7.238	7.32	7.18	7.18	7.21	7.26	
96	40	7.238	7.32	7.18	7.18	7.21	7.26	
97	40	7.238	7.32	7.18	7.18	7.21	7.26	
98	50	8.150	8.20	8.11	8.09	8.13	8.15	
99	50	8.150	8.20	8.11	8.09	8.13	8.15	
100	50	8.150	8.20	8.11	8.09	8.13	8.15	
101	64	8.655	8.67	8.66	8.62	8.67	8.64	
102	64	8.655	8.67	8.66	8.62	8.67	8.64	
103	64	8.655	8.67	8.66	8.62	8.67	8.64	
104	76	8.787	8.79	8.82	8.77	8.81	8.76	
105	76	8.787	8.79	8.82	8.77	8.81	8.76	
106	76	8.787	8.79	8.82	8.77	8.81	8.76	
107	90	8.833	8.83	8.88	8.83	8.87	8.81	
108	90	8.833	8.83	8.88	8.83	8.87	8.81	
109	90	8.833	8.83	8.88	8.83	8.87	8.81	
110								
111	A	3.951	3.907	3.944	3.973	3.917	3.973	
112	C	4.900	4.940	4.959	4.876	4.973	4.852	
113	B	0.095	0.097	0.090	0.090	0.092	0.095	
114	M	30.341	29.726	30.463	30.383	30.321	30.115	
115	sse	0.238816	0.253706	0.254936	0.303308	0.211335	0.115958	0.
116								
117	Tot SSE	11.55302						
118	Targ	11.94079						

Figure 4

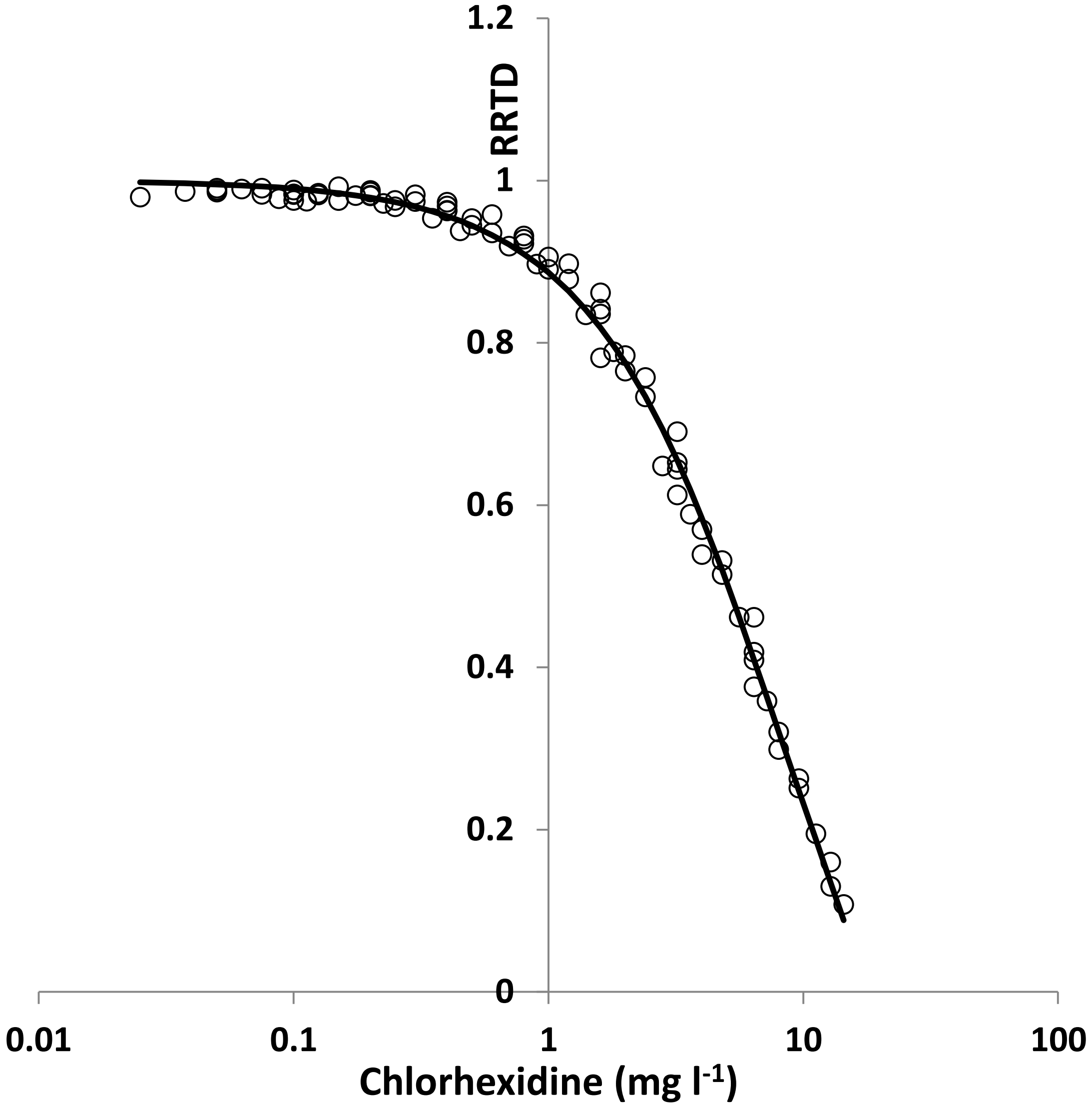


Figure 5

Clipboard | Font | Alignment | Number

=IF($A172=0,1,IF($A172<C$255,EXP(-1*($A172/C$255)^C$256),IF((1/EXP(1))*(1-LN(($A172/C$255)^C$256))<0,0,(1/EXP(1))*(1-LN(($A17

	A	B	C	D	E	F	CW	CX
1								
2	RMSE	0.0218361						
3		obs	1	2	3	4	99	100
4		1	-0.0660	0.0035	-0.0189	0.0191	0.0114	0.0003
5		2	-0.0229	0.0358	-0.0234	0.0286	0.0112	-0.0121
6		3	-0.0035	-0.0160	0.0043	0.0018	0.0172	0.0013
82		79	-0.0021	0.0490	0.0411	0.0336	-0.0062	0.0169
83		80	0.0488	-0.0065	-0.0240	0.0045	-0.0028	0.0234
84		81	-0.0254	-0.0056	-0.0149	-0.0001	0.0126	-0.0125
85								
86								
87	model	Obs	1	2	3	4	99	100
88	1	1	0.93399	1.00350	0.98109	1.01907	1.01144	1.00026
89	0.99764	2	0.97477	1.03347	0.97421	1.02628	1.00886	0.98558
166	0.12804	79	0.12594	0.17699	0.16915	0.16167	0.12181	0.14490
167	0.12804	80	0.17683	0.12153	0.10406	0.13250	0.12528	0.15140
168	0.07826	81	0.05288	0.07265	0.06337	0.07815	0.09086	0.06571
169								
170								
171	Conc	model	1	2	3	4	99	100
172	0	1	=IF($A172=0,1,IF	1	1	1	1	1
173	0.0375	0.99764	0.99797	0.99771	0.99761	0.99721	0.99773	0.99758
174	0.05	0.99672	0.99716	0.99681	0.99668	0.99616	0.99684	0.99663
251	12.8	0.12804	0.12633	0.12472	0.12574	0.13950	0.12756	0.12732
252	14.4	0.07826	0.07540	0.07465	0.07601	0.09119	0.07749	0.07773
253								
254		Mod	1	2	3	4	99	100
255	P1	7.1213	7.3217	7.2238	7.2140	7.3341	7.2728	7.2288
256	P2	1.0796	1.1754	1.1554	1.1478	1.1148	1.1556	1.1444
257								
258		SSE	0.044169068	0.05177825	0.045781245	0.040551295	0.03294538	0.04006947
259		Tot SSE	3.78178					
260		Target	2.18361					
261								
262	MIC	17.9824	17.1433	17.1647	17.2400	17.9861	17.2793	17.3200
263	NIC	1.4498	1.6972	1.6327	1.6145	1.5701	1.6441	1.6107

Figure 6

```
=(M$132/(M$133*SQRT(PI()/2)))*EXP(-2*(($C5-M$134)^2/M$133^2))+(M$135/(M$136*SQRT(PI()/2)))*EXP(-2*(($C5-M$137)^2/M$136^2))+(M$138/(M$139*SQRT(PI()/2)))*EX
```

	J	K	L	M	N	DC	DD	DE	DF	DG	DH	DI	GX	GY
1														
2		**MC analysis**												
3				**Modelled Data**							**Virtual data sets**			
4		**time**	**Modelled**	**MC1**	**MC2**	**MC95**	**MC96**		**time**	**Modelled**	**MC1**	**MC2**	**MC95**	**MC96**
5		1.04	4.9713675	=(M$132/	4.96509	4.94832	4.99872		1.04	4.9713675	4.8811	4.97615	4.94275	4.98347
6		1.06	2.8469856	2.83839	2.83586	2.79671	2.81468		1.06	2.8469856	2.84638	2.80143	2.81095	2.84357
126		3.46	0.0226895	0.02001	0.02097	0.0236	0.02328		3.46	0.0226895	0.0555	0.00527	0.06119	0.01604
127		3.48	0.0186681	0.01635	0.01718	0.01946	0.01919		3.48	0.0186681	0.02206	0.01389	0.01855	-0.03865
128		3.5	0.0152933	0.0133	0.01401	0.01598	0.01575		3.5	0.0152933	-0.04361	0.00629	0.00143	-0.02886
129														
130				**Parameter Estimates**										
131		**Parameter**	**MOD**	**MC1**	**MC2**	**MC95**	**MC96**							
132	**Pk1**	A1	0.6335031	0.63389	0.62065	0.63325	0.6421							
133		w1	0.2428981	0.24216	0.2401	0.24353	0.24118							
134		c1	1.3963352	1.39583	1.39741	1.39594	1.3978							
135	**Pk3**	A2	1.5527126	1.5761	1.57711	1.54235	1.536							
136		w2	0.4182147	0.42184	0.42183	0.418	0.41376							
137		c2	1.8745919	1.87693	1.87453	1.8731	1.8741							
138	**pK2**	A3	1.2868928	1.26079	1.26732	1.30255	1.29881							
139		w3	0.6086458	0.59643	0.60091	0.61345	0.61288							
140		c3	2.5666363	2.57194	2.57021	2.56324	2.56287							
141	**Artefact**	A4	5.5773766	4.58378	5.3543	10.6411	10.641							
142		w4	0.14447	0.14202	0.14304	0.15886	0.158							
143		c4	0.90182	0.91032	0.90432	0.8665	0.86763							
144														
145		**SSE**	0.09986	0.08157	0.07409	0.09013	0.09645							
146		**Tot SSE**	9.759271											

Figure 7