Journal of Fiber Bioengineering and Informatics 7:1 (2014) 53–65
doi: 10.3993/jfbi03201405

Time-Varying Moving Average Model for Autocovariance Nonstationary Time Series

Wanchun Fei [a,*], Lun Bai [b]

[a] *College of Textile and Clothing Engineering, Soochow University, Suzhou 215021, China*
[b] *National Engineering Laboratory for Modern Silk, Suzhou 215123, China*

Abstract

In time series analysis, fitting the Moving Average (MA) model is more complicated than Autoregressive (AR) models because the error terms are not observable. This means that iterative nonlinear fitting procedures need to be used in place of linear least squares. In this paper, Time-Varying Moving Average (TVMA) models are proposed for an autocovariance nonstationary time series. Through statistical analysis, the parameter estimates of the MA models demonstrate high statistical efficiency. The Akaike Information Criterion (AIC) analyses and the simulations by the TVMA models were carried out. The suggestion about the TVMA model selection is given at the end. This research is useful for analyzing an autocovariance nonstationary time series in theoretical and practical fields.

Keywords: MA Model; Autocovariance; Parameter Estimation; Simulation; Model Selection

1 Introduction

For the past decades, time series analysis has become a highly developed subject, and there are now well-established methods for fitting a wide range of models to time series data as in the books [1-4] and in the articles [5-8]. However, none of these studies focused on autocovariance nonstationary time series. Virtually all the established methods rest on one fundamental assumption, namely, that the process is *autocovariance* stationary, or locally stationary. At least the statistical characteristics of the nonstationary processes class are changing *smoothly* over time. The nonstationarity of the time series means at least one statistical characteristic is variant with time points. The mean (or trend) nonstationary time series can usually be reduced to mean stationary by some simple transformation, such as the difference method. The autocovariance is one of the most important statistical characteristics of the nonstationary time series [9] and can be used as an important index when evaluating simulation by some time series analysis models. Needless to say, for a single time series sampled from some time series, the assumption of autocovariance stationarity is a mathematical idealization which, in some cases, may be valid only as an approximation to the real situation. In practical applications, the most one could hope for is, over the

* Corresponding author.
Email address: feiwch@suda.edu.cn (Wanchun Fei).

observed time interval, the series would not depart *too far* from autocovariance stationarity for the results of the analysis to be valid.

It is not difficult to see why the notion of autocovariance stationarity is such an appealing one. It endows the process with *statistical stability* so that autocovariances originally defined as ensemble averages can be estimated from a single realization by computing the corresponding time domain averages. (Strictly speaking, the property to which we are referring here is *ergodicity* rather than stationarity, but in practical terms the two ideas are very closely related.) However, like virtually all mathematical concepts, stationarity is an idealization, and in practice, it can at best be expected to hold only as an approximation. In order to carry out meaningful statistical analysis on a single time series, the time series is usually considered to be in some sense locally stationary although globally nonstationary [3, 5, 8]. It is clearly of interest, therefore, to examine if any type of the analysis is available for those cases where the assumption of autocovariance nonstationarity becomes realistic.

Usually, it is hard to statistically analyze a nonstationary time series for its complexity. Nonstationary models offer greater complexity than stationary models, and the statistical problems of model identification and parameter estimation are similarly more intricate. Experience gained so far has shown that even simple nonstationary models can capture examples of time series behavior which would be impossible to describe with stationary models. For a zero mean autocovariance nonstationary time series with finite length, the full order Time-varying Parameter Autoregressive (TVPAR) model with time varying AR coefficients and variances of residuals [10] was used for simulation [11] and pattern recognition [12]. The term TVPAR model was adopted to differentiate from the TVAR model in which the variance of residual is time invariant [13], at least in a time segment. There are three TVPAR models: full order TVPAR (TVPAR(T)), time-unvarying order TVPAR (TVPAR(p)), and time-varying order TVPAR (TVPAR(p_t)) [14]. In this paper, three time-varying moving average (TVMA) models are presented and analyzed by means of AIC [15], and TVMA model selection is suggested. The AIC is a measure of the relative quality of a statistical model, for a given set of data. As such, AIC provides a means for model selection. In the general case, the AIC is AIC=$2k - 2\ln(L)$, where k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model. We believe that although these models are analyzed by AIC, the result will still remain valid if analyzed by other information criterion [16]. The AIC is used only as an information criterion and is not compared with other criterion in this study.

In time series analysis, the Moving-average (MA) model is a common approach for modeling univariate stationary time series. Time series can be expressed to be generated by passing white noises through a non-recursive linear filter. The notation MA(q) refers to the moving average model of order q:

$$X_t = \mu_t + \sum_{k=0}^{q} \theta_k \varepsilon_{t-k},$$

where μ_t is the mean of the series at the time point t, the θ_k are the parameters of the model, and the ε_{t-k} are white noise error terms, $\varepsilon_t \sim N(0, 1)$. The value of q is called the order of the MA model. We can assume, without loss of generality, that μ_t is zero at each time point t.

A moving average model is conceptually a linear regression of the current value of the series against previous (unobserved) white noise error terms or random shocks. In stationary time series analysis, the random shocks at each point are assumed to be mutually independent and to come from the same distribution, typically a normal distribution, with location at zero and constant scale. The distinction in this model is that these random shocks are propagated to future values of the time series. Fitting the MA estimates is more complicated than with autoregressive models, e.g. $X_t = \sum_{j=1}^{p} \varphi_j X_{t-j} + \varepsilon_t$, because the lagged error terms are not observable while X_t and X_{t-j} are observable. This means that iterative nonlinear fitting procedures need to be used in place of linear least squares. It is meaningful to estimate the parameters of the MA models since MA models have a less obvious interpretation than AR models. Sometimes the autocorrelation function (ACF) and partial autocorrelation function (PACF) will suggest that a MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model [1]. The moving average model is essentially a finite impulse response filter applied to white noise, with some additional interpretation placed on it.

Fitting the MA estimates is so complicated that the MA model is rarely analyzed and applied in practical engineering. The most general problem of time series model building may be stated in the following terms: *given some observations sampled from some stochastic process, find the function which reduces the observations to a strict white noise process.* The main purpose of this paper is to present a novel type of MA model for autocovariance nonstationary time series. Fortunately, after statistical analysis and simulation, the resulting estimates of the MA model demonstrate statistical efficiency.

This paper focuses on TVMA models for autocovariance nonstationary time series which are repeatedly sampled as size series of cocoon filament from same category cocoons. In this study, the assumption of autocovariance stationarity is abandoned. The TVMA models can be used to analyse and simulate autocovariance nonstationary time series including the size series of cocoon filament.

2 The Time Series and the TVMA Models

Autocovariance nonstationary time series is expressed as follows:

For time series data $\{z_{i,t}\}$ repeatedly sampled from some nonstationary linear stochastic process, where $i = 1, 2, \cdots, n$, n is the series number repeatedly sampled, which is great enough in statistical significance, $t = 0, 1, \cdots, T$, T is a positive integer. The series is called an autocovariance nonstationary time series if it satisfies the following:

A. The average at every time point, $\mu_t = E\{z_{\cdot,t}\} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} z_{i,t}$ is finite;

B. The autocovariance, $c_{t,\tau} = \mathrm{cov}(z_{\cdot,t}, z_{\cdot,\tau}) = E\{(z_{\cdot,t} - \mu_t)(z_{\cdot,t} - \mu_\tau)\} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (z_{i,t} - \mu_t)(z_{i,t} - \mu_\tau)$ is finite, where t, $\tau \in \{0, 1, \cdots, T\}$, meanwhile, the autocovariance $c_{t,\tau}$ depends not only on the separation between the time points t and τ, but also on their individual locations.

If the average at every time point μ_t is equal to zero, the time series described above is called zero mean autocovariance nonstationary time series.

Fig. 1 shows a real example of a raw autocovariance nonstationary time series and a preprocessed series. (I) The upper left portion shows the raw data of filament size (the thickness of a segment of cocoon filament) series sampled from some category cocoon [11], measured in dtex, $\{x_{i,t} | i = 1, 2, \cdots, n, t = 0, 1, \cdots, T_i\}$. (II) The bottom left portion shows the preprocessed length unified (to the mean length $T + 1$) size series of the cocoon filaments, $\{y_{i,t} | t = 0, 1, \cdots, T\}$. (III) The upper right and the bottom right portions show the deterministic (trend) component $\{d_{i,t}\}$ and the stochastic component $\{z_{i,t}\}$ of the unified size series respectively. Obviously $\{y_{i,t}\} = \{d_{i,t}\} + \{z_{i,t}\}$. Fig. 2 shows the autocovariance $c_{t,\tau}$ of the time series both shown in the bottom of Fig. 1.

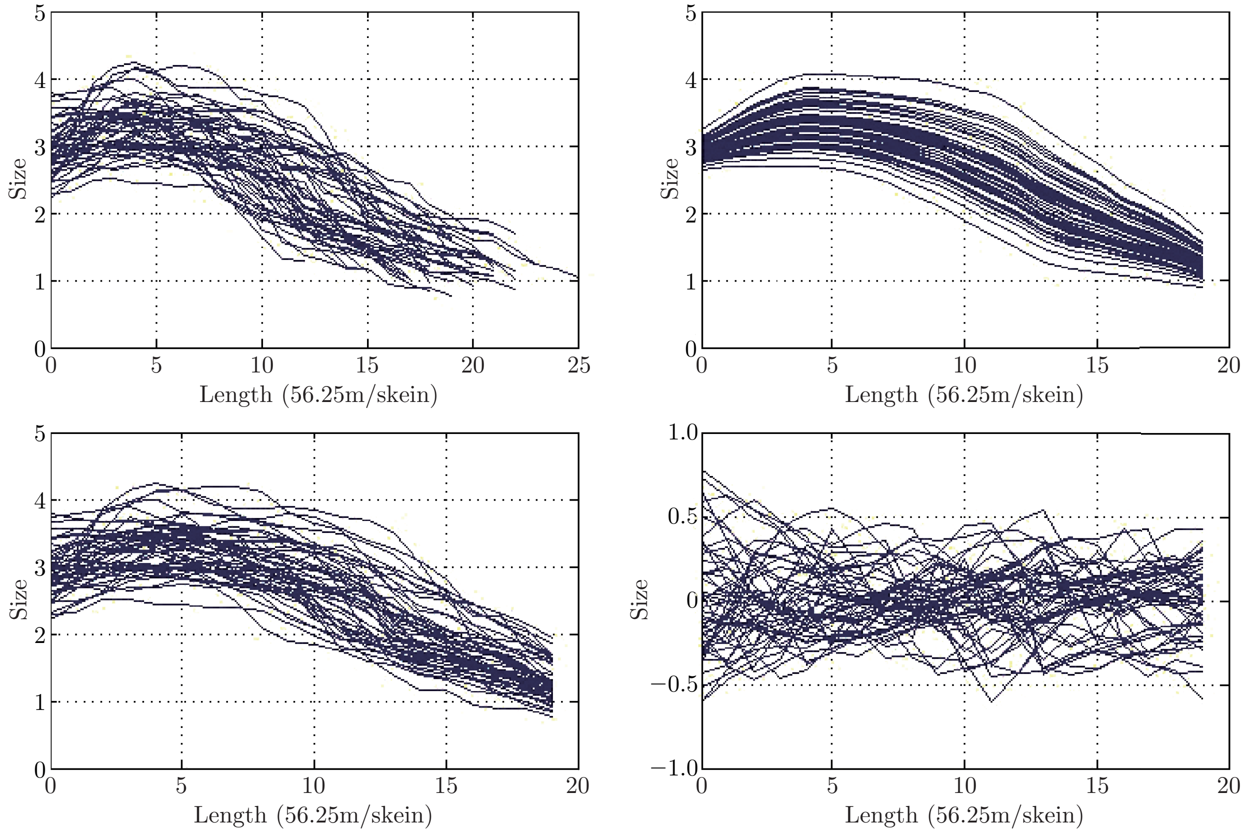

Fig. 1: A real example of an autocovariance nonstationary time series

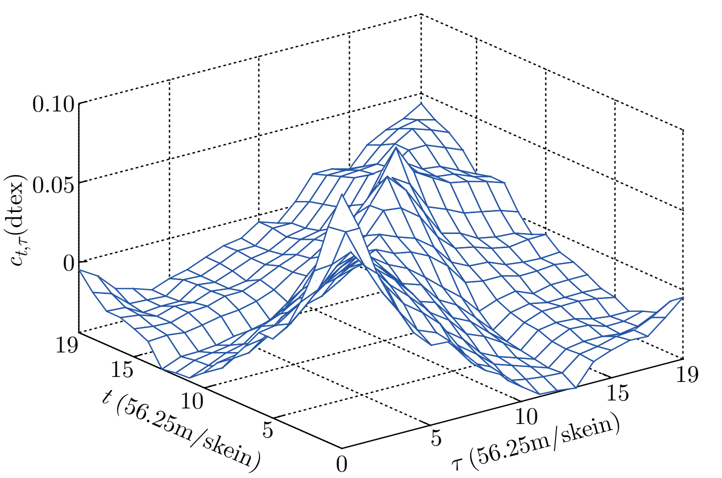

Fig. 2: The autocovariance $c_{t,\tau}$ of the process shown in the bottom of Fig. 1

The time series data $\{z_{i,t}\}$ with the same physical meaning are sampled n times from the same zero mean autocovariance nonstationary process. This differs from the panel series data which usually has various meanings which refer to multi-dimensional data sampled from some entity only one time over multiple time periods.

To analyze a zero mean autocovariance nonstationary time series $\{z_{i,t}\}$, the following three TVMA models are presented now:

A. Full order TVMA model, TVMA(T):

$$z_{i,t} = \sum_{k=0}^{t} \theta_{k,t}\varepsilon_{i,t-k} \tag{1}$$

B. Time-unvarying order TVMA model, TVMA(q):

$$z_{i,t} = \sum_{k=0}^{\eta} \theta_{k,t}\varepsilon_{i,t-k} \tag{2}$$

where $\eta = \min(t, q)$.

C. Time-varying order TVMA model, TVMA(q_t):

$$z_{i,t} = \sum_{k=0}^{q_t} \theta_{k,t}\varepsilon_{i,t-k} \tag{3}$$

where $q_t \leq t$.

The common points of the above three TVMA models are: their parameters $\theta_{k,t}$ (including the variances $\theta_{0,t}^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(z_{i,t} - \sum_{k=1}^{\tau}\theta_{k,t}\varepsilon_{i,t-k})^2$, where $\tau = t$, η or q_t) are time-varying on the time points, do not have to change smoothly over time; and residuals $\varepsilon_{i,t}$ unrelated to previous time series value(s) $z_{i,t}$ are independent normal random variables at all time points. Series values at a present time point can be expressed as the addition of two parts, one linearly dependent on the residual(s) of previous time point(s) and the other which is not. The correlation between values at different time points is, of course, zero, and we may therefore write:

$$E\{z_{\cdot,t}\} = 0, \mathrm{var}\{\varepsilon_{\cdot,t}\} = 1, \text{ every } t;$$
$$\mathrm{cov}\{\varepsilon_{\cdot,t}, \varepsilon_{\cdot,\tau}\} = 0, \text{ all } t \neq \tau.$$

The only difference among the above three TVMA models is the model order. The order of the full order TVMA model is always the time point t; the order of the time-unvarying order TVMA model is a fixed lag unvarying with the time point except at the beginning of the time series; and the order of the time-varying order TVMA model is usually variable with the time point.

It had been proved [10] that the autocovariances of the time series simulated by a full order time-varying parameter autoregressive model are equal to those of the original time series. It is not difficult to prove that the autocovariances of the time series simulated by the TVMA(T) model are equal to those of the original time series.

The TVMA(q) model will become the TVMA(T) model if the order q equals T. Virtually, the order q is selected between 1 and $T/2$ while analyzing the longer time series. The TVMA(q_t) model will become the TVMA(q) model if q_t happens to be invariant with the time point after the beginning of the time series.

3 Model Parameters and AIC Analysis

3.1 For TVMA(q)

For the time-unvarying order TVMA model expressed in Equation (2), q is the time-unvarying order usually less than $T/2$ while the model is used to analyze the longer time series. The residual $\varepsilon_{i,t}$ is a white noise, $\varepsilon_{i,t} \sim N(0,1)$. Obviously, $\hat{\theta}_{0,0}^2 = c_{0,0}, \varepsilon_{i,0} = z_{i,t}/\hat{\theta}_{0,0}$. Considering the independence between $\varepsilon_{i,t}$ and $\varepsilon_{i,\tau}(t \neq \tau)$, where $\varepsilon_{i,t}$ is unrelated with previous time series value(s) $z_{i,t-k}$, according to sampled time series data $z_{i,t}$ and Equation (2), when $t = 1$, we can obtain:

$$c_{0,1} = \theta_{0,0}\theta_{1,1}, \quad c_{1,1} = \theta_{0,1}^2 + \theta_{1,1}^2.$$

where $c_{0,0}, c_{0,1}$ and $c_{1,1}$ are known items through calculation on sampled time series data. After finishing,

$$\hat{\theta}_{1,1} = c_{0,1}/\hat{\theta}_{0,0}, \quad \hat{\theta}_{0,1}^2 = c_{1,1} - \hat{\theta}_{1,1}^2.$$

Similarly, increasing t progressively, $j = \eta, \eta-1, \cdots, 1, 0$, we can obtain the estimates of MA parameters $\hat{\theta}_{k,t}$ and $\hat{\theta}_{0,t}^2$ in Equation (2) by use of the following iterative Equation (4) and Equation (5 or 5′).

$$\hat{\theta}_{j,t} = \left(c_{t-j,t} - \sum_{k=1}^{\eta-j} \hat{\theta}_{k,t-j}\hat{\theta}_{j+k,t} \right) \Big/ \hat{\theta}_{0,t-j}, \tag{4}$$

$$\hat{\theta}_{0,t}^2 = c_{t,t} - \sum_{k=1}^{\eta} \hat{\theta}_{k,t}^2, \quad \text{if } c_{t,t} > \sum_{k=1}^{\eta} \hat{\theta}_{k,t}^2, \tag{5}$$

otherwise,

$$\hat{\theta}_{0,t}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(z_{i,t} - \sum_{k=1}^{\eta} \hat{\theta}_{k,t}\varepsilon_{i,t-k} \right)^2. \tag{5'}$$

The likelihood function of $z_{i,0}, \cdots, z_{i,T}$ will be:

$$L(\theta_{0,t}, \theta_{1,t}, \cdots, \theta_{\eta,t}) = \prod_{t=0}^{T} f(z_{i,t}|\varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-\eta})$$

where $f(z_{i,t}|\varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-\eta})$ is the normal density function with the mean $\sum\limits_{k=1}^{\eta} \theta_{k,t-j}\varepsilon_{i,t-k}$ and variance $\theta_{0,t}^2$, expressed as:

$$f(z_{i,t}|\varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-\eta}) = \frac{1}{\sqrt{2\pi\theta_{0,t}^2}} \exp\left\{ -\frac{1}{2\theta_{0,t}^2} \left(z_{i,t} - \sum_{k=1}^{\eta} \theta_{k,t}\varepsilon_{i,t-k} \right)^2 \right\}.$$

Therefore, the likelihood function of the TVMA(q) model on the ith series is:

$$L_i(\{\theta_{k,t}\},\theta_{0,t}^2)=\left(\frac{1}{2\pi}\right)^{(T+1)/2}\prod_{t=0}^{T}\frac{1}{\sqrt{\theta_{0,t}^2}}\exp\left\{-\frac{1}{2}\sum_{t=0}^{T}\frac{1}{\theta_{0,t}^2}\left(z_{i,t}-\sum_{k=1}^{\eta}\theta_{k,t}\varepsilon_{i,t-k}\right)^2\right\}.$$

The log-likelihood function of the TVMA(q) model on the ith series and all the n series are respectively given by the following two equations:

$$l_i(\{\theta_{k,t}\},\theta_{0,t}^2)=-\frac{T+1}{2}\log 2\pi-\frac{1}{2}\sum_{t=0}^{T}\log\theta_{0,t}^2-\frac{1}{2}\sum_{t=0}^{T}\frac{1}{\theta_{0,t}^2}\left(z_{i,t}-\sum_{k=1}^{\eta}\theta_{k,t}\varepsilon_{i,t-k}\right)^2, \quad (6)$$

$$l(\{\theta_{k,t}\},\theta_{0,t}^2)=-\frac{n(T+1)}{2}\log 2\pi-\frac{n}{2}\sum_{t=0}^{T}\log\theta_{0,t}^2-\frac{1}{2}\sum_{t=0}^{T}\frac{1}{\theta_{0,t}^2}\sum_{i=1}^{n}\left(z_{i,t}-\sum_{k=1}^{\eta}\theta_{k,t}\varepsilon_{i,t-k}\right)^2, \quad (7)$$

Substituting parameters $\hat{\theta}_{k,t}$ and $\hat{\theta}_{0,t}^2$ for Equation (7), the maximum log-likelihood value can be expressed as:

$$l(\{\hat{\theta}_{k,t}\},\hat{\theta}_{0,t}^2)=-\frac{n(T+1)}{2}(\log 2\pi+1)-\frac{n}{2}\sum_{t=0}^{T}\log\hat{\theta}_{0,t}^2$$

There are $T+1+q(2T+1-q)/2$ parameters in the TVMA(q) model. The AIC(q) of the TVMA(q) model can be obtained by:

$$\mathrm{AIC}(q)=-2l(\{\hat{\theta}_{k,t}\},\hat{\theta}_{0,t}^2)+2(T+1)+q(2T+1-q)=\kappa+n\sum_{t=0}^{T}\log\hat{\theta}_{0,t}^2+q(2T+1-q) \quad (8)$$

where $\kappa=n(T+1)(\log 2\pi+1)+2(T+1)$. The value of q yielding the minimum AIC specifies the best TVMA(q) model.

If series with longer length ($T\gg q$) are stationary and ergodic and satisfy $\theta_{0,0}^2=\theta_{0,1}^2=\cdots=\theta_{0,T}^2=\hat{\sigma}^2$ for any given single series in this series cluster, the number of parameters in this MA model can be considered to be $q+1$. Hence,

$$\mathrm{AIC}(q)=(T+1)(\log 2\pi+1)+(T+1)\log\hat{\sigma}^2+2(q+1).$$

Ignoring the constant $(T+1)(\log 2\pi+1)$, the above equation can be theoretically used as the AIC formula for a single autocovariance stationary time series.

3.2 For TVMA(T)

For a zero-mean autocovariance nonstationary time series, if the order q of the TVMA(q) model becomes T, the TVMA(q) model will become the TVMA(T) model. Therefore, the estimates of the MA parameter $\hat{\theta}_{k,t}$ and variances $\hat{\theta}_{0,t}^2$ of residual $\varepsilon_{i,t}$ in Equation (1) can also be obtained. The AIC(T) of the TVMA(T) model can be calculated through the following equation:

$$\mathrm{AIC}(T)=-2l(\{\hat{\theta}_{k,t}\},\hat{\theta}_{0,t}^2)+2(T+1)+T(2T+1-T)=\kappa+n\sum_{t=0}^{T}\log\hat{\theta}_{0,t}^2+T(T+1) \quad (9)$$

where $\kappa=n(T+1)(\log 2\pi+1)+2(T+1)$.

3.3 For TVMA(q_t)

For the time-varying order TVMA model expressed in Equation (3), q_t is the time-varying order usually less than $T/2$ where the model is used to analyze the longer time series. The residual $\varepsilon_{i,t}$ is a white noise, $\varepsilon_{i,t} \sim N(0,1)$. Obviously, $\hat{\theta}^2_{0,0} = c_{0,0}$. Considering the independence between $\varepsilon_{i,t}$ and $\varepsilon_{i,t}(t \neq \tau)$, where $\varepsilon_{i,t}$ is unrelated with previous time series value(s) $z_{i,t-k}$, let $t = 1, 2, \cdots, T, j = q_t, q_t - 1, \cdots, 1, 0$. We can obtain the estimates of the MA parameter $\hat{\theta}_{k,t}$ and variance $\hat{\theta}^2_{0,t}$ of residual $\varepsilon_{i,t}$ in Equation (3) by use of Equation (10) and Equation (11 or 11′).

$$\hat{\theta}_{j,t} = \left(c_{t-j,t} - \sum_{k=0}^{q_t - j} \hat{\theta}_{k,t-j} \hat{\theta}_{j+k,t} \right) \Big/ \hat{\theta}_{0,t-j} \tag{10}$$

$$\hat{\theta}^2_{0,t} = c_{t,t} - \sum_{k=1}^{q_t} \hat{\theta}^2_{k,t} \quad \text{if } c_{t,t} > \sum_{k=1}^{q_t} \hat{\theta}^2_{k,t}, \tag{11}$$

otherwise,

$$\hat{\theta}^2_{0,t} = \frac{1}{n} \sum_{i=1}^{n} \left(z_{i,t} - \sum_{k=1}^{q_t} \hat{\theta}_{k,t} \varepsilon_{i,t-k} \right)^2. \tag{11'}$$

The likelihood function of $z_{i,0}, \cdots, z_{i,T}$ will be:

$$L(\theta_{0,t}, \theta_{1,t}, \cdots, \theta_{q_t,t}) = \prod_{t=0}^{T} f(z_{i,t} | \varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-q_t})$$

where $f(z_{i,t} | \varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-q_t})$ is the normal density function with the mean $\sum_{k=1}^{q_t} \theta_{k,t-j} \varepsilon_{i,t-k}$ and variance $\theta^2_{0,t}$, expressed as:

$$f(z_{i,t} | \varepsilon_{i,t-1}, \cdots, \varepsilon_{i,t-q_t}) = \frac{1}{\sqrt{2\pi\theta^2_{0,t}}} \exp\left\{ -\frac{1}{2\theta^2_{0,t}} \left(z_{i,t} - \sum_{k=1}^{q_t} \theta_{k,t} \varepsilon_{i,t-k} \right)^2 \right\}.$$

Therefore, the likelihood function of the TVMA(q_t) model on all n series at time point t will be:

$$L_t(\{\theta_{k,t}\}, \theta^2_{0,t}) = \left(\frac{1}{\sqrt{2\pi\theta^2_{0,t}}} \right)^n \exp\left\{ -\frac{1}{2\theta^2_{0,t}} \sum_{i=1}^{n} \left(z_{i,t} - \sum_{k=1}^{q_t} \theta_{k,t} \varepsilon_{i,t-k} \right)^2 \right\}.$$

The log-likelihood function is expressed as the following:

$$l_t(\{\theta_{k,t}\}, \theta^2_{0,t}) = -\frac{n}{2} \log 2\pi\theta^2_{0,t} - \frac{1}{2\theta^2_{0,t}} \sum_{i=1}^{n} \left(z_{i,t} - \sum_{k=1}^{q_t} \theta_{k,t} \varepsilon_{i,t-k} \right)^2 \tag{12}$$

On substituting parameters $\hat{\theta}_{k,t}$ and $\hat{\theta}^2_{0,t}$ for Equation (12), the maximum log-likelihood value can be expressed as:

$$l_t(\{\theta_{k,t}\}, \theta^2_{0,t}) = -\frac{n}{2}(\log 2\pi + 1) - \frac{n}{2} \log \hat{\theta}^2_{0,t}$$

There are $q_t + 1$ parameters in the TVMA(q_t) model at time point t. The AIC(q_t) of the TVMA(q_t) model can be obtained by:

$$\mathrm{AIC}_t(q_t) = -2l(\{\hat{\theta}_{k,t}\}, \hat{\theta}^2_{0,t}) + 2(q_t + 1) = n(\log 2\pi + 1) + n \log \hat{\theta}^2_{0,t} + 2(q_t + 1).$$

The value of q_t yielding the minimum $\mathrm{AIC}_t(q_t)$ specifies the best TVMA(q_t) model at time point t.

The total maximum log-likelihood value at all the time points ($t = 0, 1, \cdots, T$) of the TVMA(q_t) model can be expressed as follows:

$$l(\{\hat{\theta}_{k,t}\}, \hat{\theta}^2_{0,t}) = -\frac{n(T+1)}{2}(\log 2\pi + 1) - \frac{n}{2}\sum_{t=0}^{T} \log \hat{\theta}^2_{0,t}$$

There are $\sum_{t=1}^{T} q_t + T + 1$ parameters at all the time points in the TVMA(q_t) model. The AIC(q_t) of the TVMA(q_t) model at the all time points can be obtained by Equation (13):

$$\mathrm{AIC}(q_t) = -2l(\{\hat{\theta}_{k,t}\}, \hat{\theta}^2_{0,t}) + 2\sum_{t=1}^{T} q_t + 2(T+1) = \kappa + n\sum_{t=0}^{T} \log \hat{\theta}^2_{0,t} + 2\sum_{t=1}^{T} q_t \quad (13)$$

where $\kappa = n(T + 1)(\log 2\pi + 1) + 2(T + 1)$. If q_t could be selected up to t, even if $t > T/2$, the AIC(q_t) of the TVMA(q_t) model should be the minimum among the AIC values for the three TVMA models proposed above.

4 Simulation

A real, original zero mean nonstationary time series, shown at the right bottom in Fig. 1, has been analyzed through the above three TVMA models, and 1500 time series have been simulated through the three TVMA models respectively. The experimental results verified that the autocovariances of time series simulated by the TVMA(T) model are almost completely equal to those of the original time series as shown in Fig. 2, and therefore, the TVMA(T) model is appropriate for describing this type of time series with finite length.

AIC analysis is carried out for TVMA(q) and TVMA(T). The AIC values are shown in Table 1. If the order of the TVMA(q) model can only be selected from 1 to $T/2$, we suggest that TVMA(6) should be selected. Otherwise, if the order of the TVMA(q) model can be selected from 1 to $T(= 19)$, we suggest that TVMA(T) should be selected. 1500 time series have been simulated through the TVMA(6) model. The autocovariances of the simulated time series are shown in Fig. 3.

Table 1: AIC(q) of TVMA(q)

q	1	2	3	4	5	6	7	8	$\cdots$	16	17	18	19
AIC	−45	0.6	48	105	22	−109	−11	110	$\cdots$	−62	−185	−462	-1.05×10^3

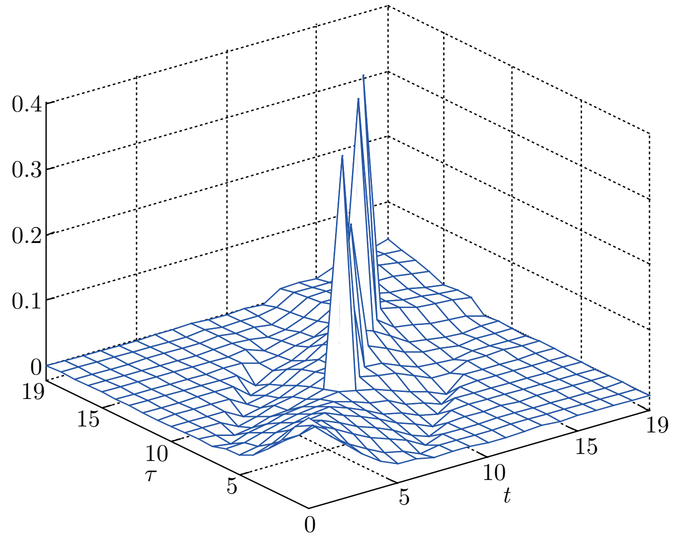

Fig. 3: The autocovariance of simulated time series by TVMA($q = 6$)

AIC analysis is also carried out for TVMA(q_t). Table 2 indicates the AIC values if the order q_t can be only selected within the numbers listed on the first row. If the order of the TVMA(q_t) model can only be selected from 1 to $T/3$ from Table 2, we suggest the order $q_t \leq 4$ should be selected. The autocovariances of time series simulated by TVMA($q_t \leq 4$) are shown as Fig. 4, and the orders q_t in the TVMA(q_t) are shown in Table 3.

Table 2: AIC(q_t) of TVMA(q_t)

$q_t \leq$	1	2	3	4	5	6	7	8	9	10	11	⋯	19
AIC	−45	−32	−103	−340	−289	−151	−291	−342	−353	−345	−345	⋯	−345

Table 3: The orders q_t of TVMA($q_t \leq 4$)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
q_t	1	2	3	4	4	1	4	1	1	1	4	3	3	3	4	4	1	1	1

According to the AIC analysis indicated in Table 1 and Table 2 for a zero mean nonstationary time series shown at the right bottom in Fig. 1, the TVMA(T) model with the minimum AIC specifies it as the best among these three TVMA models, followed by TVMA($q_t \leq 4$) and TVMA($q = 6$).

5 Discussion

The TVMA orders q and q_t affect the separation between the time points t and τ of the autocovariance $c_{t,\tau}$ of the simulated time series. The greater the TVMA orders, the longer the relevant separation.

Although analytical results are verified by simulations on some real time series, it should be noted that the objective time series in this study are zero mean autocovariance nonstationary

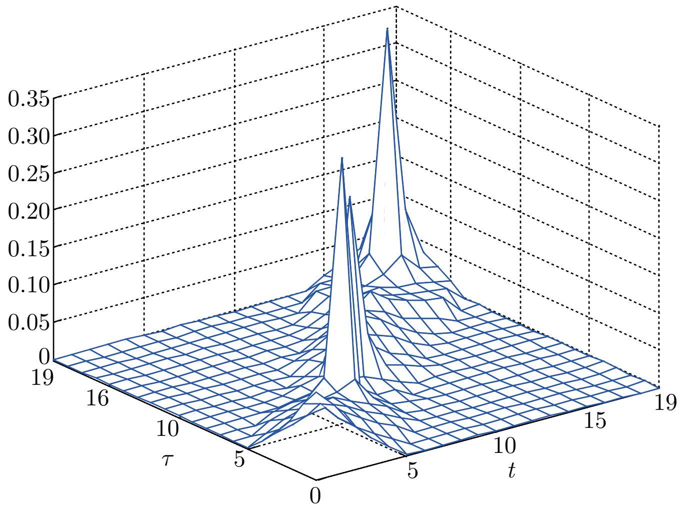

Fig. 4: The autocovariance of simulated time series by TVMA($q_t \leq 4$)

time series which are repeatedly sampled from some univariate stochastic process, differing from the panel data. The TVMA models presented above are effective only on clusters of zero mean autocovariance nonstationary time series but should rarely be used to analyze some single time series sampled only once from univariate stochastic processes.

For simulation of zero mean autocovariance nonstationary time series through learning the original series and obtaining the corresponding parameters $\theta_{k,t}$, simulation can be carried out using the TVMA(T) model. 1500 time series have been successfully simulated using TVMA(T). The results verified that the autocovariances of simulated time series are almost completely equal to those of the original time series.

The AIC values of the TVMA models in Table 1 are greater than those of the TVPAR models [14], for example, the AIC(6) of TVMA(6) (-109) is greater than the AIC(3) of TVPAR(3) (-4.302×10^3), and the AIC(T) of TVMA(T) (-1.05×10^3) is also greater than that of TVPAR(T) (-6.362×10^3). According to Akaike information criterion, the TVMA models proposed here are probably not as good as the TVPAR models. That is to say, the autocovariance error between the original and the simulated (by the TVMA models) size series of cocoon filament is probably greater than that by the TVPAR models. Nevertheless, in the analysis of nonstationary time series, the TVMA model is still a good choice besides the TVPAR model.

6 Conclusion

1. The TVMA models proposed above can be applied to analyze nonstationary time series in statistical analysis and practical engineering.

2. When selecting moving average orders for the above TVMA models according to the above AIC analysis in practical engineering, for convenience, the constant K in Equation (8), (9), and (13) can be ignored. The selection result will remain invariable. For the same cluster of zero mean autocovariance nonstationary time series, n and T are invariable in the three AIC calculation formulas, but the variances $\hat{\theta}^2_{0,t}$ are usually different.

3. For selecting one TVMA model from the above three models to analyze zero mean autocovariance nonstationary time series according to analysis, calculation and experience, it is suggested that: if the time series length is not too long, the TVMA(T) model should be selected for precise simulation; if the length is too long, the TVMA(q_t) model should be selected; and generally, the TVMA($q < T$) model should not be selected unless analyzing an autocovariance stationary time series.

Acknowledgement

We cordially thank the reviewers for their helpful comments. This study is supported by the National Engineering Laboratory for Modern Silk, China, and funded by the First Phase of Jiangsu Universities' Distinctive Discipline Development Program for Textile Science and Engineering of Soochow University.

References

[1] Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis, forecasting and control. Post and Telecom-munication Press, Beijing: 2005. p. 46-130.

[2] Priestley MB. Spectral analysis and time series, Volume 2 Multivariate series, prediction and control. Academic Press, London. 1981. p. 816-890.

[3] Priestley MB. Non-linear and non-stationary time series analysis. Academic Press, London: 1988. p. 140-183.

[4] Fan J, Yao Q. Nonlinear time series: nonparametric and parametric methods. Springer, New York: 2005. p. 89-124.

[5] Dahlhaus R. Fitting time series models to nonstationary processes. Ann Statist: 1997: 1: 1-37.

[6] Dahlhaus R, Giraitis L. On the optimal segment length for parameter estimates for locally stationary time series. J Time Series Anal, 1998: 6: 629-655.

[7] Guo W, Dai M. Multivariate time-dependent spectral analysis using cholesky decomposition. Statistica Sinica: 2006: 825-845.

[8] Davis RA, Lee TCM, Rodriguez-Yam Gabriel A. Structural break estimation for nonstationary time series models. J Am Stat Assoc: 2006: 223-239.

[9] Fei W, Bai L. Statistical characteristics of nonstationary time series. Int J Nonlin Sci Num: Supplement. 2010: 11: 295-299.

[10] Fei W, Bai L. Auto-regressive models of non-stationary time series with finite length. Tsinghua Sci Technol: 2005: 2: 162-168.

[11] Fei W, Bai L, Morikawa H, Miura M. Simulation of size series of cocoon filament with time varying parameter autoregressive model. J Silk Sci Tech Jpn: 2005: 87-91.

[12] Fei W, Bai L. Pattern recognition of non-stationary time series with finite length. Tsinghua Sci Technol: 2006: 5: 611-616.

[13] Grenier Y. Time-dependent ARMA modeling of nonstationary signals. IEEE Transactions on Acoustics, Speech, and Signal Processing: ASSP31(4) 1983: 899-911.

[14] Fei W, Bai L. Time-varying parameter auto-regressive models for autocovariance nonstationary time series. Sci China Ser A: 2009: 3: 577-584.

[15] Akaike H. Information theory and an extention of the maximum likelihood principle. Petrov B. N. and Csaki F. (eds.), 2nd Inter. Symp. on Information Theory, Akademiai Kiado, Budapest: 1973. p. 267-281.

[16] Akaike H. A new look at the statistical model identification. IEEE Trans. Autom. Contr: 1974: AC-19: 716-723.