International Journal of Forecasting 16 (2000) 521–530
www.elsevier.com/locate/ijforecast

The theta model: a decomposition approach to forecasting

V. Assimakopoulos*, K. Nikolopoulos

Department of Electrical and Computer Engineering, Forecasting Systems Unit, National Technical University of Athens, 15773 Zografou, Athens, Greece

Abstract

This paper presents a new univariate forecasting method. The method is based on the concept of modifying the local curvature of the time-series through a coefficient 'Theta' (the Greek letter θ), that is applied directly to the second differences of the data. The resulting series that are created maintain the mean and the slope of the original data but not their curvatures. These new time series are named Theta-lines. Their primary qualitative characteristic is the improvement of the approximation of the long-term behavior of the data or the augmentation of the short-term features, depending on the value of the Theta coefficient. The proposed method decomposes the original time series into two or more different Theta-lines. These are extrapolated separately and the subsequent forecasts are combined. The simple combination of two Theta-lines, the Theta = 0 (straight line) and Theta = 2 (double local curves) was adopted in order to produce forecasts for the 3003 series of the M3 competition. The method performed well, particularly for monthly series and for microeconomic data. © 2000 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.

Keywords: M3-Competition; Time series; Univariate forecasting method

1. Introduction

There have been many attempts to develop forecasts based directly on decomposition (Makridakis et al., 1984). The individual components that are usually identified are the trend-cycle, seasonality and the irregular component. These are projected separately into the future and recombined to form a forecast of the underlying series. This approach in practice is not frequently used. The main difficulties are in isolating successfully the error component as well as in producing adequate forecasts for the trend-cycle. Perhaps the only technique that has been found to work relatively well is to forecast the seasonally adjusted data using Holt's method (Makridakis et al., 1984) or the dampen trend method (Gardner & McKenzie, 1985) and then adjust the forecasts using the seasonal components from the end of the data.

The Theta-model proposes a different approach to decomposition: a decomposition of the seasonally adjusted series into short and long term components.

The challenge for the proposed method was

*Corresponding author. Tel.: +30-1-772-3742; fax: +30-1-772-3740.

E-mail address: vassim@epu.ntua.gr (V. Assimakopoulos).

0169-2070/00/$ – see front matter © 2000 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.
PII: S0169-2070(00)00066-2

to increase the degree of exploitation of the embedded useful information in the data, before the application of a forecasting method. Viewed intuitively, such information has long and short-term components. These components are identified using the Theta-model and are then extrapolated separately. The Theta-model operation is analogous to the operation of a magnifying glass through which the time series fluctuations are minimized or maximized accordingly. The combination of the components-forecasts thus becomes more effective while retaining the benefits from combining.

Combining under certain circumstances improves forecasting accuracy (Clemen, 1989). The reason lies in the averaging of errors that are produced by each individual forecasting method. These errors relate to the instability of patterns or relationships, to the minimization procedures for the selection of the best model to use, or even to measurement weaknesses (Makridakis, Wheelwright & Hyndman, 1998). Above all, errors are associated with the nature of the chosen model. Each model or functional form imposes its own logic on the data in a more or less flexible way, and this specific logic is subsequently extrapolated to the future. If there is an amount of useful information within the time series, then there is also an accompanying degree of exploitation of this information associated with each distinct forecasting method.

In this sense Theta can be seen as an alternative decomposition approach or/and as an extension to the concept of combining.

2. The Theta-model

The model is based on the concept of modifying the local curvatures of the time series. This change is obtained from a coefficient, called Theta-coefficient (as a symbol is used the Greek letter Theta), which is applied directly to the second differences of the time series:

$$X''_{\text{new}}(\theta) = \theta \cdot X''_{\text{data}}, \quad \text{where } X''_{\text{data}} = X_t - 2X_{t-1} + X_{t-2} \text{ at time } t.$$

If the local curvatures are gradually reduced then the time series is deflated as it is shown in Fig. 1. The smaller the value of the Theta-coefficient, the larger the degree of deflation. In the extreme case where $\Theta = 0$ the time series is transformed to a linear regression line. The progressive decrease of the fluctuations di-

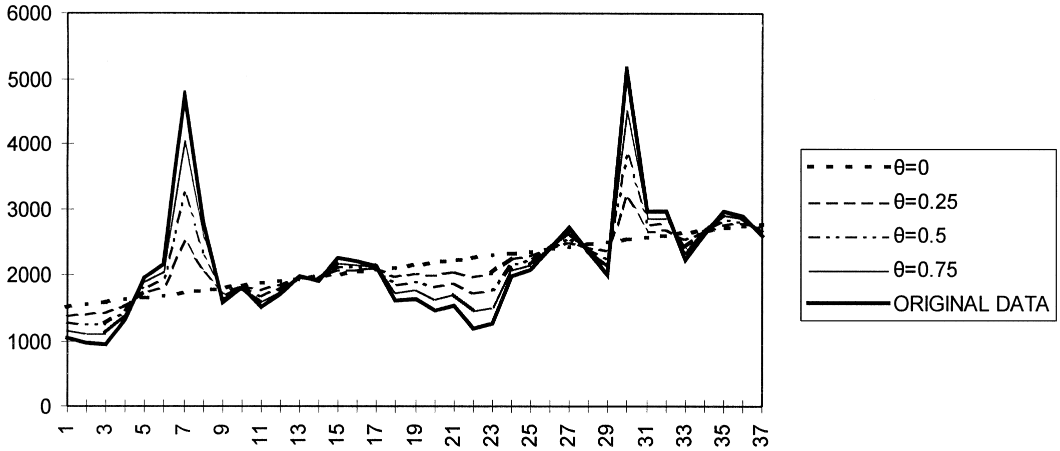

Fig. 1. M3-Comp. Series 200, the Theta-model deflation.

minishes the absolute differences between successive terms in the derived series and is related, in qualitative terms, to the emergence of long-term trends in the data (Assimakopoulos, 1995).

The Θ-coefficient can also take negative values but they are of no interest in the present context and are not discussed further.

Conversely if the local curvature is increased ($\Theta > 1$), then the time series is dilated as it is shown in Fig. 2. The larger the degree of dilation, the larger the magnification of the short-term behavior.

Following this procedure, a set of new time series, the so-called Theta-lines, are constructed. The placement of these lines in relation to the original data can be done in many different ways. If the fitting is an OLS estimation procedure then the mean and the slope of the Theta-lines remain the same compared to those of the original data (see Appendix A).

The general formulation of the method becomes as follows:

The initial time series is decomposed into two or more Theta-lines. Each of the Theta-lines is extrapolated separately and the forecasts are simply combined. Any forecasting method can be used for the extrapolation of a Theta-line according to existing experience (Fildes, Hibon, Makridakis & Meade, 1998). A different combination of Theta-lines can be employed for each forecasting horizon.

This is demonstrated by considering one of the simplest cases in which the initial time series is decomposed into two Theta-lines, i.e. $\Theta = 0$ and $\Theta = 2$:

$$\text{Data} = 1/2(L(\Theta = 0) + L(\Theta = 2))$$

where $L(\Theta = 0)$ stands for the Theta Line for Θ parameter equal to zero.

The first Theta-line ($\Theta = 0$) is the linear regression line of the data (see Appendix B) and the second one has second differences exactly twice the initial time series. This is a case where two, extreme and symmetrical to 1, Theta-lines are composed (see Appendix B). The first component $L(\Theta = 0)$ describes the time series through a linear trend. The second one, $L(\Theta = 2)$, has doubled the local curvatures magnifying the short-term behavior. The first Theta-line is extrapolated in the usual way for a linear trend. The second is extrapolated via simple exponential smoothing. The simple combination of the two forecasts gives the final forecast of the Theta-model for the specific time series as it is shown in Fig. 3.

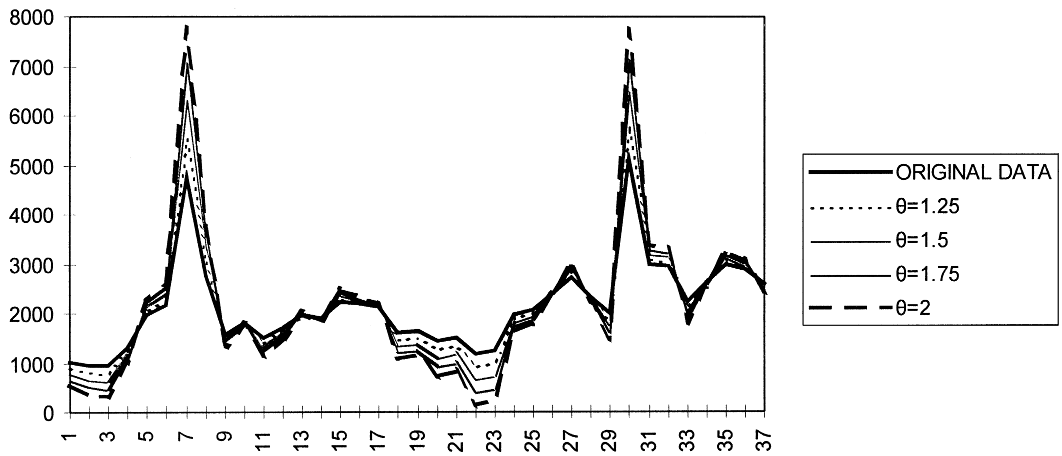

Fig. 2. M3-Comp. Series 200, the Theta-model dilation.

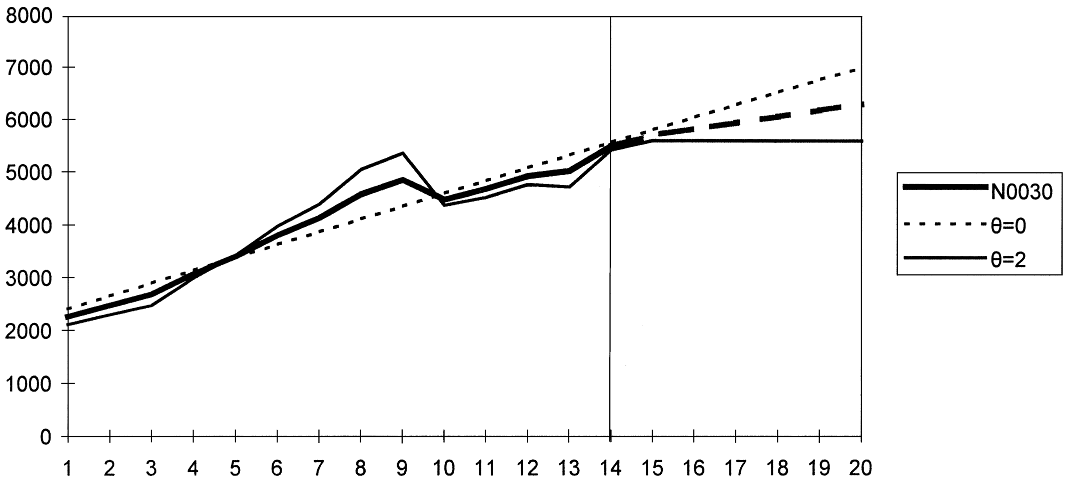

Fig. 3. M3-Comp. Series 30, the Theta-model forecasts.

This combination of Theta-lines $\Theta=0$ and $\Theta=2$ was employed to produce forecasts for the 3003 time-series of the M3 competition.

The steps followed are:

Step 0. (Seasonality testing) Firstly each time series was tested for statistical significant seasonal behavior. The criterion was the *t*-test value for the auto-correlation function value with lag one year (that is for monthly time series 12 observations and for quarterly time-series 4 observations) compared to 1.645 which is the *t*-statistic value for 0.1 probability.[1]

Step 1. (Deseasonalisation) The time-series were deseasonalised via the classical decomposition method (multiplicative).

Step 2. (Decomposition) Each time-series was decomposed into two Theta-lines, the linear regression line ($\Theta=0$) and the Theta line for $\Theta=2$.

[1]The set of seasonal indices given by M. Hibbon and S. Makridakis after ISF98 for the 3003 time series of the M3-competition were not used in the seasonal adjustment procedure.

Step 3. (Extrapolation) The linear regression line is extrapolated in the usual way while the second line is extrapolated via simple exponential smoothing.

Step 4. (Combination) The forecasts produced from the extrapolation of the two lines were combined with equal weights.

Step 5. (Reseasonalisation) The forecasts were reseasonalised.

3. Evaluation

The strong point of the method lies in the decomposition of the initial data. The two components include information, which is useful for the forecasting procedure but is lost or cannot completely be taken into account by the existing methods when they are directly applied to the initial data. Especially in the case of $L(\Theta = 0)$ this phenomenon is more comprehensible. The straight line includes information for the long-term trend of the time series which is ‘‘neglected’’ when a method tries to adapt to more recent trends. On the other hand, when the linear trend is used exclusively all valuable information on short-term fluctuations is ignored.

The Theta-model performance in the monthly time series of the M3 competition constitutes a characteristic example. The monthly data of the competition were characterised, in general, by a relatively large amount of volatility. This fact does not allow most methods to keep in memory the long-term trend and thus to take it into serious consideration in their forecasting function. In the case of Theta-model the long-term trend is incorporated into the method as a major component through the $L(\Theta = 0)$ and extrapolation is straightforward. At the same time, the existence of $L(\Theta = 2)$ operates as a counterbalance to the simplification of using a plain linear trend model. $L(\Theta = 2)$ increases the roughness of the monthly time series and augments the most recent trends. The effect of this augmentation is that the combined starting point reaches the ''correct'' level and since the extrapolation of $L(\Theta = 2)$ is horizontal the simple combination of both preserves a conservative but constant continuation of the long-term trend.

4. Perspectives for future research

The two-line variant ($\Theta = 0$ and $\Theta = 2$) is only one of the several possibilities that result from the general formulation of the method. The first extension is to use more than two Theta-lines with Θ-coefficients which are not symmetric to one (see Appendix B). In this case the data are not decomposed and each Theta line is used only to produce a set of forecasts which will be combined accordingly. This is expected to add to the robustness of the method, under the condition that a relatively efficient method for each Theta-line will be selected. On the other hand, it is not certain that the use of several Theta-lines will contribute to more accurate forecasts since the separation of the incorporated information in the initial time-series may not be sufficiently well defined.

Another option is to use different Theta-lines combinations for each forecasting horizon. There is empirical evidence (Collopy & Armstrong, 1992) that for longer horizons forecasts should be biased more to long-term behavior while for shorter-term forecasts we should mostly take into account the recent trends. This can be accomplished easily by using different pairs of Theta-lines for each forecasting horizon. For example if the couple $\Theta = 0$ and $\Theta = 1.5$ is used then greater emphasis is placed on the long-term trend of the time-series while in the case of the Theta-lines $\Theta = 0$ and $\Theta = 2.5$, the short-term behavior gains more importance.

The last and most promising characteristic of the model is the utilisation of different Theta-lines for each time series. A pair of Theta-lines will correspond to each time series according to its qualitative and/or quantitative characteristics (Armstrong & Collopy, 1993). This is the objective of the further research regarding the Theta-model.

Appendix A

The data X_i can be written as:

$$X_i = X_1 + (i-1)(X_2 - X_1) + \left(\sum_{t=2}^{i-1} (i-t) X''_{t+1} \right)$$

The points Y_i of a Theta − Line by definition are:

$$Y_i = Y_1 + (i-1)(Y_2 - Y_1) + \theta\left(\sum_{t=2}^{i-1} (i-t)X''_{t+1}\right)$$

The minimization problem becomes:

$$\begin{aligned}\min\left(\sum_i e_i^2\right) &= \min\left(\sum_i (Y_i - X_i)^2\right)\\ &= \min\left(\sum_i \left(Y_1 + (i-1)(Y_2 - Y_1) + \theta\left(\sum_{t=2}^{i-1} (i-t)X''_{t+1}\right) - X_1 - (i-1)(X_2 - X_1)\right.\right.\\ &\left.\left.\quad -\left(\sum_{t=2}^{i-1} (i-t)X''_{t+1}\right)\right)^2\right)\end{aligned}$$

Applying calculus,

$$\left\{\begin{aligned}&\frac{\partial \sum_i e_i^2}{\partial Y_1} = 2\sum_i \frac{\partial (Y_i - X_i)}{\partial Y_1}(Y_i - X_i) = 0\\ &\frac{\partial \sum_i e_i^2}{\partial (Y_2 - Y_1)} = 2\sum_i \frac{\partial (Y_i - X_i)}{\partial (Y_2 - Y_1)}(Y_i - X_i) = 0\end{aligned}\right\}\Leftrightarrow$$

$$\left\{\begin{aligned}&2\sum_i \left(Y_1 + (i-1)(Y_2 - Y_1) + \theta\left(\sum_{t=2}^{i-1}(i-t)X'_{t+1}\right) - X_1 - (i-1)(X_2 - X_1) - \left(\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right)\right) = 0\\ &2\sum_i (i-1)\left(Y_1 + (i-1)(Y_2 - Y_1) + \theta\left(\sum_{t=2}^{i-1}(i-t)X_{t+1}\right) - X_1 - (i-1)(X_2 - X_1) - \left(\sum_{t=2}^{i-1}(i-t)X_{t+1}\right)\right) = 0\end{aligned}\right\}\Leftrightarrow$$

$$\left\{\begin{aligned}&nY_1 + \frac{n(n-1)}{2}(Y_2 - Y_1) = nX_1 + \frac{n(n-1)}{2}(X_2 - X_1) + (1-\theta)\left(\sum_i \sum_{t=2}^{i-1}(i-t)X''_{t+1}\right)\\ &\frac{n(n-1)}{2}Y_1 + \frac{n(n-1)(2n-1)}{6}(Y_2 - Y_1) = \frac{n(n-1)}{2}X_1 + \frac{n(n-1)(2n-1)}{6}(X_2 - X_1)\\ &+(1-\theta)\left(\sum_i (i-1)\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right)\end{aligned}\right\}\Leftrightarrow$$

$$\left\{\begin{aligned}&Y_1 + \frac{(n-1)}{2}(Y_2 - Y_1) = X_1 + \frac{(n-1)}{2}(X_2 - X_1) + \frac{(1-\theta)}{n}\left(\sum_i \sum_{t=2}^{i-1}(i-t)X''_{t+1}\right), &\text{(A.1)}\\ &Y_1 + \frac{(2n-1)}{3}(Y_2 - Y_1) = X_1 + \frac{(2n-1)}{3}(X_2 - X_1) + \frac{2(1-\theta)}{n(n-1)}\left(\sum_i (i-1)\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right), &\text{(A.2)}\end{aligned}\right.$$

The mean value of a Theta-Line is:

$$\bar{Y}_i = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n}\left(Y_1 + (i-1)(Y_2 - Y_1) + \theta \sum_{t=2}^{i-1}(i-t)X''_{t+1}\right) \Rightarrow$$
$$\bar{Y}_i = \frac{1}{n}\left(Y_1 \sum_{i=1}^{n} + (Y_2 - Y_1)\sum_{i=1}^{n}(i-1) + \theta \sum_{i=1}^{n}\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right) \Rightarrow$$
$$\bar{Y}_i = \frac{1}{n}\left(nY_1 + \frac{n(n-1)}{2}(Y_2 - Y_1) + \theta \sum_{i=1}^{n}\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right) \Rightarrow$$
$$\bar{Y}_i = Y_1 + \frac{(n-1)}{2}(Y_2 - Y_1) + \frac{\theta}{n}\sum_{i=1}^{n}\sum_{t=2}^{i-1}(i-t)X_{t+1} \overset{(1)}{\Rightarrow}$$
$$\bar{Y}_i = X_1 + \frac{(n-1)}{2}(X_2 - X_1) + \frac{1}{n}\sum_{i=1}^{n}\sum_{t=2}^{i-1}(i-t)X''_{t+1} \Rightarrow$$
$$\bar{Y}_i = \bar{X}_i$$

The formula for the slope of a Theta-Line is:

$$b_\theta = c_1 \sum_i iY_i - c_2 \sum_i Y_i, \quad \begin{bmatrix} c_1 = \dfrac{12}{n(n+1)(n-1)} \\ c_2 = -\dfrac{6}{n(n-1)} \end{bmatrix},$$

or

$$b_\theta = Y_2 - Y_1 + \theta c(i, X''_t), \quad \left[c(i, X''_t) = \sum_i (ic_1 + c_2) \sum_{t=2}^{i-1}(i-t)X''_{t+1} \right]$$

Subtracting Eq. (A.1) from (A.2) gives,

$$\left(\frac{2n-1}{3} - \frac{n-1}{2}\right)(Y_2 - Y_1) = \left(\frac{2n-1}{3} - \frac{n-1}{2}\right)(X_2 - X_1)$$
$$+ (1-\theta)\left(\sum_i \left[\frac{2}{n(n-1)}(i-1) - \frac{1}{n}\right]\sum_{t=2}^{i-1}(i-t)X''_{t+1}\right) \Leftrightarrow$$

$$Y_2 - Y_1 = X_2 - X_1 + (1-\theta)c'(i, X''_t), \quad \text{where } c'(i, X''_t) = c(i, X''_t)$$

Thus,

$$b_\theta = X_2 - X_1 + (1-\theta)c'(i, X''_t) + \theta c(i, X''_t) \Rightarrow$$

$$b_\theta = X_2 - X_1 + c(i, X''_t) \Rightarrow$$
$$b_\theta = b_{\text{time-series}}$$

where $b_{\text{time-series}} = b_1$ is the slope of the raw data, since if $\Theta = 1$ the time series remains untouched.

Appendix B

From Appendix A:

$$X_i = X_1 + (i-1)(X_2 - X_1) + \sum_{t=2}^{i-1} (i-t)X''_{t+1}$$

$$\Rightarrow \sum_{t=2}^{i-1} (i-t)X''_{t+1} = X_i - X_1 - (i-1)(X_2 - X_1)$$

Substituting this summation in Eqs. (A.1), (A.2) yields to:

$$\begin{cases} Y_1 + \frac{(n-1)}{2}(Y_2 - Y_1) = X_1 + \frac{(n-1)}{2}(X_2 - X_1) \\ \quad + \frac{(1-\theta)}{n}\left(\sum_i (X_i - X_1 - (i-1)(X_2 - X_1))\right), \\ Y_1 + \frac{(2n-1)}{3}(Y_2 - Y_1) = X_1 + \frac{(2n-1)}{3}(X_2 - X_1) \\ \quad + \frac{2(1-\theta)}{n(n-1)}\left(\sum_i (i-1)(X_i - X_1 - (i-1)(X_2 - X_1))\right), \end{cases} \quad \begin{matrix} \text{(B.1a)} \\ \\ \\ \text{(B.2a)} \end{matrix}$$

Evaluating the summations leads to:

$$\begin{cases} Y_1 + \frac{(n-1)}{2}(Y_2 - Y_1) = \theta X_1 + \frac{\theta(n-1)}{2}(X_2 - X_1) + (1-\theta)\overline{X_n}, \\ Y_1 + \frac{(2n-1)}{3}(Y_2 - Y_1) = \theta X_1 + \frac{\theta(2n-1)}{3}(X_2 - X_1) + \frac{2(1-\theta)}{n(n-1)} \\ \quad \left(\sum_i (i-1)X_i, \right. \end{cases} \quad \begin{matrix} \text{(B.1b)} \\ \\ \text{(B.2b)} \end{matrix}$$

By setting

$$\frac{1}{n}\sum (i-1)X_i - \frac{(n-1)}{2}\overline{X_n} = \mathrm{cov}_n(i-1, X_i) = C_n$$

the previous equations become:

$$\begin{cases} Y_2 - Y_1 = \theta(X_2 - X_1) + \frac{12(1-\theta)}{(n-1)(n+1)} C_n \\ Y_1 = \theta X_1 + (1-\theta)\overline{X_n} - \frac{6(1-\theta)}{(n+1)} C_n \end{cases}$$

By setting

$$V_n = Var(i-1) = \frac{1}{n}\left[\sum (i-1)^2 - \left(\frac{n-1}{2}\right)^2\right] = \frac{(n-1)(n+1)}{12}$$

the equations become:

$$\begin{cases} Y_2 - Y_1 = \theta(X_2 - X_1) + (1-\theta)b_n, & \text{(B.3a)} \\ Y_1 = \theta X_1 + (1-\theta)\overline{X_n} - (1-\theta)b_n \dfrac{(n-1)}{2}, & \text{(B.3b)} \end{cases}$$

From this set of equations the following results are obvious:

1. For Θ-coefficients = 0 the linear regression line is produced,

$$\theta = 0 \Rightarrow \begin{cases} Y_2 - Y_1 = b_n \\ Y_1 = \overline{X_n} - b_n \dfrac{(n-1)}{2} \end{cases}$$

   which are the standard LS equations.

2.

$$Y_i(\theta),\, Y_i(-\theta) \Rightarrow \begin{cases} \dfrac{1}{2}\,[Y_1(\theta) + Y_1(-\theta)] = Y_1(0) \\ \dfrac{1}{2}\,[Y_2(\theta) + Y_2(-\theta)] = b_n + Y_1(0) \end{cases}$$

3.

$$\theta = 1 + a \Rightarrow \begin{cases} \dfrac{1}{2}\,[Y_1(1+a) + Y_1(1-a)] = X_1 \\ \dfrac{1}{2}\,[Y_2(1+a) + Y_2(1-a)] = X_2 \end{cases}$$

From the above result becomes obvious that if two *lines* are produced from symmetric to 1 Θ-coefficients, for example $\Theta_1 = 1 + a$ and $\Theta_2 = 1 - a$, the average of these two lines reproduces the original time series. That is:

$$\theta = 1 \pm a \Rightarrow \frac{1}{2}\,[Y_i(1+a) + Y_i(1-a)] = X_i$$

The points of a new Theta line are calculated via the formula

$$Y_i = Y_1 + (i-1)(Y_2 - Y_1) + \theta\left(\sum_{t=2}^{i-1} (i-t)X''_t\right)$$

So for Θ_1 and Θ_2 it is derived that:

$$\begin{cases} Y_i(1+a) = Y_1(1+a) + (i-1)(Y_2(1+a) - Y_1(1+a)) + (1+a)\left(\displaystyle\sum_{t=2}^{i-1} (i-t)X''_{t+1}\right) \\ Y_i(1-a) = Y_1(1-a) + (i-1)(Y_2(1-a) - Y_1(1-a)) + (1-a)\left(\displaystyle\sum_{t=2}^{i-1} (i-t)X''_{t+1}\right) \end{cases}$$

Thus:

$$\frac{Y_i(1+a)+Y_i(1-a)}{2} = \left(\frac{Y_i(1+a)+Y_i(1-a)}{2}\right)$$
$$+ (i-1)\left(\frac{Y_2(1+a)+Y_2(1-a)}{2} - \frac{Y_1(1+a)+Y_1(1-a)}{2}\right)$$
$$+\left(\frac{1+a+1-a}{2}\right)\sum_{t=2}^{i-1}(i-t)X''_{t+1}$$
$$= X_1 + (i-1)(X_2 - X_1) + \left(\sum_{t=2}^{i-1}(i-t)x''_{t+1}\right)$$
$$= X_i$$

References

Armstrong, J. S., & Collopy, F. (1993). Casual forces: structuring knowledge for time series extrapolation. *Journal of Forecasting 12*, 103–115.

Assimakopoulos, V. (1995). A successive filtering technique for identifying long-term trends. *Journal of Forecasting 14*, 35–43.

Clemen, R. (1989). Combining forecasts: a review and annotated bibliography with discussion. *International Journal of Forecasting 5*, 559–584.

Collopy, F., & Armstrong, J. S. (1992). Rule-based forecasting. Development and validation of an expert systems approach to combine time series extrapolations. *Management Science 38*, 1394–1414.

Fildes, R., Hibon, M., Makridakis, S., & Meade, N. (1998). Generalising about univariate forecasting methods: further empirical evidence. *International Journal of Forecasting 14*, 339–358.

Gardner, Jr. S., & McKenzie, E. (1985). Forecasting trends in time series. *Management Science 31*, 1237–1246.

Makridakis, S., Wheelwright, S., & Hyndman, R. (1998). *Forecasting methods and applications*, Wiley, New York.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1984). *The forecasting accuracy of major time series methods*, John Wiley & Sons.

Biographies: Vassilis ASSIMAKOPOULOS is Associate Professor of Forecasting Informational Systems and Director of the Forecasting Systems Unit at the NTUA (National Technical University of Athens) in the ECE (Electrical and Computer Engineering) division. His research interests are in statistics, time series forecasting, neural networks, advanced mathematics and forecasting informational systems. He has published papers in the areas of forecasting, decision systems, statistics and energy modeling. He is the author of the book (in Greek) 'Forecasting Techniques'.

Konstantinos NIKOLOPOULOS is Research Assistant in the Forecasting Systems Unit at the NTUA (National Technical University of Athens) in the ECE (Electrical and Computer Engineering) division. He obtained a diploma in Electrical and Computer Engineering from NTUA. His is in the third year of Ph.D. Studies at NTUA in the research area of intelligent forecasting informational systems. His research interests are in statistics, time series forecasting, neural networks, control systems databases, software engineering and forecasting informational systems. He has published papers in several conference proceedings.