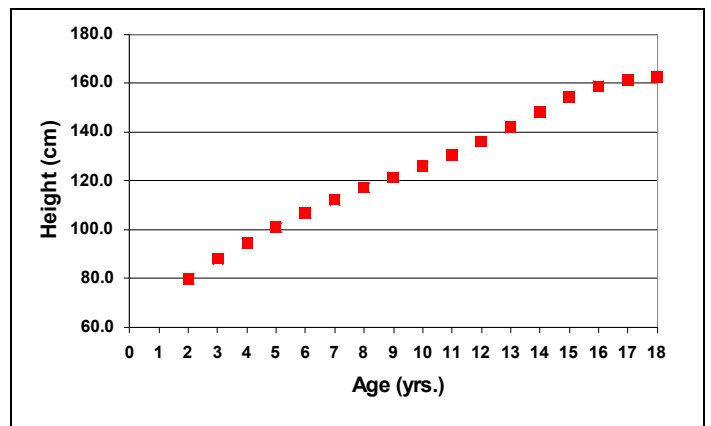


# chapter 2-12

## Mathematical Modeling of Serial Data

### Introduction

Serial data (also known as longitudinal data) are produced when we measure the same subjects repeatedly. The unique feature of serial data is that the data points are not independent of each other. The value of any data point is dependent upon the value of the previous data point in the series. It is this dependency that is made use of in the procedures



**Figure 2-12.1: Serial data – growth in height of a boy**

described here. Examples of the use of serial data in a growth study are illustrated in Figure 2-12.1, or in an intervention study where there is pre- and post-intervention testing often with testing at more than two time points. Another form of serial data are produced when electronic signals are sampled at a fixed frequency and saved as digital information in the computer (A/D - Analog to Digital Data Conversion). Data collected in this way have the same characteristics as the growth data described above. They are discrete data points at each point in time that are interdependent.

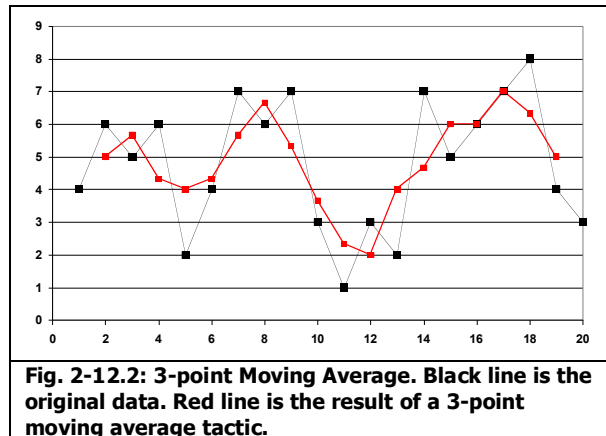
Earlier we discussed the use of the paired t test for simple pre-post (2 measurement occasions) evaluation of the difference between means. With more than two occasions the repeated measures ANOVA can be employed. Serial data opens itself up to more than just testing differences between means, however. Whether, the data are collected manually or by A/D, a common feature is that of noise. In A/D, that noise is a signal superimposed upon the real signal as a consequence of poor setup or even a feature of the acquisition hardware. In manually collected data it is a consequence of measurement error on each occasion. Either way you have sequential data that is less smooth than the real data for the underlying phenomenon. The next section addresses methods that can be used to smooth serial data.

There are various tactics available to smooth serial data. We will deal with:

- Moving “Running” Averages
- Weighted Moving Averages
- Fitting Mathematical Equations
- Mathematical Modeling

## Moving Averages

One method used in digital conditioning is to carry out moving (or running) averaging to smooth data. The challenge is to smooth the data without distorting the biological meaning of the data. The concept of a moving average acknowledges that serial data points are not independent of each other and in fact makes the points more dependent on each other because of the averaging process.



**Fig. 2-12.2: 3-point Moving Average. Black line is the original data. Red line is the result of a 3-point moving average tactic.**

In carrying out moving averages you simply average adjacent data points to give an estimate to the central point of that group of points. You may choose 3-point, 10-point or indeed any number of points. The number of points however is usually an odd number so that there is an easily identifiable central point. Fig. 2-12.2 shows the result of 3-point moving averaging of data. The black line represents the original data. The red line represents the result of the averaging. The moving average in addition to smoothing the data, results in a reduction in the amplitude. The intention however, is that the averaging minimizes noise whilst maintaining the true phenomenon of the data. This will not be the case if the amplitude is reduced substantially. This problem of moving averages is shown in the four plots in Figure 2-12.3.

A simple sinusoid is our initial data, shown in the first plot. To this data was added a random error term to produce the second plot. This now represents a signal we have collected from some device. It is a sinusoid but it is contaminated with some noise, the effect of which we will try to overcome with a moving average smoothing technique. The third and fourth plots show the results of 5-point and 7-point moving averages. Notice that the amount of noise has been reduced and that the waveform still looks similar to the original sinusoid. The last plot resulted from a 21-point average, which resulted in a smoother curve but unfortunately at the sacrifice of the original sinusoidal information. Although the result is a sinusoid with the same frequency, the resultant curve has about half the amplitude of the original curve; thus, the smoothing has altered the original data as well as smoothing out the noise. This exemplifies the problem of

moving averages, in that it smoothes out the peaks or valleys in your data, which might be noise but it also might be your valuable signal.

A weighted average is an alternative, in that points closer to the middle are weighted more.

e.g. 5 data points 10, 9, 13, 12, 1  
average = 12

with a weighting scheme of 1 3 5 3 1

weighted average = 11.8

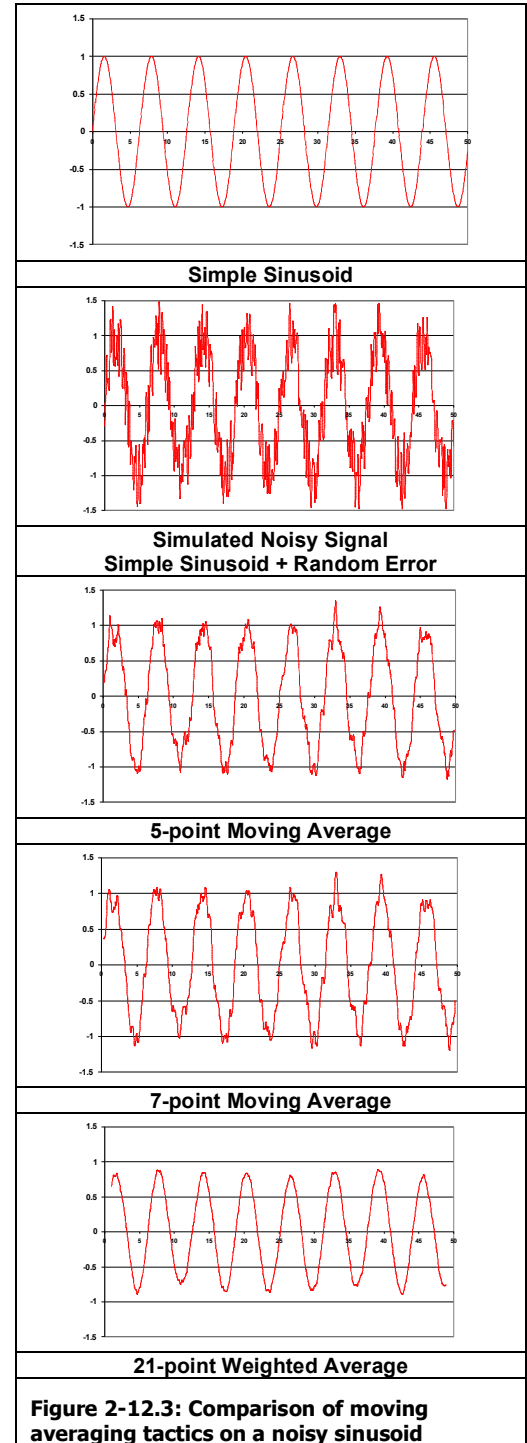
$$[(1 \times 10) + (3 \times 9) + (5 \times 13) + (3 \times 12) + (1 \times 16)] / 13$$

This will tend not to flatten the peaks and valleys quite so much, since more importance is given to the central points. In the end the decision as to what is the appropriate weighting scheme is based on which scheme will smooth the data but not distort the meaningful part of the signal.

### Fitting Mathematical Equations

A method often used to smooth noisy data is to fit an equation to the data. For instance, a power curve might look like the shape of the data, so that is used to smooth through the data. This, however, imposes quite a restrictive shape to the data, and this might not actually be the best fit to the underlying phenomenon.

A 5<sup>th</sup> order polynomial is often used because it can fit the most complex of curves (see Figure 2-12.4), but obviously has the restriction that it has no biological relevance. As will be discussed later, the better approach is to use a model (an equation with a theoretical basis) to smooth the data. In this way a smoothing occurs but the shape of the true phenomenon is preserved. The major assumption here, however, is that your model truly reflects the underlying phenomenon.



### Estimating Missing Values (Imputation)

Another reason to fit a curve to your data can be to estimate missing values in serial data. This process is called imputation. In this process, a curve is fit to your data, and then a prediction can be made based upon any value of the independent variable. There is a high degree of confidence associated with such predictions, since they rely heavily on information from adjacent points. This confidence is enhanced if the curve is fit by a mathematical model that truly represents the shape of the curve of the true phenomenon.

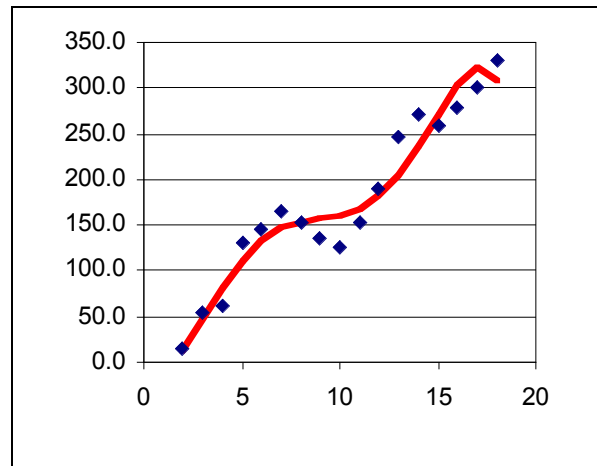


Figure 2-12.4: 5<sup>th</sup> order Polynomial Curve Fit

This confidence is enhanced if the curve is fit by a mathematical model that truly represents the shape of the curve of the true phenomenon.

### Mathematical Modeling

A real biological phenomenon can be modeled with one or more mathematical equations. There are many reasons why one might embark on modeling, with the most obvious being a desire to show understanding of a phenomenon in an attempt to explain the course of events. Modeling is common in kinesiology and you will no doubt encounter models for systems such as muscle function and temperature regulation etc. A good model can be used as evidence for underlying mechanisms, however, a poor model that “works” under certain conditions can actually serve to mislead as much as it illuminates.

One of the main reasons to use mathematical modeling is to **explain** a phenomenon. However, models are also often used to predict events, such as local weather forecasts. Models are often less perfect in their predictions than we would like. The key to this type of modeling is how precise does the prediction need to be? This question will be discussed later.

In this chapter we will present two models. First, we will present a model describing the growth of height of an individual, which can be used to smooth real data for the production of growth norms. Later, we will show a mechanical model of dynamic skinfold compression where the aim was to try to explain, in mechanical terms, the shape of the dynamic compression of a skinfold over the first 2 seconds of compression. This served to give understanding to the elastic and viscous nature of the skinfold under compression.

### Deterministic and Probabilistic Models

There are two basic forms of mathematical models, deterministic and probabilistic. Deterministic models make precise predictions. An example of a deterministic model would be Newton's law of motion relating the force of a moving body to its mass and acceleration:

$$\text{Force} = \text{Mass} \times \text{Acceleration}$$

A precise prediction does not mean an exact prediction but a prediction with a very small error that can be regarded as negligible. When the error is not small enough to be regarded as negligible we have a probabilistic model.

In house design, models can be developed to predict what size of wooden beam should be used to support the weight of any given roof. However, there would be considerable error in this prediction because of the variable nature of wood itself. It would therefore not be a safe procedure to predict the minimum size of wood to use. Rather, a safety allowance is added to the predicted value in order to accommodate the variability of wood and also additional variable unknown loads on the roof such as heavy snow fall.

As in linear regression, we are relating a response,  $y$  to a variable  $x$  based upon the results of  $n$  experiments (the number of subjects). In fitting a straight line through the observations we have chosen a model of the form,

$$y = \alpha x + \beta$$

where  $\alpha = \text{slope}$  and  $\beta = \text{intercept}$

The model,  $y = \alpha x + \beta$ , is a deterministic mathematical model for the relationship between  $y$  and  $x$ . Given a value of  $x$ , the model predicts a single value of  $y$  which is unique and thus determined. A requirement of our deterministic models is that they predict with an error that is negligible for practical purposes. A model can still be useful even when there is large error, as long as the size of the expected error in prediction is known and is taken into in application of the results. In making predictions we must take into account the size of the Standard Error of Estimate (S.E.E.). By way of example, regression analysis was carried out on young university men and women to produce equations to predict standing height from tibial height (perpendicular distance from the floor to the superior border of the lateral surface of the tibia). Tibial height is used as an estimate of lower leg length.

The result of modeling this relationship is the production of the following two regression equations.

$$\text{Men:} \quad \text{Height} = (2.34 \times \text{Tibial Height}) + 67.72$$

$$r = 0.88 \quad \text{S.E.E.} = 3.04 \text{ cm}$$

$$\begin{aligned} \text{Women:} \quad \text{Height} &= (2.37 \times \text{Tibial Height}) + 62.26 \\ r &= 0.88 \quad \text{S.E.E.} = 3.74 \text{ cm} \end{aligned}$$

In men, height can be predicted from tibial height with an S.E.E. of 3.04 cm and in women the S.E.E. equals 3.74 cm. Our first question is do we have a good model? The answer is based on our judgment of whether the S.E.E. is small enough to be an acceptable margin of error when predicting height? When Sitting Height is added via stepwise regression to produce the multiple regression equations below, the standard error of estimate is reduced to 2.33 cm for men and 2.03 cm for women.

$$\begin{aligned} \text{Men:} \quad \text{Height} &= (1.68 \times \text{Tibial Height}) + (0.90 \times \text{Sitting Height}) + 14.45 \\ r &= 0.93 \quad \text{S.E.E.} = 2.33 \text{ cm} \\ \text{Women:} \quad \text{Height} &= (2.34 \times \text{Tibial Height}) + (2.34 \times \text{Sitting Height}) + 5.92 \\ r &= 0.97 \quad \text{S.E.E.} = 2.03 \text{ cm} \end{aligned}$$

When predicting with such an equation one has to bear in mind this margin of error, and decide, based upon practical considerations, whether this is precise enough for the task at hand.

We have actually now moved to a probabilistic model for  $y$ . In a probabilistic model there are one or more random components that are present to try to account for the random variability of  $y$  for a given value of  $x$ . Our straight line probabilistic model could be represented as:

$$\text{Height} = m_t \text{Tibial height} + c + \epsilon$$

Where:  $m_t$  is the regression coefficient for tibial height,  $c$  is a constant (the intercept of Height when Tibial Height = 0) and  $\epsilon$  is a random variable representing the variability of  $Y$  about the straight line relationship

This description of the error variable ( $\epsilon$ ) is generally in the form of assumptions about the nature of the random variable  $\epsilon$ . A simple form is to state that  $\epsilon$  has some specific probability distribution with expected value and variance, and that repeated values of  $\epsilon$  are independent of one another in repeated sampling. The assumed nature of error about the straight line is:

- For each value of  $X$  there is a normal distribution of  $Y$  from which the sample value of  $Y$  is drawn
- The population of values of  $Y$  corresponding to a selected  $X$  has a mean that lies on the straight line, such that the mean error = 0
- In each population the standard deviation of  $Y$  about its mean has the same value (SEE)

Thus, the error term is a random value from a normal distribution with mean = 0, and standard deviation = standard error of estimate of the regression line.

## Ideal Characteristics of a Model

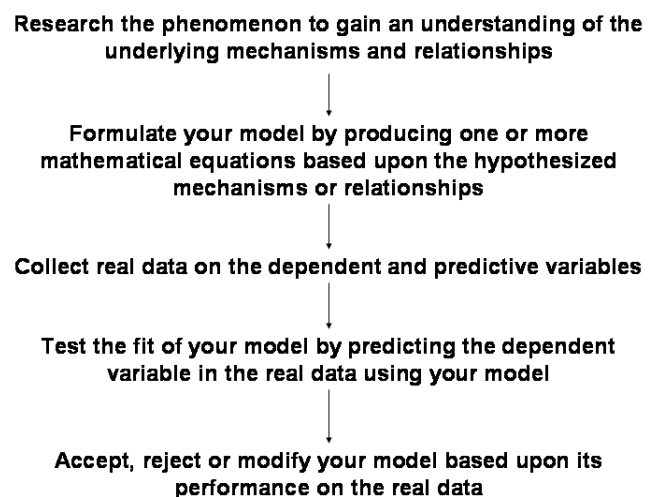
A mathematical model is considered accurate if it is able to predict some variable and do so with an error that will be negligible for practical purposes. Many mathematical models of phenomena in physics predict with errors that are often extremely small, sometimes so small as to be unobservable by the measuring instrument. Many other models employed in engineering, the physical sciences, business, and the social and the biological sciences predict with a much larger error. Whatever purpose you have for modeling, there are basic ideal qualities that you strive for in your model.

1. **Simple form;**
2. **Good fit to experimental data;**
3. **Biologically meaningful parameters.**

It does not matter what system you are working with, the basic steps in the modeling process are the same. These steps are outlined in Figure 2-12.5. The researcher generally goes into this process with a firm foundation of the underlying physiology or mechanisms at play and with evidence or intuition for promoting a given model. The hardest and most important part of all however, is this initial determination of the basis for the model. What are the underlying mechanisms or relationships that will determine the structure of your model, and its ability to predict the dependent variable? The success of the model is determined by the depth of your knowledge of the answers to this question. Once the system is understood then the task is then to translate this into mathematical expressions.

In essence, a model is nothing more than one or more mathematical equations describing the phenomenon. What distinguishes it from simple fitting of mathematical equations is that the parameters in the model have some biological meaning, and therefore explanation. Once any model has been developed it must be tested on real data. Select a sample with characteristics appropriate for your model. You

may need to restrict subjects by age, sex, or some other characteristics that are important to

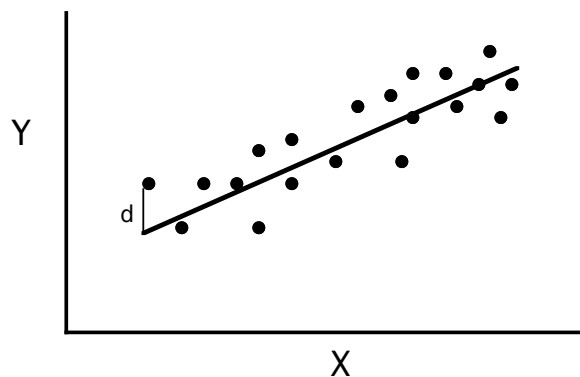


**Figure 2-12.5: Steps in the Modeling Process**

control for in application of your model, if you are unable to include them as predictive variables in your model. Once you have decided upon the sample, collect data on the dependent variable that you are predicting and the predictive independent variables. You can now embark on the most exciting phase of modeling, and that is, comparing real data with predictions from your model. The tough task is working out what those modifications should be. If your model fits reasonably, then you probably have a handle on the underlying mechanisms and have made a contribution. After further modifications are made, your new model must be tested against more real data, and so on until you settle upon the model that serves your needs.

### Using SOLVER for Least Sum of Squares Curve Fitting

Earlier, SOLVER was briefly introduced as an EXCEL function that allowed a value for a target cell to be found based upon changing values in cells that are linked to the target cell via formulae. This is one of the so-called "What If" functions in EXCEL. If you only had EXCEL available to you, you could still do some very complex modeling using the SOLVER function.



**Figure 2-12.6: Least Sum of Squares Curve Fitting**

In EXCEL under the TOOLS menu there is the DATA ANALYSIS package that contains many statistical tests. One of these tests is the linear regression analysis. If you need linear regression analysis this is the best way to do it in EXCEL. However, in order to illustrate the use of SOLVER in EXCEL, we will start our modeling by considering a linear regression analysis without using the Data Analysis Package. Earlier the process of least sum of squares curve fitting was explained using Figure 2-12.6, where the sum of squared deviations ( $d$ ) is used as the criterion of best fit. This can be reproduced in EXCEL using SOLVER.

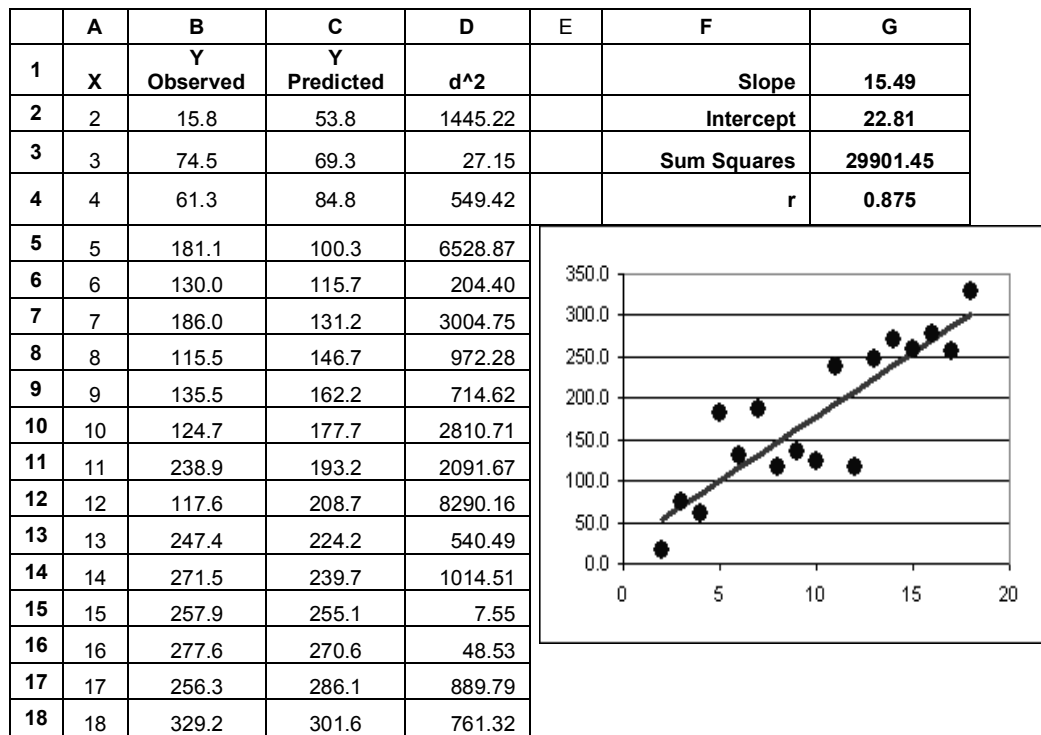
In our example below (Figure 2-12.7), we have measures of some variable  $X$  and associated  $Y$  values for a group of 9 subjects. These are not real data, and a high correlation coefficient is not expected. However, our aim is to find out the equation of the best fitting straight line to predict  $Y$  from  $X$ , and in turn find out how good this equation is at prediction.

The process starts by entering the formula for a straight line,  $Y = mX + c$ , where  $m$  = slope &  $c$  = intercept, into the work sheet. A column of predicted  $Y$  values ( $Y_p$ ) is calculated by entering the formula  $= (A2 * \$G\$1) + \$G\$2$ , into cell C2 and filling down the column. G1 and G2 are the cells containing the values assigned to the slope and intercept respectively. The \$ signs



indicate the absolute referencing required for each of these cells. The values in these cells are set at some initial estimated values, however, these values will change to appropriate values automatically when we run SOLVER. Next, calculate the squared differences between Y predicted and Y observed (the residuals or *d* values), by entering into cell D2 the formula  $= (B2 - C2)^2$ . Fill down the column.

The best fitting line is the one where the sum of squared deviations is the smallest. Therefore, we now calculate the sum of the squared deviations by entering the formula  $= \text{SUM}(D2:D18)$  into cell G3. At this point, if you change the value of either the slope or the intercept, the predicted values of Y will change, and hence the sum of squares. You could laboriously go through various combinations of values for slope and intercept, until you found the values that gave you the minimal value for the sum of squares, and hence was the best fitting straight line through the data. Fortunately, the EXCEL option called SOLVER will do this for you.



**Figure 2-12.7: EXCEL Spreadsheet of SOLVER set up for fitting a straight line**

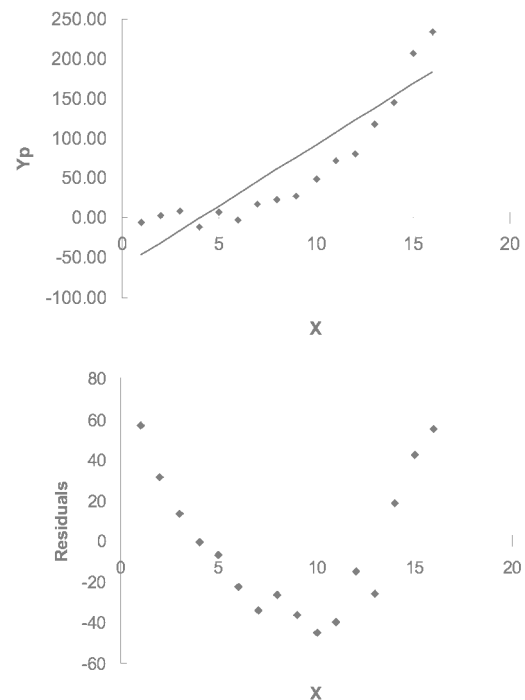
Look under the TOOLS menu and you will see SOLVER listed. Select it, and after a slight delay for loading, the SOLVER dialogue box will appear. You have to select a target cell, which in our case is the sum of squares cell (G3). Then you have to indicate the target value. In our case we select minimum. Next, SOLVER needs to know which cell or cells it is allowed to change in order to minimize G3. The cells to be used are those of the slope and intercept

(G1:G2). Click the Solve button and SOLVER will go through a series of iterations and come up with a solution that fits your criteria. SOLVER simply goes through a process of changing slope and/or intercept and checking whether the sum of squares has got smaller. When it can no longer get the sum of squares any smaller, it reports that it has been minimized. The cells for slope and intercept now contain values for the best fitting line through the data. Figure 2-12.7 shows the results of the SOLVER analysis with the values of slope = 15.49 and intercept = 22.81 being returned for the best fitting straight line. Note that cell G4 contains the correlation coefficient between predicted and observed Y scores. This was achieved by entering the equation =CORREL(B2:B18,C2:C18), which returns the correlation coefficient between the two columns of data.

### Goodness of Fit of Models

Once you have collected real data, the test of your model is how well it fits that data. The question is how do we judge how well that data has been fit? When dealing with the common situation of predicting a dependent variable, the sum of squares between real and predicted values is a good measure of fit. In addition the shape of the predicted curve in comparison to the real data is important to consider. It is possible to have similar sum of squares for two models fit to the same data but that one has some shape fitting characteristics that are superior to the other model.

Often in complex curves, a residual plot is used to determine if there is some pattern or systematic way in which the model fails to fit the data (Figure 2-12.8). In our example the residual plot is not necessary to show that the straight line does not reflect the shape of this data. In more complex models, the residual plot should appear to have no pattern with residuals oscillating around zero. If there is a pattern in the residual plot, then this indicates that your model fails in some systematic way and must be modified.



**Figure 2-12.8: Line fit plot (top) and Residual plot (bottom) for straight line fit of curvilinear data.**

### Modeling Height Growth Curves using SOLVER (an MS EXCEL function)

Our first example of modeling is the Preece-Baines model 1 (1978) that was developed to explain the complex curve of growth of humans. The curve of human growth is a complex s-shaped curve. Over the years many have tried to describe this in mathematical terms. To date this has best been achieved by the work of Preece and Baines in their production of a double exponential equation:

$$\text{Preece-Baines Model 1} \quad h_t = h_1 - \frac{2(h_1 - h_q)}{e^{[s_0(t-q)]} + e^{[s_1(t-q)]}}$$

where:  $h_t$  is height at time  $t$ ,

$h_1$  is final height,

$s_0$  and  $s_1$  are rate constants,

$q$  is a time constant and

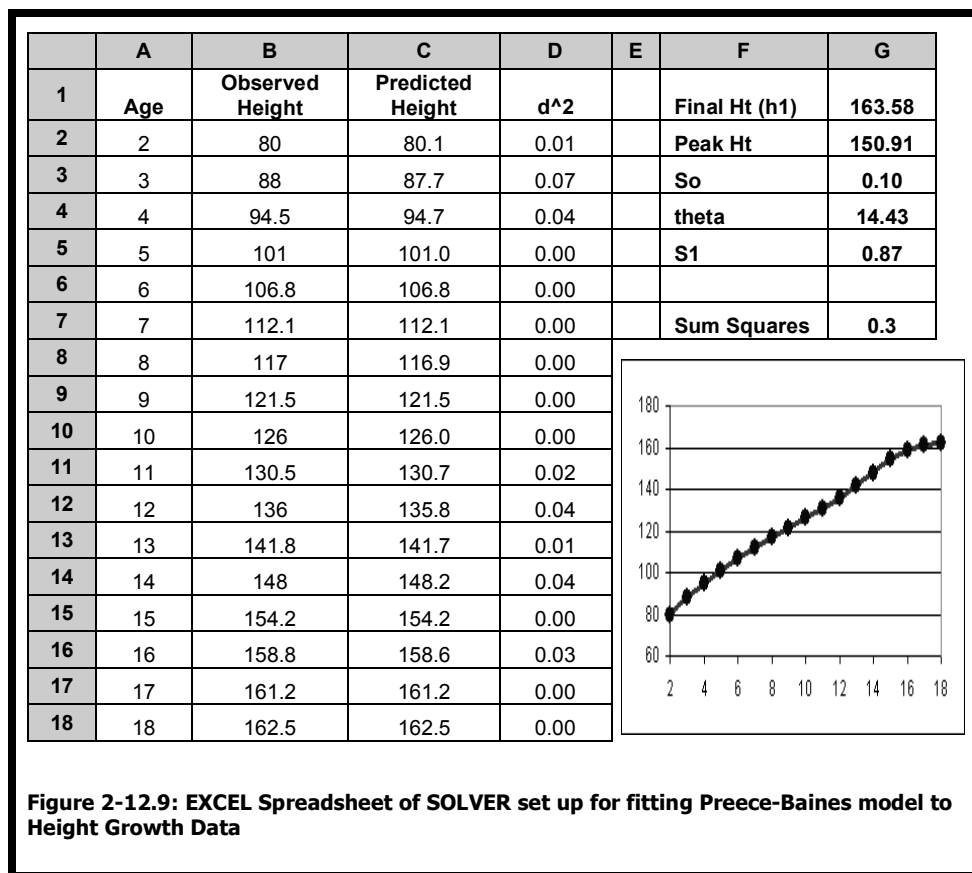
$h_q$  is height at  $t = q$ .

The basic concept that we will be using in this modeling process, is that of calculating the best fitting equations to describe real data, based on the criterion of the Least Sum of Squares. Previously the SOLVER setup for least sum of squares linear curve fitting was described. Once you have set up the spreadsheet you can fit very complex equations by simply changing the equation in the predicted Y column. This is how the complex Preece-Baines model 1 shown above can be fit to real data easily with EXCEL.

The result of using SOLVER and this equation on height growth data is shown in Figure 2-12.9. The differences between this and the set up for the straight line in Figure 2-12.8 are that the equation in column C is different and that there are 5 unknown's (cells G1:G5) rather than the two unknown's (slope and intercept) in the straight line example. The equation typed into cell C2 and then copied down the column was:

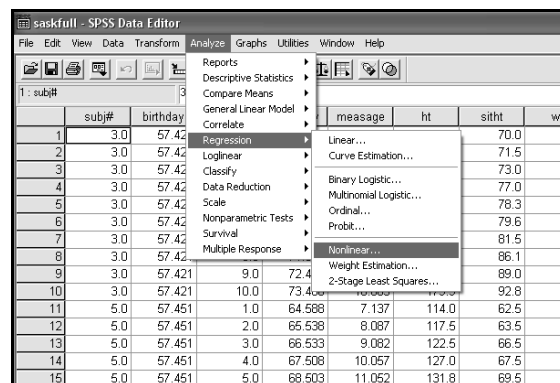
$$=G\$1-((2*(G\$1-G\$2))/((EXP(G\$3*(A2-G\$4)))+(EXP(G\$5*(A2-G\$4))))))$$

The graph shows how similar the observed values for height (bullets) and the predicted values of height (line) are.



**Modeling Height Growth Curves using SPSS**

Although SOLVER can be used effectively to find out your model parameter estimates, it is limited to analysis of one curve at a time. However, if you have SPSS available, then you can carry out multiple curve fitting and also receive the more detailed report of model statistics and parameter estimates. To carry out modeling with SPSS you use the nonlinear regression option of the analyze menu shown in Figure 2-12.10. Once nonlinear regression has been selected, the model definition dialog box appears (Figure 2-12.11 - top). This requires that you define the parameters of the model. In the case of our Preece-Baines model these are final ht, peak ht, theta and the rate constants s0 and s1.



**Figure 2-12.10: SPSS analyze menu for selection of nonlinear regression**

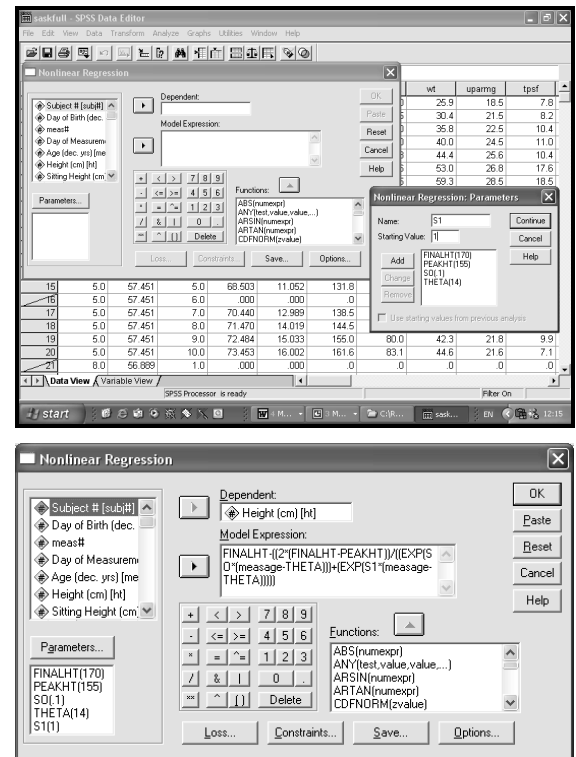
As with the use of SOLVER, this is an iterative procedure that requires initial starting values for

each parameter. The data being used in this example are a file of anthropometric data from the Saskatchewan Growth Study kindly made available by Dr. Don Bailey. In this dataset, height and other anthropometric measures were collected annually on boys from age 7 to 18 years. The data file is organized such that each subject has several lines of data, one for each measurement occasion. When running nonlinear regression on longitudinal data you should turn on a Split File by subject ID #, such that the model will then be fit for each individual subject sequentially. The bottom panel of Figure 2-12.11 shows the completed model definition for fitting the Preece-Baines model to the longitudinal Saskatchewan data. Initial parameter estimates of FINALHT = 170, PEAKHT = 155, SO = 0.1, THETA = 14 and S1 = 1

were used. There is nothing magical about these values, other than they are ballpark estimates of the expected values. However, sometimes SPSS fails to find a solution to the nonlinear fit because the initial parameter estimates are not close enough. This can be cured by changing one or more initial values to allow a solution to be found.

Figure 2-12.12 shows the SPSS output for the model defined in Figure 2-12.11 and run on subject #3 of the Saskatchewan data. The output reports the parameter estimates for each iteration. Notice that it took 14 iterations until the solution was found. A solution was declared when the residual sum of squares failed to decrease by the accepted tolerance. Note that the final solution had a residual sum of squares of 1.60796 with an associated r squared of 0.99944. It gave final estimates of the parameters as:

FINALHT	182.3
PEAKHT	171.94
SO	0.13
THETA	14.34



**Figure 2-12.11: SPSS nonlinear regression dialog box for definition of model parameters (top) and the dialog box showing the completely defined Preece-Baines model (bottom)**

S1

13

**Non-linear Regression**

SUBJ#: 3.0

Iteration	Residual SS	FINALHT S1	PEAKHT	SO	THETA
1	2750.379534	170.000000 1.00000000	155.000000	.100000000	14.0000000
1.1	24.76460563	182.618873 1.08144292	172.330440	.131556135	14.3330856
2	24.76460563	182.618873 1.08144292	172.330440	.131556135	14.3330856
2.1	1.614926667	182.455552 1.12599347	171.985610	.135573756	14.3509246
3	1.614926667	182.455552 1.12599347	171.985610	.135573756	14.3509246
3.1	1.607987874	182.396122 1.13308354	171.948522	.135743790	14.3426556
4	1.607987874	182.396122 1.13308354	171.948522	.135743790	14.3426556
4.1	1.607960619	182.386708 1.13452404	171.945145	.135817339	14.3419360
5	1.607960619	182.386708 1.13452404	171.945145	.135817339	14.3419360
5.1	1.607959635	182.384746 1.13480182	171.944345	.135831626	14.3417675
6	1.607959635	182.384746 1.13480182	171.944345	.135831626	14.3417675
6.1	1.607959599	182.384368 1.13485531	171.944194	.135834394	14.3417354
7	1.607959599	182.384368 1.13485531	171.944194	.135834394	14.3417354
7.1	1.607959597	182.384297 1.13486529	171.944165	.135834910	14.3417294

Run stopped after 14 model evaluations and 7 derivative evaluations.  
Iterations have been stopped because the relative reduction between successive residual sums of squares is at most SSSCON = 1.000E-08

Nonlinear Regression Summary Statistics			Dependent Variable HT	
Source	DF	Sum of Squares	Mean Square	
Regression	5	241471.61204	48294.32241	
Residual	5	1.60796	.32159	
Uncorrected Total	10	241473.22000		
(Corrected Total)	9	2894.30400		

R squared = 1 - Residual SS / Corrected SS = .99944

Asymptotic 95 %

Parameter	Estimate	Asymptotic Std. Error	Confidence Interval	
			Lower	Upper
FINALHT	182.38429707	1.523467472	178.46809926	186.30049488
PEAKHT	171.94416514	1.028414497	169.30054152	174.58778877
SO	.135834910	.010217585	.109569771	.162100049
THETA	14.341729435	.202010828	13.822444071	14.861014800
S1	1.134865287	.174165533	.687158532	1.582572042

**Figure 2-12.12: SPSS Output for nonlinear regression of Preece-Baines height prediction model on male subject #3 from the Saskatchewan Growth study**

### Mechanical Modeling of Skinfold Compression

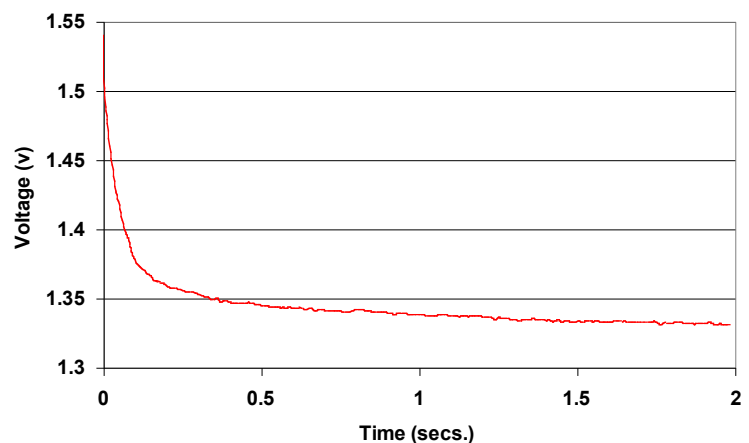
Our second example of modeling is the mechanical modeling of dynamic skinfold compression. The measurement of subcutaneous adipose tissue with skinfold calipers is a routine laboratory and field technique used in the assessment of body composition and nutritional status. Because skinfold calipers are spring operated, they exert a finite pressure on the tissue. Skinfold measurements therefore represent the measured compressed thickness of a double fold of skin and underlying adipose tissue. In an ideal situation, however, skinfold thickness would be measured at zero pressure without tissue compression.

Dynamic compressibility refers to the exponential decline in caliper reading after the initial application of the caliper to the skinfold. In order to investigate dynamic compression, Ward et al. modified a simple Slim Guide skinfold caliper with the addition of a potentiometer, which provided an analogue signal proportional to caliper jaw opening (Chapter 3-3).

This was used to collect compression data at three sites in 8 men and 8 women. The data were then modeled with a combination of spring and viscous elements, producing coefficients representing the elasticity and viscosity of the modeled components of individual skinfolds. The modified skinfold caliper used for this data collection is described in chapter 3-3 as an example of a simple A/D set up.

A skinfold is composed of a double layer of skin plus subcutaneous adipose tissue. The characteristics of the skinfold under compression by a skinfold caliper exerting a constant pressure are determined by the biophysical properties of the component tissues. These biophysical properties are dependent upon both the properties of the skin and the underlying subcutaneous adipose tissue. Under compression there appears to be a two component curve (2-12.9):

1. The initial fast decline phase influenced by the properties of the skin; elasticity and resting tension



**Figure 2-12.9: Typical dynamic skinfold compression curve using modified slim guide caliper**

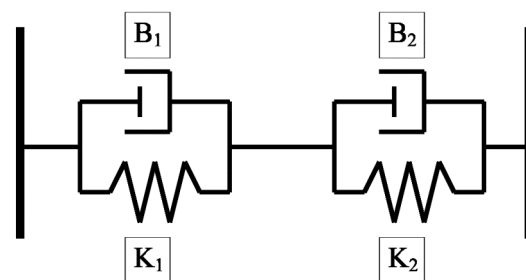
## 2. The slower second phase of compression of skin plus underlying adipose tissue

A mechanical model comprised of a spring and dashpot (viscous element) has been successfully used to model the response of a tissue to stretch and compression. There is an reorientation of the fibres in the initial phase followed by a stretching of the elastic components, each phase being modified by the hydration (viscosity) of the tissues, with the resultant displacement curve being similar to that of a damped spring. The hypothesized model is that there is a damped spring representing the response of the skin to initial compression, the characteristics of this displacement curve being determined by skin elasticity, original tension and hydration. Then, in series with that is another damped spring representing the response of the skin plus adipose tissue to compression in the second slower phase of the displacement curve. Even though our initial intention was to collect a diverse sample of individuals to evaluate our model, it was still possible to show significant differences between the sexes in the model coefficients despite the small sample sizes. The above observations from the literature predict that women have more elastic and less viscous tissues, observations confirmed by the direction of the differences in the elasticity and viscosity coefficients of the model.

Having determined that a two component model was the most appropriate a mechanical model based upon two tissues with both elastic and viscous elements was developed. The model, shown diagrammatically in Figure 2-12.10 is comprised of a spring and viscous element in parallel which is in series with another parallel spring and viscous element (dashpots).  $k_1$  and  $k_2$  are the spring constants and  $b_1$  and  $b_2$  are the coefficients of viscosity of the two dashpots. This model is represented by the following equation:

$$Tt = T_{initial} + F \left( \frac{1}{k_1} - \left[ \frac{e^{-\frac{k_1 t}{b_1}}}{k_1} \right] \right) + F \left( \frac{1}{k_2} - \left[ \frac{e^{-\frac{k_2 t}{b_2}}}{k_2} \right] \right)$$

Figure 2-12.11 Shows the result of signal averaging 10 trial curves, with superimposed the results of the least sum of squares curve fitting of the two component mechanical model. The two curves are virtually identical, the sum of squares being 1.417 with a standard error of estimate of 0.084 mm.



**Figure 2-12-10: Mechanical model of skinfold compression**

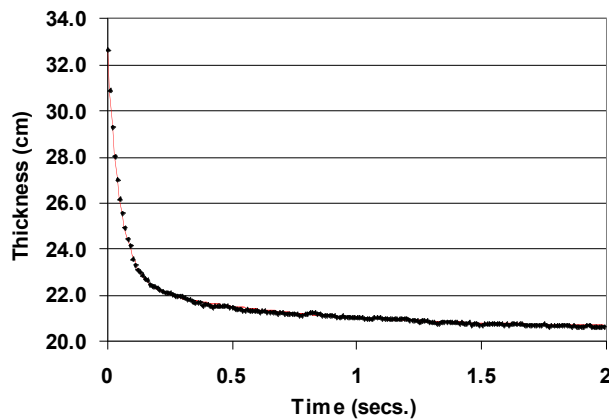


Table 2-12.1 shows the mean coefficients for the males and females when all three sites are combined. Using an independent T-test significant differences ( $p < 0.05$ ) were found between the sexes in each of the viscosity and elasticity coefficients, although no significant difference was found in  $T_{init}$  the estimated thickness at zero compression. The coefficients indicated that females were more elastic and less viscous than males in the dynamic compression of their skinfold tissues.

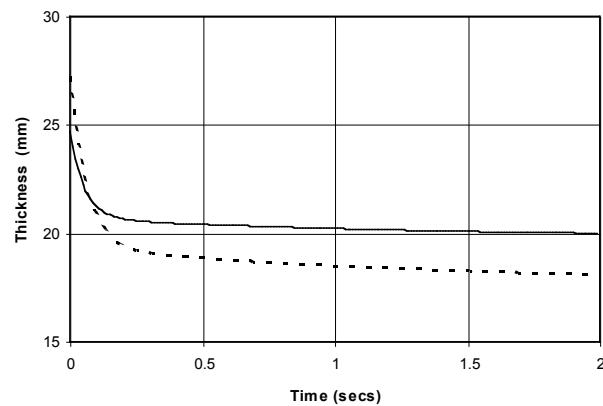
<b>Women</b>	<b>Mean</b>	-1.275	-6.276	-0.077	-7.738	27.255
	<b>S.E.M.</b>	0.251	0.970	0.010	1.672	3.684
<b>Men</b>	<b>Mean</b>	-2.499*	-11.362*	-0.135*	-14.59*	24.692
	<b>S.E.M.</b>	0.883	3.618	0.039	5.796	7.807
* = significantly different from Women mean using independent T-test ( $p < 0.05$ )						

**Table 2-12.1: Means and standard errors of mean for values of model coefficients  $k_1$ ,  $k_2$ ,  $b_1$ ,  $b_2$  and  $T_{init}$  for women ( $n=8$ ) and men ( $n=8$ ) all three skinfold sites**

Table 2-12.1 shows the means and standard errors of mean for values of model coefficients  $k_1$ ,  $k_2$ ,  $b_1$ ,  $b_2$  and  $T_{init}$  for females ( $n=8$ ) and males ( $n=8$ ) at three skinfold sites. This shows a significant difference in model coefficients between males and females. This difference is depicted in the average curves for males and females in Figure 2-12.12, where female skinfolds are seen to be more elastic and less viscous.



**Figure 2-12.11: Predicted skinfold thickness (red line) superimposed upon actual skinfold thicknesses (black bullets)**



**Figure 2-12.12: Predicted skinfold thickness (red line) superimposed upon actual skinfold thicknesses (black bets)**

## References

- Preece, M.A., and M.J. Baines. 1978. A new family of mathematical models describing the human growth curve. *Ann. Hum. Biol.* 5:1-24.
- Ward, R., J. Schlenker and G.S. Anderson. 2001. Simple method for developing percentile curves for height and weight. *Am. J. Phys. Anthrop.* 116:246-250.