

# Estimation of Missing Values for BL (p, 0, p, p) Time Series Models with Student-t Innovations

Poti Abaja Owili

Mathematics and Computer Science Department, Laikipia University, Nyahururu, Kenya

**Abstract** In this study optimal linear estimators of missing values for bilinear time series models BL (p, 0, p, p) whose innovations have a student-t distribution are derived by minimizing the h-steps-ahead dispersion error. Data used in the study was simulated using the R Statistical Software where 100 samples of size 500 each were generated for the bilinear model BL (1, 0, 1, 1). The time series data generated was numbered from 1 to 500. In each sample, three data positions 48, 293 and 496 were selected at random and the value at these points removed to create artificial missing values. For comparison purposes, two commonly used non-parametric techniques of artificial neural network (ANN) and exponential smoothing (EXP) estimates were also computed. The performance criteria used to ascertain the efficiency of these estimates were the mean squared error (MSE) and Mean Absolute Deviation (MAD). The study found that ANN estimates were the most efficient for estimating missing values of the bilinear time series with student-t innovations. The study recommends the use of ANN for estimating missing values in bilinear time series model with student errors.

**Keywords** ANN, Exponential smoothing, MSE, Performance criterion, Simulation

## 1. Introduction

Data analysts are frequently faced with the missing value problem. Missing values may occur for various reasons which may include poor record keeping, lost records, technical error, collecting data at irregular times, etc., ([1], [2]). In addition, a peculiar case can arise when one may be interested in determining the likely value of a variable of interest at a time that may not coincide with a particular measurement or observation [3]. These can result in one or several observations missing.

These missing values must be accounted for since missing values have negative effects on the modeling of the data [4]. There are many ways of handling missing values. The common approach is to use imputation techniques. This involves using a substitute value to replace the missing observation as in [5]. According to [6], imputation broadly comprises several techniques that have been developed to compute missing values. These techniques may employ strategies such as mean substitution and artificial neural networks approach. It may also involve the use of appropriate statistical prediction or forecasting models such as regression, time series models, and Markov chain and Monte Carlo methods.

Estimation of missing values for bilinear time series has

been done for a particular order of the bilinear time series BL (1, 0, 2, 0) by [4]. They used estimating functions criterion to derive the estimates of missing values. Other studies have also been done to estimate missing values for pure bilinear time series when the innovation sequence has the GARCH distribution [7]. Still the same authors have estimated missing values for pure bilinear time series when the innovation sequence has the normal distribution [8]. [9] also estimated missing values for bilinear time series for the pure bilinear case when the innovation sequence has the student-t distribution. They found that the estimates of the missing values were equivalent to a one-step-ahead forecast. Further, [10] used the linear interpolation criterion to estimate missing values for the BL (p, 0, p, p) when the error term follows the normal distribution.

The distribution of interest in this study is the student-t distribution. This distribution is characterized by long tails and is suitable for modeling financial data which is known to be highly skewed. There is no evidence to show that optimal linear interpolation approach based on the dispersion error has been used to estimate missing values for bilinear BL (p, 0, p, p) with student-t distribution.

### 1.1. Identification of Bilinear Time Series Models

Given a time series data, the first step in the identification process of bilinear time series model is to test whether the data can be modeled either as a linear time series or belongs to the broader class of nonlinear time series models. This involves testing a null hypothesis that the data is linear. This can be done using one of the statistical tests of linearity ([11];

\* Corresponding author:

abajapoti@gmail.com (Poti Abaja Owili)

Published online at <http://journal.sapub.org/statistics>

Copyright © 2016 Scientific & Academic Publishing. All Rights Reserved

[12]). If the null hypothesis is rejected then the data can be appropriately modeled by a nonlinear series model and a bilinear model is one of the candidate models to be considered ([13]; [14]). If the data is nonlinear then the second step is used.

The second step in the identification process is to determine the class of the nonlinear models to which the data belongs. This involves the use of moments and cumulants. It has been noted that BL (p, 0, p, 1) and ARMA (p, 1) models have similar second order moments and hence these moments cannot be conclusively used in the identification of the bilinear time series models [15]. Consequently, it is imperative to use higher order moments. The higher moments are known to satisfy the Yule-Walker type difference equations ([15], [16]). Thus these equations could be used for model identification of the bilinear time series models. The difference between the moments of bilinear time series and those of the other nonlinear time series models is that the higher order moments of a bilinear time series (including the fourth moments) decay slowly as the lag tends to infinity. However, the fourth moments of the other nonlinear time series models do not behave in a similar manner.

After determining that the data is bilinear, then the order of the model is computed using canonical correlation analysis carried between the linear combinations of the observations and linear combinations of higher powers of the observations.

For some super diagonal and diagonal bilinear time series, the third order moments are not equal to zero. This pattern of nonzero moments can be used to discriminate between white noise and the bilinear models and also between different bilinear models [16]. Using the patterns presented in a table of third order moments, one can easily distinguish bilinear models from pure ARMA or mixed ARMA models. Third order moments may also be useful in detecting non-normality in the distribution of the innovation sequence.

This technique of model identification can be extended to more general bilinear models provided that difference equations for higher order moments and cumulants can be obtained [15].

### 1.2. Estimation of Parameters of the Bilinear Time Series Models

Several estimation techniques have been proposed for the estimation of the parameters of the bilinear time series in the literature. Most of them deal with particular classes of the bilinear time series models [9]. [13] proposed two methods for the estimation of the model parameters of a bilinear time series models, namely the use of Newton Raphson technique and the Marquart Algorithm. He applied both methods to the estimation of the parameters of a bilinear time series model identified for sunspot and seismology data. Secondly, he proposed estimation of the parameters using maximum likelihood method. More recently [18] proposed a generalized autoregressive conditional heteroskedasticity-

type maximum likelihood estimator for estimating the unknown parameters for a special bilinear model. They showed that their proposed estimator was consistent and asymptotically normal under only finite fourth moment of errors. [19] proposed the use of covariance estimates based on the least squares method on the parameters of the bilinear model BL (p, 0, p, 1). [20] estimated the parameter of the simple diagonal bilinear model BL (0, 0, 1, 1) using the least squares method.

## 2. Literature Review

### 2.1. Student-t Distributions

Most of the data encountered in practice show departure from the linearity and thus may be modeled by nonlinear time series models [21]. These models have innovations that can adequately be described by student-t, ARCH and stable distributions. For financial data, models with the student-t distribution play an important role in modeling. The student-t distribution can be integrated with other distributions such as GARCH to produce even better models. For example, GARCH (1, 1) model with student-t distribution is able to reproduce the volatility dynamism of financial data. Given a model, specification for log of return disturbances can be modeled using either the student-t distribution or the normal distribution. However, the student-t distribution is particularly useful since it can describe the excess kurtosis in the conditional distribution that is found in financial time series unlike the models with normal innovations (Owili, 2015c).

[22] focused on the bilinear time series model with GARCH innovations (BL-GARCH). It has an important property that it can take into account explosions and related volatility features of non-linear time series. The most common model used in financial data is the BL-GARCH (1,1) model and may be used in practical applications with either the normal, the student-t or GED noises.

### 2.2. Empirical Studies on Bilinear Models

The bilinear time series models find applications in many areas such as hydrology, economics and finance. Since it is a complicated model, only specific classes of the bilinear models have been studied. For example, [13] considered model BL (p, 0, p, q); [11] studied the asymptotic behavior of the correlation function for the simple bilinear model BL (0, 0, 1, 1); [23] and [24] studied the model BL (1, 0, 1, 1); [25] considered the model BL (0, 0, 1, 1). [26] estimated the coefficients of a bilinear model BL (1, 0, 1, 1) using the maximum likelihood method. [24] claimed that estimating bilinear models is quite challenging.

It can be seen from the literature that several studies on inferences based on bilinear time series models have been done. These include model identification, determining conditions necessary for stationarity and invertibility and estimation of the parameters of the bilinear time series

models. On missing value, Owili, Poti and Orawo (2015) only studied pure bilinear model BL (0, 0, p, p). No study has been done on the general bilinear model BL (p, 0, p, p). Therefore, the aim of this study was to derive estimators of missing values for BL (p, 0, p, p) using the optimal linear estimation approach of [28].

### 2.3. Estimation of Missing Values using Linear Interpolation Method

Suppose we have one value  $x_m$  missing out of a set of an arbitrarily large number of  $n$  possible observations generated from a time series process  $\{x_t\}$ . Let the subspace  $S_m^*$  be the allowable space of estimators of  $x_m$  based on the observed values  $\{x_t, x_{t-1}, x_{t-2}, \dots, x_1\}$  i.e.,  $S_m^* = \text{sp} \{x_t : 1 \leq n, t \neq m\}$  where  $n$ , the sample size, is assumed large. The projection of  $x_m$  onto  $S_m^*$  (denoted  $P_{S_m^*}^{x_m}$ ) such that the dispersion error of the estimate (written  $\text{disp}(x_m - P_{S_m^*}^{x_m})$ ) is a minimum would simply be a minimum dispersion linear interpolator. The missing value  $x_m$  is estimated as

$$x_m^* = \hat{x}_m + \sum_{k=m+1}^n a_k (x_k - \hat{x}_k) \tag{1}$$

where  $\hat{x}_m$  is the estimate obtained from the model based on the previous lagged observations of the data before the point  $m$ , the missing data point and  $x_m$  the missing value, the coefficients  $a_k$  ( $k=1, 2, \dots, k-m$ ) are to be estimated by minimizing the dispersion error ( $\text{disp } x_m$ ) given by equation (1) as in [28].

### 3. Research Methodology

Data was simulated from the general bilinear time series models BL (1, 0, 1, 1) with student-t innovations using R statistical software. A program code in R was used. 100 samples of size 500 each were generated and missing artificial points were created at data point 48, 293 and 496. These points were selected at random. Data analysis was done using the following software: Microsoft Excel, TSM, R and Matlab7. The mean squared error (MSE) and mean absolute deviation (MAD) were used as performance measures.

## 4. Results

### 4.1. Derivation of the Missing Values for Bilinear Time Series Models with Student-t Innovations

Estimates of missing values for BL (p, 0, p, p) bilinear time series models whose innovations follow student-t distributions were derived based on minimizing the h-steps ahead dispersion error. Two assumptions were made in the process of the derivations. The first one was that the time series data is stationary. Secondly, the higher powers (of orders greater than two or products of coefficients of orders greater than two) of the coefficients are approximately negligible. This was consequence of the result of the first assumption.

#### 4.1.1. Estimating Missing Values for BL (1, 0, 1, 1) with t- Errors

The bilinear model BL (1, 0, 1, 1) with t- errors is expressed as

$$x_t = \phi x_{t-1} + b_{11} x_{t-1} e_{t-1} + e_t, \text{ where } e_t \sim t(0, 1)$$

The missing value is obtained using theorem 4.1  
Theorem 4.1

The optimal linear estimate for missing value for BL (1, 0, 1, 1) with student errors is given by

$$x_m^* = \hat{\phi} x_{m-1} + \hat{b}_{11} x_{m-1} e_{m-1} + \sum_{k=m+1}^n \frac{\hat{\phi}_1^{k-m}}{\left[ \frac{n}{n-2} (\hat{\phi}_1^2 + 1) + b_{11}^2 v(4) \right]} (x_k - \hat{x}_k)$$

Proof

The stationary BL (1, 0, 1, 1) can be expressed as

$$x_t = \sum_{i=1}^{\infty} \left( \prod_{j=1}^i (\phi + b_{11} e_{t-j}) \right) e_{t-i} + e_t$$

The h-steps ahead forecast is given by

$$x_{t+h} = \sum_{i=1}^{\infty} \left( \prod_{j=1}^i (\phi + b_{11} e_{t+h-j}) \right) e_{t+h-i} + e_{t+h}$$

and the h-steps ahead forecast error is given by

$$x_{t+h} - \hat{x}_{t+h} = \sum_{i=1}^{h-1} \left( \prod_{j=1}^i (\phi + b_{11} e_{t+h-j}) \right) e_{t+h-i} + e_{t+h} \tag{3}$$

Substituting equation (3) in equation (1), we obtain

$$disp x_m = E(\hat{e}_m^2) - 2E \left\{ \hat{e}_m \cdot \left[ \sum_{k=m+1}^n a_k \sum_{i=1}^{k-m} \left( \prod_{j=1}^i (\varphi_1 + b_{11} e_{k-j}) \right) e_{k-i} + \sum_{k=m+1}^n a_k e_k \right] \right\} \\ \left( \sum_{k=m+1}^n a_k \sum_{i=1}^{k-m} \left( \prod_{j=1}^i (\varphi_1 + b_{11} e_{k-j}) \right) e_{k-i} + \sum_{k=m+1}^n a_k e_k \right)^2 \quad (2)$$

Simplifying the terms RHS of equation (2), we obtain

$$E(\hat{e}_m^2) = \frac{n}{n-2} \\ E \left\{ 2\hat{e}_m \cdot \left( \sum_{k=m+1}^n a_k \sum_{i=1}^{k-m} \left( \prod_{j=1}^i (\varphi_1 + b_{11} e_{k-j}) \right) e_{k-i} + \sum_{k=m+1}^n a_k e_k \right) \right\} \\ = E \left\{ 2\hat{e}_m \cdot \left( \sum_{k=m+1}^n a_k \sum_{i=1}^{k-m} \left( \prod_{j=1}^i (\varphi_1 + b_{11} e_{k-j}) \right) e_{k-i} + \sum_{k=m+1}^n a_k e_k \right) \right\} = \sum_{k=m+1}^n a_k b^{k-m} \frac{n}{n-2} \\ E \left( \sum_{k=m+1}^n a_k \sum_{i=1}^{k-m} \left( \prod_{j=1}^i (\varphi + b_{11} e_{k-j}) \right) e_{k-i} + \sum_{k=m+1}^n a_k e_k \right)^2 \approx \left[ \sum_{k=m+1}^n a_k^2 \left( \varphi^2 \frac{n}{n-2} + b_{11}^2 v(4) + \frac{n}{n-2} \right) \right]$$

$$\text{where } v(4) = v(4) = E(e_t^4) = \frac{n^2 \Gamma(\frac{5}{2}) \Gamma(\frac{n-4}{2})}{\Gamma(\frac{n}{2}) \Gamma(\frac{1}{2})}$$

Hence equation (2) becomes

$$disp x_m = \frac{n}{n-2} - 2 \sum_{k=m+1}^n a_k \varphi^{(k-m)} \frac{n}{n-2} + \sum_{k=m+1}^n a_k^2 \left( \varphi^2 \frac{n}{n-2} + b_{11}^2 v(4) + \frac{n}{n-2} \right) \quad (3)$$

Differentiating equation (3) with respect to the coefficients, we get

$$\frac{d}{da_k} \left[ -2 \frac{n}{n-2} \sum_{k=m+1}^{m+1} a_k b^{k-m} + 2 \sum_{k=m+1}^n a_k \left( \varphi^2 \frac{n}{n-2} + b_{11} v(4) + \frac{n}{n-2} \right) \right] = 0 \\ \Rightarrow -2 \frac{n}{n-2} \varphi^{k-m} + 2 a_k \left( \varphi^2 \frac{n}{n-2} + b_{11} v(4) + \frac{n}{n-2} \right) = 0$$

$$\Rightarrow a_k = \frac{\hat{\phi}_1^{k-m}}{\left[ \frac{n}{n-2} (\hat{\phi}_1^2 + 1) + b_{11}^2 v(4) \right]}$$

The optimal linear estimator of  $x_m$ , denoted  $x_m^*$ , that minimizes the error dispersion error is

$$\begin{aligned} x_m^* &= \hat{x}_m + \sum_{k=m+1}^n a_k (x_k - \hat{x}_k) \\ &= \hat{x}_m + \sum_{k=m+1}^n \frac{\hat{\phi}_1^{k-m}}{\left[ \frac{n}{n-2} (\hat{\phi}_1^2 + 1) + b_{11}^2 v(4) \right]} (x_k - \hat{x}_{k1}) \\ &= \hat{\phi} x_{m-1} + \hat{b}_{11} x_{m-1} e_{m-1} \\ &\quad + \sum_{k=m+1}^n \frac{\hat{\phi}_1^{k-m}}{\left[ \frac{n}{n-2} (\hat{\phi}_1^2 + 1) + c^2 v(4) \right]} (x_k - \hat{x}_k) \\ \hat{x}_m &= b_{11} x_{m-1} e_{t-1} + \theta e_{m-1} \end{aligned}$$

4.1.2. Estimating Missing Values for Bilinear Time Series Model BL (p, 0, p, p) with Student-t Innovations

The pure bilinear time series model BL (p, 0, p, p) with student t errors is

$$x_t = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{i=1}^p \sum_{j=1}^p b_{ij} x_{t-i} e_{t-j} + e_t \quad \text{where} \quad e_t \sim t(0,1)$$

The missing values can be estimated using theorem 4.2.

Theorem 4.2

The optimal linear estimate for one missing value  $x_m$  for the general bilinear time series model BL (p, 0, p, p) with student t-errors is given by

$$x_m^* = \sum_{i=1}^p \hat{\phi}_i x_{t-i} + \sum_{i=1}^p \sum_{j=1}^p \hat{b}_{ij} x_{t-i} \hat{e}_{t-j} + \sum_{k=m+1}^n \frac{\frac{n}{n-2} \left\{ \hat{\phi}_s^{k-m} + 2 \sum_{s \neq r}^p \hat{\phi}_s \hat{\phi}_r \right\} \{x_k - x_k\}}{\left\{ \frac{n}{n-2} \left[ \sum_{s=1}^p \hat{\phi}_s^2 + \sum_{j=1, s \neq j}^p \hat{b}_{sj}^2 \right] + \sum_{j=1}^p v(4) \hat{b}_{jj}^2 \right\}}$$

Where  $v(4)$  is the fourth moment of the data given.

$$v(4) = E(e_t^4) = \frac{n^2 \Gamma(\frac{5}{2}) \Gamma(\frac{n-4}{2})}{\Gamma(\frac{n}{2}) \Gamma(\frac{1}{2})}$$

Where  $v(4) = \text{kurtosis} * (\text{variance})^2$ .

Proof

The stationary bilinear time series model BL (p, 0, p, p) is of the form

$$\begin{aligned}
x_t &= \sum_{i=1}^p \sum_{i=1}^{\infty} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t-j} \right\} e_{t-sj} \\
&+ \sum_{s < r}^p \sum_r^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{t-j} \right\} e_{t-s-r} \\
&+ \sum_{r > s}^p \sum_s^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{t-j} \right\} \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t-j} \right\} e_{t-s-r} + e_t
\end{aligned} \tag{4}$$

The h steps ahead forecast based on equation (4) is given by

$$\begin{aligned}
x_{t+h} &= \sum_{i=1}^p \sum_{i=1}^{\infty} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t+h-sj} \\
&+ \sum_{s < r}^p \sum_r^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t+h-s-r} \\
&+ \sum_{r > s}^p \sum_s^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{t+h-j} \right\} \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t+h-s-r} + e_{t+h}
\end{aligned} \tag{5}$$

Or

$$\begin{aligned}
x_k &= \sum_{i=1}^p \sum_{i=1}^{k-1} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-sj} \\
&+ \sum_{s < r}^p \sum_r^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} + \\
&+ \sum_{r > s}^p \sum_s^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{k-j} \right\} \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} + e_t
\end{aligned}$$

The forecast error is

$$\begin{aligned}
x_k - \hat{x}_k &= \sum_{i=1}^p \sum_{i=1}^{k-m} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-sj} \\
&+ \sum_{s < r}^p \sum_r^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} \\
&+ \sum_{r > s}^p \sum_s^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{k-j} \right\} \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} + e_t
\end{aligned} \tag{5}$$

Substituting equation (5) in equation (1) and simplifying, we obtain

$$\begin{aligned}
 & -2E\hat{e}_m \cdot E \sum_{k=m+1}^n a_k \cdot \sum_{i=1}^p \sum_{i=1}^{\infty} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t-sj} \\
 & + -2E\hat{e}_m \cdot \sum_{k=m+1}^n a_k \cdot \sum_{s < r}^p \sum_{r}^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t+h-s-r} \\
 & \cdot \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} \\
 & = -2E\hat{e}_m \cdot \sum_{k=m+1}^n a_k \cdot \sum_{r > s}^p \sum_{s}^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{t+h-j} \right\} \cdot \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} e_{t+h-s-r}
 \end{aligned}$$

First term  $E(\hat{e}_m^2) = \frac{n}{n-2}$

Second term

$$\begin{aligned}
 & -2\hat{e}_m \cdot \sum_{k=m+1}^n a_k \cdot \left\{ \sum_{r > s}^p \sum_{s}^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{t+h-j} \right\} \right\} \\
 & -2E(x_m - \hat{x}_m) \sum_{k=m+1}^n a_k (x_k - \hat{x}_k) \cdot \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{t+h-j} \right\} \{ e_{t+h-s-r} + e_t \} \\
 & = -2E\hat{e}_m \cdot \sum_{k=m+1}^n a_k \sum_{i=1}^p \sum_{i=1}^{k-m} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-si} = -2 \sum_{k=m+1}^n \frac{n}{n-2} \varphi^{(k-m)/s} \\
 & \quad -2\hat{e}_m \cdot \sum_{k=m+1}^n a_k \\
 & \quad \sum_{s < r \text{ and } s > r}^p \sum_{r}^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} \\
 & = \begin{cases} \sum_{k=m+1}^n \frac{4n}{n-2} a_k \varphi_r \varphi_s & \text{for } s+r = k-m \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

Third term

$$E \left\{ \begin{aligned} & \sum_{i=1}^p \sum_{i=1}^{k-m} \prod_{j=1}^i \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{t-sj} \\ & + \sum_{s < r}^p \sum_{r}^p \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} \left\{ \varphi_r + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} \\ & + \sum_{r > s}^p \sum_{s}^p \left\{ \varphi_r + \sum_{j=1}^p b_{rj} e_{k-j} \right\} \left\{ \varphi_s + \sum_{j=1}^p b_{sj} e_{k-j} \right\} e_{k-s-r} + e_t \end{aligned} \right\}^2$$

This simplifies to

$$\begin{aligned} disp x_m = & -2 \frac{n}{n-2} \sum_{k=m+1}^n \varphi_s^{(k-m)/s} + \sum_{k=m+1}^n \frac{n}{n-2} 4a_k \varphi_r \varphi_s \\ & + \sum_{k=m+1}^n a_k^2 \left\{ \begin{aligned} & \sum_{s=1}^p \frac{n}{n-2} \varphi_s^2 + \\ & \sum_{j=1, s \neq j}^p \frac{n}{n-2} b_{sj}^2 + e_{k-j} + \sum_{j=1}^p v(4) b_{jj}^2 \end{aligned} \right\} \end{aligned} \quad (6)$$

where

$$v(4) = E(e_t^4) = \frac{n^2 \Gamma(\frac{5}{2}) \Gamma(\frac{n-4}{2})}{\Gamma(\frac{n}{2}) \Gamma(\frac{1}{2})}$$

Differentiating equation (6) with respect to  $a_k$ , we have

$$\begin{aligned} \frac{d}{da_k} disp = & -2 \frac{n}{n-2} \sum_{k=m+1}^n \varphi_s^{(k-m)/s} + \frac{n}{n-2} \varphi_r \varphi_s + \\ & 2a_k \left\{ \sum_{s=1}^p \frac{n}{n-2} \varphi_s^2 + \sum_{j=1, s \neq j}^p \frac{n}{n-2} b_{sj}^2 + e_{k-j} + \sum_{j=1}^p v(4) b_{jj}^2 \right\} = 0 \end{aligned}$$

Solving for  $a_k$ , we get

$$\hat{a}_k = \frac{\sum_{K=m+1}^n \frac{n}{n-2} \hat{\varphi}_s^{k-m} + 2 \sum_{s \neq r}^p \frac{n}{n-2} \hat{\varphi}_s \hat{\varphi}_r}{\left\{ \frac{n}{n-2} \left[ \sum_{s=1}^p \hat{\varphi}_s^2 + \sum_{j=1, s \neq j}^p \hat{b}_{sj}^2 \right] + \sum_{j=1}^p v(4) \hat{b}_{jj}^2 \right\}}$$

Corollary

For  $p=1$ , we have the bilinear model BL (1, 0, 1, 1). The best linear estimate is given by

$$x_m^* = \hat{\varphi}_1 x_{m-1} + \hat{b}_{11} x_{m-1} \hat{e}_{m-1} + \sum_{k=m+1}^n \frac{\hat{\varphi}^{(k-m)}}{(1 + \hat{\varphi}_1^2 + v(4) \hat{b}_{11}^2)} (x_k - \hat{x}_k)$$

## 4.2. Simulation Results

In this section, the results of the optimal linear estimator, artificial neural networks and exponential are given in table 1.

**Table 1.** Efficiency Measures for BL (1, 0, 1, 1) with student-t innovations

POINT	MAD			MSE		
	OLE	ANN	EXP	OLE	ANN	EXP
48	1.05	1.023	1.921	2.55	2.1023	2.889
293	1.35	1.073	1.43	2.73	2.4562	3.781
496	1.44	0.9	1.47	3.84	2.139	4.623
Total	3.85	2.996	4.821	9.13	6.5975	11.23
Mean	1.28	0.999	1.607	3.04	2.1992	3.743

From table 1, it can be concluded that ANN estimates most efficient (MSE=2.1992) for the different missing data point positions followed by OLE estimates (MSE=3.04). This is in contradiction to the results obtained for pure bilinear time series model BL (0, 0, 1, 1) by [9].

## 5. Conclusions

In this study we have derived estimates for missing values for the bilinear time series model BL (p, 0, p, p) with student-t innovations. The study found that ANN estimates were the most efficient compared to both the OLE and EXP. Further the estimates of the missing values were found to be dependent on the observations before and after the missing value point.

## Appendix

Appendix A: Program Codes used in simulation

```
# The R program (BL1011_studentErrors)
b12<-0.4; b1=0.2
sigma<-1
h<-c()
e<-c()
z<-c()
e[1:2]<- rt(2,7)
x <-c()
x[1:2]<-c(0, e[1])
set.seed(02848151)
for (i in 3:1500) {
  e[i] = rt(1,7)          # generate noise value
  x[i] = b1*x[i-1]+b12*x[i-1]*e[i-2]+ e[i] # calculate x using the model
}
t<- x[-1:-1000]
y<-round(t,7)
z<-0.1+0.8*(y-min(y))/(max(y)-min(y))
n<-round(z,7)
Y
```

---

## REFERENCES

- [1] Pigot, D., T., 2001, A Review of Methods for Missing Data. Educational Research and Evaluation Vol. 7, No. 4, pp. 353±383.
- [2] Fung, D. S. C., 2006, Methods for the Estimation of Missing Values in Time Series <http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1063&context=theses> retrieved on 11/2/2006.
- [3] Musial, J. P., Verstraete1, M. M., and Gobron N., 2011, Technical Note: Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series.
- [4] Abraham, B. and Thavaneswaran, A., 1991, A Nonlinear Time Series and Estimation of missing observations. Ann. Inst. Statist. Math. Vol. 43, 493-504.
- [5] Jones, R. H. 1980, Maximum likelihood fitting of ARMA models to time series with missing observations. Technometrics, 22, 389 -395.
- [6] Abrahantes, J. C., Sotto, C., Molenberghs, G., Vromman, G., and Bierinckx, B. (2011). A comparison of various software tools for dealing with missing data via imputation .J. Stat. Comput. Simulation 81, No. 11, 1653-1675. Am Stat. 2007 Feb; 61(1): 79–90.
- [7] Owili, P. A., Nassiuma, D., and Orawo, L., 2015b, Efficiency of Imputation Techniques for Missing Values of Pure Bilinear Models with GARCH Innovations, American Journal of Mathematics and Statistics, Vol. 5 No. 5. pp. 316-324. doi:10.5923/j.ajms.20150505.13.
- [8] Owili, P. A., Nassiuma, D., and Orawo, L., 2015a, Imputation of Missing Values for Pure Bilinear Time Series Models with Normally Distributed Innovations." American Journal of Applied Mathematics and Statistics 3.5: 199-205.
- [9] Owili, P. A., Nassiuma, D., and Orawo, L., 2015c, Estimation of Missing Values for Pure Bilinear Time Series Models with Student-t Innovations, International Journal of Statistics and Applications, Vol. 5 No. 6, 2015, pp. 293-301. doi: 10.5923/j.statistics.20150506.05.
- [10] Owili, P. A., 2015, Imputation of Missing Values for BL (p,0,p,p) Models with Normally Distributed innovations. Science Journal of Applied Mathematics and Statistics.Vol.3, No. 6, 2015, pp. 234-242. doi: 10.11648/j.sjams.20150306.12
- [11] Keenan, D. M., 1985, "A Tukey Non-additivity Type Test for Time Series Nonlinearities." &&&
- [12] Tsay, R. S. 1986, Time series model specification in the presence of outliers. Journal of the American USA: Springer-Verlag Vector Time Series. Biometrika, 84(2), 495-499.
- [13] Subba Rao T., 1981, On the Theory of bilinear time series models. J. Roy. Staist.Ser. B 43 244-255.
- [14] Subba, Rao, T.; Gabr, M.M. (1984). An Introduction to Bispectral Analysis and Bilinear Time Series Models. *Lecture notes in statistics, 24*.New York. Springer.
- [15] Sesay, S. A and Subba Rao, T., 1988., Yule Walker type difference equations for higher order moments and cumulants for bilinear time series models. J. Time Ser. Anal.9, 385-401
- [16] Sesay, S. A and Subba Rao, T., 1991., Difference equations for higher order moments and cumulants for the bilinear time series model [19] BL(p,0,p,1]. J. Time Ser. Anal.12, 159-177.
- [17] Kumar, K., 1986, On The Identification o Some Bilinear Time Series Model. Journal of Time Series Analysis Volume 7, Issue 2, Pages 117–122.
- [18] Shiqing, L. Liang, P. and Fukang Zhu, 2015, Inference For A Special Bilinear Time-Series Model. Journal of Time Series Analysis J. Time. Ser. Anal.36:61–66.
- [19] Mathews, V. J. & Moon, T. K., 1991, Parameter estimation for a bilinear time series model. In *Proceedings - ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing.* (Vol. 5, pp. 3513-3516). Piscataway, NJ, United States: Publ by IEEE.
- [20] Wu, W. B and Min, W., 2004, On linear processes with dependent innovations. Technical Report, Department of Statistics, University of Chicago.
- [21] Wu, W. B. and Min, W., 2004, On linear processes with dependent innovations. Technical Report, Department of Statistics, University of Chicago. 0
- [22] Diongue, A. K, Dominique Guegan, D. Wol, R. C., 2009, Exact Maximum Likelihood estimation for the BL-GARCH model under elliptical distributed innovations. Documents de travail. Journal of Statistical Computation and Simulation. Vol. 00, No. 0, 1–17.
- [23] Turkman, K. F. and Turkman, M. A. A., 1997, Extremes of bilinear time series models. Journal of Time Series Analysis 18: 305–319. variances and exact likelihood equations. J. of Time Series Analysis, 24, 739-754.
- [24] Basrak, B. Davis, R. A., Mikosch T., 1999, The sample ACF of a simple bilinear process. Stochastic Processes and their Applications 83: 1–14.
- [25] Zhang, Z. and Tong, H., 2001, on some distributional properties of a first-order nonnegative bilinear time series model. Journal of Applied Probability 38: 659–671.
- [26] Hristova, D., 2004, Maximum Likelihood Estimation of a Unit Root Bilinear Model with an Application to Prices.
- [27] Ling, S, Peng, L. and Zhu, F., 2015, Inference for a special bilinear time-series model. Journal of time series analysis J. time. ser. anal. 36: 61–66).
- [28] Nassiuma, D. K., 1994, Symmetric stable sequence with missing observations. J.T.S.A. volume 15, page 317.