



Pan, Jiazhu and Xia, Qiang and Liu, Jinshan (2016) Bayesian analysis of multiple thresholds autoregressive model. Computational Statistics. ISSN 1613-9658 , <http://dx.doi.org/10.1007/s00180-016-0673-3>

This version is available at <https://strathprints.strath.ac.uk/57442/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

Bayesian analysis of multiple thresholds autoregressive model

JIAZHU PAN^a QIANG XIA^b JINSHAN LIU^{b,*}

^a *Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK*

^b *Department of Applied Mathematics, South China Agricultural University, Guangzhou, 510642, P.R. China*

Abstract. Bayesian analysis of threshold autoregressive (TAR) model with various possible thresholds is considered. A method of Bayesian stochastic search selection is introduced to identify a threshold-dependent sequence with highest probability. All model parameters are computed by a hybrid Markov chain Monte Carlo (MCMC) method, which combines Metropolis-Hastings (M-H) algorithm and Gibbs sampler. The main innovation of the method introduced here is to estimate the TAR model without assuming the fixed number of threshold values, thus is more flexible and useful. Simulation experiments and a real data example lend further support to the proposed approach.

Key words: Threshold autoregressive model; Bayesian inference; MCMC; Metropolis-Hastings algorithm; Model selection.

AMS Subject classification: 62E15, 62H99

*Corresponding author, Fax: +86 20 85280322. Email addresses: jiazhu.pan@strath.ac.uk(J.Z.Pan), xiaqiang@scau.edu.cn(Q.Xia), liujs58@scau.edu.cn(J.S.Liu)

1 Introduction

The study of nonlinear time series models has received a great deal of attention during the last three decades. One of the most popular nonlinear time series models is the threshold autoregressive (TAR) model proposed by Tong (1978,1983). For this model see also Tong and Lim (1980), Tong (1990) and Tasy (2005), among others. The TAR models can be used to describe many nonlinear phenomena such as limit cycles, chaos, harmonic distortion, jump phenomena and time irreversibility. Hansen (2011) summarized the impact of the TAR model in the fields of econometrics and economics.

Some authors have applied Bayesian approach to estimating unknown parameters in the TAR model to avoid complicated analytical works and numerical multiple integrations in its statistical inference. The intractability of posterior distributions has led to some interesting approaches to model selection and parameter estimation based on Markov chain Monte Carlo (MCMC) methods. For example, Geweke and Terui (1993) considered Bayesian TAR model for nonlinear time series. McCulloch and Tsay (1993a, 1993b) proposed a Bayesian procedure for detecting threshold values in the TAR model via posterior probability plots. Chen and Lee (1995) applied Gibbs sampler and Metropolis-Hastings (M-H) algorithm to inference of TAR models. Chen (1998) gave a Bayesian analysis of generalized TAR models. Chen et al. (2008) considered Bayesian estimation for parsimonious TAR models. Moreover, Sáfadi and Morettin (2000) considered Bayesian analysis of threshold autoregressive moving average models. Ismail and Charif (2003) considered Bayesian inference of threshold moving average models. So et al. (2005) proposed a Bayesian threshold nonlinearity test for financial time series. Some authors have considered multiple-regime TAR models. For example, Chen, Gerlach and Lin (2010) investigated three-regime TAR models with GARCH errors and Bayesian methods for estimation and model selection as well. Brooks and Garrett (2002) considered a three-regime TAR model (SETAR) by the frequentist approach to explain the (mean) dynamics of spot and future markets. In this work, we consider to analyze possible multiple-regime TAR models,

where the number of threshold values is unknown, which is an alternative to the existing works.

We want to propose a data-driven Bayesian approach to analyze of possible multiple threshold values in the TAR model. The main idea of the proposed method is to introduce a sequence of random variables which take the value 1 at those positions associated with threshold values, and 0 otherwise. In our Bayesian framework, the unknown threshold-dependent parameters are estimated using their posterior distributions via maximum a posteriori (MAP) estimation, which possesses good statistical properties (see, e.g., Lavielle and Moulines, 2000). A hybrid MCMC method, which combines the basic Metropolis-Hastings (M-H) algorithm and Gibbs sampler, is used to estimate the threshold-dependent variables and other model parameters. Since the number of the regimes in the TAR model is not assumed to be fixed, so the method introduced here is more flexible than those proposed in the existing literatures on Bayesian analysis of TAR model.

The rest of this paper is arranged as follows. Section 2 presents the TAR model and the methodology of our Bayesian approach. Section 3 gives the details of the procedure of computing MAP estimation via MCMC method. Simulation results and a real data example are provided in Section 4. Section 5 is our conclusion. The proof of our theoretical result is given in Appendix.

2 Threshold Autoregressive Model and Bayesian Inference

2.1 TAR Model

A time series $\{y_t, t = 1, 2, \dots\}$ is said to follow a TAR model with k regimes if it satisfies the following equation

$$y_t = \theta_0^{(j)} + \sum_{i=1}^{q_j} \theta_i^{(j)} y_{t-i} + \varepsilon_t^{(j)}, \quad \text{for } r_{j-1} < y_{t-d} \leq r_j, \quad (2.1)$$

where $j = 1, 2, \dots, k$. For each j , $\{\varepsilon_t^{(j)}\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with normal distribution $N(0, \sigma_j^2)$. The threshold values r_j 's satisfy $-\infty = r_0 < r_1 < \dots < r_k = \infty$ and form a partition of the space of y_{t-d} . The positive integer d

is referred to as the delay (or threshold lag) parameter of the model. We denote the TAR model (2.1) by $\text{TAR}(k; q_1, \dots, q_k)$. The TAR model is a piecewise linear model in the space of y_{t-d} , but not a piecewise linear model in time. Tong (1990) provided an excellent review of this type of models.

Suppose there exists a positive integer q such that $0 \leq k, d, q_i \leq q, i = 1, \dots, k$, and the first q observations $\{y_1, \dots, y_q\}$ are given. Let π_i be the time index of the i th smallest observation of $\{y_{q+1-d}, y_{q+2-d}, \dots, y_{n-d}\}$. Then $y_{\pi_1} < y_{\pi_2} < \dots < y_{\pi_{n-q}}$ and $y_{\pi_i} \in \{y_{q+1-d}, \dots, y_{n-d}\}$. Let $Y = (y_{\pi_1+d}, y_{\pi_2+d}, \dots, y_{\pi_{n-q}+d})$, and $\Theta_i = (\theta_0^{(i)}, \theta_1^{(i)}, \dots, \theta_{q_i}^{(i)})', i = 1, 2, \dots, k$. Given the first q observations, the (conditional) probability density function (p.d.f.) of Y is expressed as

$$\begin{aligned} & f(Y|\Theta_i, \sigma_i^2, r_i, 1 \leq i \leq k; d) \\ \propto & \prod_{i=1}^k (2\pi\sigma_i^2)^{-n_i/2} \exp \left\{ -\frac{1}{2\sigma_i^2} \sum_{j=s_{i-1}+1}^{s_i} \left(y_{\pi_j+d} - \theta_0^{(i)} - \sum_{l=1}^{q_i} \theta_l^{(i)} y_{\pi_j+d-l} \right)^2 \right\} \\ \propto & \prod_{i=1}^k (\sigma_i^2)^{-n_i/2} \exp \left\{ -\frac{1}{2\sigma_i^2} (Y_i^* - X_i^* \Theta_i)' (Y_i^* - X_i^* \Theta_i) \right\}, \end{aligned} \quad (2.2)$$

where the symbol \propto means direct proportion, $s_i, i = 1, 2, \dots, k-1$, satisfy $y_{\pi_{s_i}} \leq r_i < y_{\pi_{s_i+1}}$, $s_0 = 0, s_k = n - q$, and

$$Y_i^* = (y_{\pi_{s_{i-1}+1}+d}, y_{\pi_{s_{i-1}+2}+d}, \dots, y_{\pi_{s_i}+d})'$$

is the observations generated by regime i in order of occurrence,

$$X_i^* = (x_{i,s_{i-1}+1}, x_{i,s_{i-1}+2}, \dots, x_{i,s_i})'$$

is an $n_i \times (q_i + 1)$ matrix with $x_{i,l} = (1, y_{\pi_l+d-1}, \dots, y_{\pi_l+d-q_i})'$, while $n_i = s_i - s_{i-1}$ is the width of regime i . The parameters to be estimated for the TAR model are $k, \Theta_i, \sigma_i^2, r_i, i = 1, \dots, k$, and d .

In order to estimate the threshold values $\{r_i\}$, we introduce a random process $\{\gamma_i\}$ defined by

$$\gamma_{t-q} = \begin{cases} 1, & \text{if there exists } i \text{ such that } t = \pi_{s_i} + d, \\ 0, & \text{otherwise,} \end{cases} \quad (2.3)$$

$t = q + 1, 2, \dots, n - 1$. Then the estimation of the threshold values $r_i, i = 1, \dots, k$, reduces to the estimation of the vector $\gamma = (\gamma_1, \dots, \gamma_{n-q-1})$ and the delay parameter d . In fact, if $\gamma_i = 1$, then $\pi_{s_i} = i + q - d$ is determined, so that an interval that a threshold value r belongs to can be determined by the inequality $y_{\pi_{s_i}} \leq r < y_{\pi_{s_i+1}}$, where $y_{\pi_{s_i}} = y_{i+q-d}$, while $y_{\pi_{s_i+1}}$ is just greater than y_{i+q-d} in the set $\{y_{q+1-d}, \dots, y_{n-d}\}$. Note that the number k of the threshold values can be estimated by $K_\gamma = \sum_{k=1}^{n-q-1} \gamma_k + 1$. Denote $\Theta = (\Theta_1, \dots, \Theta_{K_\gamma})$ and $\sigma^2 = (\sigma_1^2, \dots, \sigma_{K_\gamma}^2)$, then the likelihood function of the parameters Θ, γ, σ^2 and d can be written as

$$L(\Theta, \gamma, \sigma^2, d|Y) \propto \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-n_k/2} \exp \left\{ -\frac{1}{2\sigma_k^2} (Y_k^* - X_k^* \Theta_k)' (Y_k^* - X_k^* \Theta_k) \right\}. \quad (2.4)$$

2.2 Bayesian Inference

We shall adopt a Bayesian approach to inference of the TAR model based on the posterior distribution of unknown parameters. To this end, we need to define the prior distribution of parameters.

Firstly, we consider $\{\gamma_i\}$ to be a sequence of i.i.d. Bernoulli random variables. Then for any $\gamma = (\gamma_1, \dots, \gamma_{n-q-1})$ in $\Omega = \{0, 1\}^{n-q-1}$, the prior probability mass function of γ given d is given by

$$\pi(\gamma|d) = \lambda^{K_\gamma-1} (1 - \lambda)^{n-q-K_\gamma}, \quad (2.5)$$

where $\lambda \in [0, 1]$ is the Bernoulli parameter, which represents the prior probability that there are thresholds at some given positions.

Given γ and d , we take the prior of $\Theta_i, i = 1, \dots, K_\gamma$, to be independent multivariate normal distribution $N(\Theta_{0i}, V_i^{-1})$ and $\sigma_i^2, i = 1, \dots, K_\gamma$, independent inverse gamma distribution $IG(v_i/2, v_i \lambda_i/2)$. The prior of d is assumed to follow a discrete uniform distribution on a set $D = \{1, \dots, d_0\}$, denoted by $\pi(d)$, where d_0 is a prescribed positive integer.

Let $\phi = (\lambda, d_0, \Theta_{0i}, V_i, v_i, \lambda_i, i \geq 1)$ denote the set of hyper-parameters, which are assumed to be known. To implement Bayesian inference, we need the joint posterior distribution of $(\Theta, \gamma, \sigma^2, d)$, which combines the prior distributions and the likelihood function.

The joint prior distribution of $(\Theta, \gamma, \sigma^2, d)$ can be expressed by

$$\pi(\Theta, \gamma, \sigma^2, d) = \pi(\Theta|\gamma, d)\pi(\sigma^2|\gamma, d)\pi(\gamma|d)\pi(d). \quad (2.6)$$

The joint posterior distribution of $(\Theta, \gamma, \sigma^2, d)$ is

$$\begin{aligned} f(\Theta, \gamma, \sigma^2, d|Y) &\propto L(\Theta, \gamma, \sigma^2, d|Y)\pi(\Theta, \gamma, \sigma^2, d) \\ &\propto \exp \left\{ - \sum_{k=1}^{K_\gamma} \frac{1}{2\sigma_k^2} [(Y_k^* - X_k^* \Theta_k)'(Y_k^* - X_k^* \Theta_k) + v_k \lambda_k] \right\} \\ &\quad \times \prod_{k=1}^{K_\gamma} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \exp \left\{ -\frac{1}{2} (\Theta_k - \Theta_{0k})' V_k (\Theta_k - \Theta_{0k}) \right\} \\ &\quad \times \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-(n_k+v_k+2)/2} \left(\frac{\lambda}{1-\lambda} \right)^{K_\gamma}. \end{aligned} \quad (2.7)$$

Using standard Bayesian techniques, we derive the conditional posterior distributions of $\Theta_i, \sigma_i^2, i = 1, 2, \dots, K_\gamma$ and d as follows.

The conditional posterior distribution of Θ_i is independent of Θ_j for $i \neq j$ and

$$\Theta_i|Y, \gamma, \sigma^2, d \sim N(\Theta_i^*, V_i^{*-1}), \quad i = 1, \dots, K_\gamma, \quad (2.8)$$

where

$$V_i^* = \left(\frac{X_i^{*'} X_i^*}{\sigma_i^2} + V_i \right)$$

and

$$\Theta_i^* = (V_i^*)^{-1} \left(\frac{X_i^{*'} X_i^*}{\sigma_i^2} \hat{\Theta}_i + V_i \Theta_{0i} \right)$$

with $\hat{\Theta}_i = (X_i^{*'} X_i^*)^{-1} X_i^{*'} Y_i^*, i = 1, \dots, K_\gamma$.

The conditional posterior distribution of σ_i^2 is independent of σ_j^2 for $i \neq j$ and

$$\sigma_i^2|Y, \Theta, \gamma, d \sim IG \left(\frac{v_i + n_i}{2}, \frac{v_i \lambda_i + n_i S_i^2}{2} \right), \quad (2.9)$$

i.e.

$$\frac{v_i \lambda_i + n_i S_i^2}{\sigma_i^2} |Y, \Theta, \gamma, d \sim \chi^2(v_i + n_i), \quad i = 1, \dots, K_\gamma,$$

where $n_i = s_i - s_{i-1}, S_i^2 = n_i^{-1} (Y_i^* - X_i^* \Theta_i)' (Y_i^* - X_i^* \Theta_i)$.

The conditional posterior distribution of d is a multinomial distribution with probability mass function

$$f(d|Y, \Theta, \gamma, \sigma^2) = \frac{L(\Theta, \gamma, \sigma^2, d|Y)}{\sum_{d=1}^{d_0} L(\Theta, \gamma, \sigma^2, d|Y)}, \quad d = 1, 2, \dots, d_0, \quad (2.10)$$

where $L(\Theta, \gamma, \sigma^2, d|Y)$ is the likelihood function defined by (2.4).

Thus, conditionally on the observations and other parameters, $\Theta_i, i = 1, \dots, K_\gamma$, remain independent and follow normal distributions, $\sigma_i^2, i = 1, \dots, K_\gamma$, remain independent and follow inverse gamma distributions. All the conditional posterior distributions of the unknown parameters, except for γ , can be identified. The estimates of the parameters should be computed by a hybrid MCMC method which combining M-H algorithm and Gibbs sampler.

The posterior distribution of γ is proportional to

$$L(\Theta, \gamma, \sigma^2, d|Y) \prod_{i=1}^{K_\gamma} \pi(\Theta_i|\gamma, d) \pi(\sigma_i^2|\gamma, d) \pi(\gamma|d) \pi(d). \quad (2.11)$$

By integrating the parameters $\Theta_i, i = 1, \dots, K_\gamma$, in (2.11), we obtain the conditional posterior distribution $f(\gamma, d|Y, \sigma^2)$ given by the following Theorem.

Theorem 1. *For any configuration of $\gamma = (\gamma_1, \dots, \gamma_{n-q-1})$, let K_γ be the number of regimes and*

$$S_k^{*2} = \frac{1}{\sigma_k^2} Y_k^{*'} Y_k^* - \Theta_k^{*'} V_k^* \Theta_k^*, \quad (2.12)$$

where Θ_k^* and V_k^* are defined as in (2.8). Then the conditional posterior distribution of (γ, d) given (Y, σ^2) is given by

$$f(\gamma, d|Y, \sigma^2) \propto \exp \{ -U(\gamma, d|Y, \sigma^2) \}, \quad (2.13)$$

where $\gamma \in \Omega = \{0, 1\}^{n-q-1}$, $d \in D = \{1, \dots, d_0\}$ and

$$U(\gamma, d|Y, \sigma^2) = \frac{1}{2} \sum_{k=1}^{K_\gamma} (S_k^{*2} + \omega_k) + \beta K_\gamma \quad (2.14)$$

in which

$$\begin{aligned} \omega_k &= (n_k + v_k + 2) \ln \sigma_k^2 + (q_k + 1) \ln(2\pi) - \ln |V_k| + \Theta_{k0}' V_k \Theta_{k0} + \frac{v_k \lambda_k}{\sigma_k^2}, \\ \beta &= \ln \frac{1 - \lambda}{\lambda}. \end{aligned}$$

The proof of Theorem 1 is presented in Appendix 1.

If $\sigma_1^2 = \dots = \sigma_k^2 = \sigma^2$ and the prior distribution of σ^2 is $IG(v/2, v\lambda/2)$, then the likelihood function (2.4) reduces to

$$L(\Theta, \gamma, \sigma^2, d|Y) \propto (\sigma^2)^{-(n-q)/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^{K_\gamma} (Y_k^* - X_k^* \Theta_k)' (Y_k^* - X_k^* \Theta_k) \right\}. \quad (2.15)$$

In this situation, the conditional posterior distribution of σ^2 is

$$\sigma^2 | Y, \Theta, \gamma, d \sim IG \left(\frac{v + n - q}{2}, \frac{v\lambda + S_\gamma^2}{2} \right), \quad (2.16)$$

where $S_\gamma^2 = \sum_{k=1}^{K_\gamma} (Y_k^* - X_k^* \Theta_k)' (Y_k^* - X_k^* \Theta_k)$, and the quantity ω_k in (2.14) reduces to

$$\omega_k = (q_k + 1) \ln(2\pi) - \ln |V_k| + \Theta_{k0}' V_k \Theta_{k0}. \quad (2.17)$$

The proof of (2.17) is given in Appendix 2.

It follows from Theorem 1 that the conditional marginal posterior distribution of γ is given by

$$f(\gamma | Y, \sigma^2) \propto \sum_{d=1}^{d_0} L(\gamma, d | Y, \sigma^2), \quad (2.18)$$

where $L(\gamma, d | Y, \sigma^2) = \exp \{-U(\gamma, d | Y, \sigma^2)\}$ and $U(\gamma, d | Y, \sigma^2)$ is defined by (2.14). Consequently, a simplified conditional posterior distribution of d is obtained with probability mass function

$$f(d | Y, \gamma, \sigma^2) = \frac{L(\gamma, d | Y, \sigma^2)}{\sum_{d=1}^{d_0} L(\gamma, d | Y, \sigma^2)}, \quad d = 1, 2, \dots, d_0. \quad (2.19)$$

It is noticed that for any $1 \leq i \leq n - q - 1$, the conditional posterior distribution $P(\gamma_i = 1 | Y, \sigma^2)$ gives the probability to have a threshold value $r \in [y_{\pi_{s_i}}, y_{\pi_{s_i+1}})$ where $\pi_{s_i} = i + q - d$. For a given estimate $\hat{\sigma}^2$ of σ^2 , the MAP estimate of γ is one of the standard Bayesian estimations defined by

$$\hat{\gamma} = \arg \max_{\gamma} f(\gamma | Y), \quad (2.20)$$

where $f(\gamma | Y) = f(\gamma | Y, \hat{\sigma}^2)$. Unfortunately, closed-form expressions of MAP estimate of γ can not be obtained. We will use an M-H procedure in the MCMC sampling to carry out numerical computation for γ .

3 MCMC method

The MAP estimate of γ is obtained by constructing a homogeneous Markov chain using the M-H algorithm with the invariant distribution $f(\gamma|Y)$. In this procedure we will use a simulated annealing (SA) algorithm. The SA algorithm defines a non-homogeneous Markov chain which converges, under appropriate conditions, to the maximizer of the posterior probability density function $f(\gamma|Y)$. A decreasing temperature schedule is introduced in the SA algorithm, which modifies the acceptance probability.

Denote the current state of Markov chain by $\gamma^{(i)} = (\gamma_1^{(i)}, \dots, \gamma_{n-q-1}^{(i)})$. The M-H algorithm is an iterative procedure. At iteration $i + 1$, we carry out the following two steps:

Step 1: a candidate $\tilde{\gamma}$ is drawn from a proposal kernel $Q(\gamma^{(i)}, \tilde{\gamma})$.

Step 2: $\tilde{\gamma}$ is accepted as the $(i + 1)$ th new state, i.e. $\gamma^{(i+1)} = \tilde{\gamma}$, with the probability

$$\alpha(\tilde{\gamma}, \tilde{\gamma}) = \min \left\{ 1, \frac{f(\tilde{\gamma}|Y)Q(\tilde{\gamma}, \gamma^{(i)})}{f(\gamma^{(i)}|Y)Q(\gamma^{(i)}, \tilde{\gamma})} \right\}; \quad (3.1)$$

otherwise, $\gamma^{(i+1)} = \gamma^{(i)}$.

In order to enhance the speed of convergence, it is important to allow more communications between the states with high probabilities. This can be done by using the following three kernels Q_1, Q_2 and Q_3 successively at each iteration:

(1) Q_1 is a proposal kernel from which the candidate $\tilde{\gamma}$ is drawn independently of the current state γ defined by $Q_1(\gamma, \tilde{\gamma}) = \pi(\tilde{\gamma})$, where $\pi(\gamma)$ is the prior density (2.5). The independent sampler allows for rapid motion to distant parts of the state space. However, the global acceptance probability for this sampler is very low for large data sets.

(2) Q_2 is a proposal kernel by which a new threshold is created or an existing threshold is removed. In this move, local changes are made from the so-called one-variable-at-a-time M-H algorithm suggested, for instance, by Chib and Greenberg (1995), to increase the speed of convergence. More precisely, a random permutation of $\{1, \dots, n - q - 1\}$ is uniformly drawn. According to this permutation, each component is flipped from 0 to 1 or from 1 to 0. The move

is accepted with the usual acceptance probability. This move visits each site randomly and all sites are visited in each scan.

(3) Q_3 is a proposal kernel by which an existing threshold is moved. In this move, two time points s_1 and s_2 are randomly chosen such that $\gamma_{s_1} = 1$ and $\gamma_{s_2} = 0$. Then, $\tilde{\gamma}_t = \gamma_t$ for all $t \neq s_1, s_2$ while $\tilde{\gamma}_{s_1} = 0$ and $\tilde{\gamma}_{s_2} = 1$. The threshold value is finally moved and accepted according to the acceptance probability. Such move is very important since it avoids trapping in a threshold neighborhood.

Each kernel is used in turn and the resulting hybrid strategy is called a cycle. The resulting cycle kernel is irreducible and aperiodic (see, e.g., Chen and Lee, 1995).

In the MAP algorithm, a Markov chain is constructed to simulate the target distribution $f(\gamma|Y)$. At each step of the cycle, the acceptance probability is defined by (3.1). A schedule for lowering the temperature is defined by $T_k = 0.9T_{k-1}$, where T_0 is greater than a numerical constant. This temperature decrease is made at each step of the independent sampler. If R is the outcome of a uniform drawing on $[0, 1]$, then

$$\gamma^{(i+1)} = \begin{cases} \tilde{\gamma}, & \text{if } R^{T_k} < \frac{f(\tilde{\gamma}|Y)Q(\tilde{\gamma}, \gamma^{(i)})}{f(\gamma^{(i)}|Y)Q(\gamma^{(i)}, \tilde{\gamma})}, \\ \gamma^{(i)}, & \text{otherwise,} \end{cases} \quad (3.2)$$

where T_k is the current temperature. After a sufficiently long burn-in, the MAP estimate of γ is determined by computing the time average of output samples of the Markov chain.

To find the MCMC estimates for all parameters $(\gamma, \sigma^2, \Theta, d)$, start with initial values for the parameter σ^2 , and then cycle through the following steps. The estimate of γ is obtained from the above M-H procedure. Then for given values of σ^2 and γ , the value of d is obtained from the conditional posterior distribution $f(d|Y, \gamma, \sigma^2)$ given by (2.19). Next for given values of (σ^2, γ, d) the values of Θ_i 's are obtained from the distributions given by (2.8), and for given values (Θ, γ, d) the values of σ_i^2 's are obtained from (2.9). After a sufficiently long burn-in, compute the mean of output samples of the Markov chain for each of the parameters, which are the MCMC estimates of the parameters. Thus the estimates of these parameters can be obtained by a hybrid MCMC method combining Gibbs sampler and M-H algorithm.

The above procedure can be implemented once again. In the first stage the threshold values should be determined under some given integers q_1, \dots, q_k , where q_j is the order of the autoregressive model in regime j . The hyper-parameters $\phi = (\lambda, d_0, \Theta_{0i}, V_i, v_i, \lambda_i, i \geq 1)$ should be selected in somewhat arbitrarily. Once the threshold values are determined, other parameters should be estimated by Gibbs sampling methods, and the order of the autoregressive model in each segment can be determined by AIC criterion. In the following stages, the orders and the hyper-parameters of the model should be determined from the early stages. With these quantities the number and positions of the threshold values should be estimated again and consequently the other parameters can be estimated again. The procedure should be implemented further until the best result is obtained.

4 Simulation and Application Examples

4.1 Simulation Experiments

In this subsection, we use three simulation examples to demonstrate the efficiency of our method. We will generate data from the models with known parameters and then using the data to estimate model $TAR(k; q_1, \dots, q_k)$. The efficiency of our method can be seen by comparing the estimated results with the source models.

Example 1: AR(1)

$$y_t = 0.5y_{t-1} + \varepsilon_t, \quad (4.1)$$

where $\varepsilon_t \sim N(0, 4)$.

Example 2: TAR(2;1,1)

$$y_t = 0.5y_{t-1}I(y_{t-1} \leq -0.4) - 0.5y_{t-1}I(y_{t-1} > -0.4) + \varepsilon_t, \quad (4.2)$$

where $\varepsilon_t \sim N(0, 4)$ if $y_{t-1} > -0.4$ and $\varepsilon_t \sim N(0, 1)$ otherwise, $I(A)$ stands for the indicator function of A , i.e. $I(A) = 1$ if A is true and 0 otherwise.

Example 3: TAR(3;1,1,1)

$$\begin{aligned}
y_t &= 0.01y_{t-1}I(y_{t-1} \leq -1.4) + 0.5y_{t-1}I(-1.4 < y_{t-1} \leq 0.8) \\
&\quad - 0.5y_{t-1}I(y_{t-1} > 0.8) + \varepsilon_t,
\end{aligned} \tag{4.3}$$

where $\varepsilon_t \sim N(0, 1)$ if $y_{t-1} > 0.8$, $\varepsilon_t \sim N(0, 4)$ if $y_{t-1} > -1.4$ and $y_{t-1} \leq 0.8$, and $\varepsilon_t \sim N(0, 2.25)$ if $y_{t-1} \leq -1.4$.

We generate 100 samples of size $n = 200$ from the models in Examples 1 and 2 respectively, and 100 samples of size $n = 400$ from the model in Example 3. The hyper-parameter λ for the three examples are chosen as $1/(n - q - 2)$. The other hyper-parameters used in the MCMC algorithm are given by $\{d_0 = 3, \theta_{1i} = 0, V_i = 0.1, i \geq 1\}$. We take $q = 3$ for all three examples. For each sample we conduct MCMC estimation based on the model TAR($k; 1, \dots, 1$). Simulation results change very little when we increased the length of Markov chain to 5000 cycles. This indicates that the Markov chains have attained convergence. The estimates of all parameters are obtained by using 5000 iterations after 5000 burn-in cycles. Fig.1, which is depicted by summarizing 100 independent samples from each model, presents the box plots of the average absolute error (AAE) between the true values and estimates of y'_i 's for all three examples. It is shown that our method is fine because the AAEs are quite small.

To illustrate our algorithm more thoroughly, we study a sample for each example. The estimates of the marginal posterior probabilities $\{P(\gamma_t = 1|Y), t = 1, \dots, n - q - 1\}$ by the MAP algorithm are plotted in Fig.2. For Example 1, no threshold was detected. For Example 2, the probability of $\hat{\gamma} = (0, \dots, 0, \hat{\gamma}_{159} = 1, 0, \dots, 0)$ is 0.82. This result indicates that $K_\gamma = 2$ and $r_1 \in [-0.5287, -0.4718)$. Similarly, for Example 3, the probability of $\hat{\gamma} = (\hat{\gamma}_1 = 0, \dots, \hat{\gamma}_{92} = 1, 0, \dots, 0, \hat{\gamma}_{222} = 1, 0, \dots, 0)$ is near to 0.7. The estimated threshold values are $r_1 \in [-1.3570, -1.4974)$ and $r_2 \in [0.7705, 0.8077)$. Tables 1 to 3 provide the averages of posterior means, medians, standard deviations, and 95% Bayesian credible intervals over the 100 replications after the burn-in period for all parameters. For integer parameters, d and k , we report the average of MAPs, with corresponding probability estimates in the brackets. This in-

icates that our method is effective. To assess the convergence of the Markov chains intuitively, we use the trace plots of the sampling process for model parameters. Fig.3 to Fig.5 present the trace plots for the three simulated models, respectively. It can be seen that the Markov chains are stationary, which indicate that the chains have attained convergence.

Table 1 Simulation results for AR(1)

Parameters	True value	Means	Medians	Standard deviations	Credible intervals	MAP
$\theta_1^{(1)}$	0.5	0.4716	0.4721	0.0613	(0.371, 0.573)	-
σ^2	4	4.1831	4.1561	0.4226	(3.539, 4.912)	-
k	1	-	-	-	-	1(96%)

Table 2 Simulation results for TAR(2;1,1)

Parameter	True value	Means	Medians	Standard deviations	Credible intervals	MAP
$\theta_1^{(1)}$	0.5	0.4228	0.4239	0.0646	(0.316, 0.527)	-
$\theta_1^{(2)}$	-0.5	-0.5554	-0.5559	0.0818	(-0.834, -0.308)	-
σ_1^2	1	1.5921	1.5814	0.0971	(1.322, 1.899)	-
σ_2^2	4	3.6678	3.5641	0.4181	(2.748, 5.028)	-
r	-0.4	-	-	-	(-0.529, -0.471)	-
d	1	-	-	-	-	1(100%)
k	2	-	-	-	-	2(82%)

Table 3 Simulation results for TAR(3;1,1,1)

Parameter	True value	Means	Medians	Standard deviations	Credible intervals	MAP
$\theta_1^{(1)}$	0.01	0.0385	0.0386	0.0349	(-0.010, 0.085)	-
$\theta_1^{(2)}$	0.5	0.8789	0.8781	0.1077	(0.554, 1.106)	-
$\theta_1^{(3)}$	-0.5	-0.5282	-0.5281	0.0887	(-0.676, -0.384)	-
σ_1^2	1	0.9903	0.9749	0.0955	(0.785, 1.255)	-
σ_2^2	4	3.9565	3.9218	0.1899	(3.222, 4.834)	-
σ_3^2	2.25	2.4988	2.4778	0.0841	(2.101, 2.957)	-
r_1	-1.4	-	-	-	(-1.497, -1.357)	-
r_2	0.8	-	-	-	(0.771, 0.808)	-
d	1	-	-	-	-	1(100%)
k	3	-	-	-	-	3(68%)

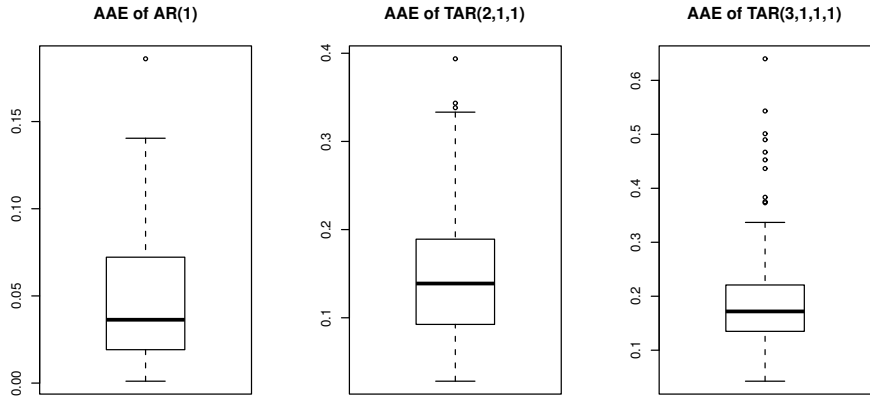


Figure 1: Boxplots of AAE between the true values and estimates of y'_i s for three models

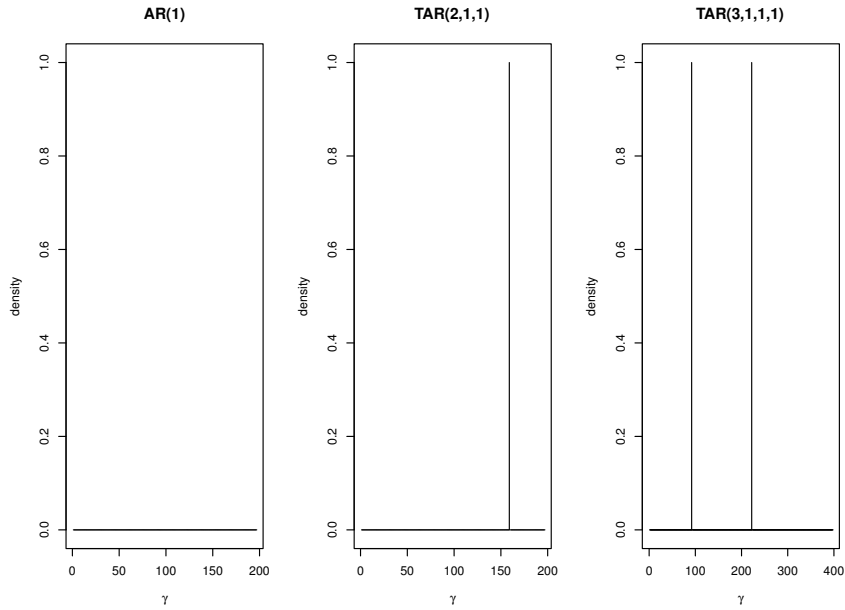


Figure 2: The marginal posterior distributions of γ for three models

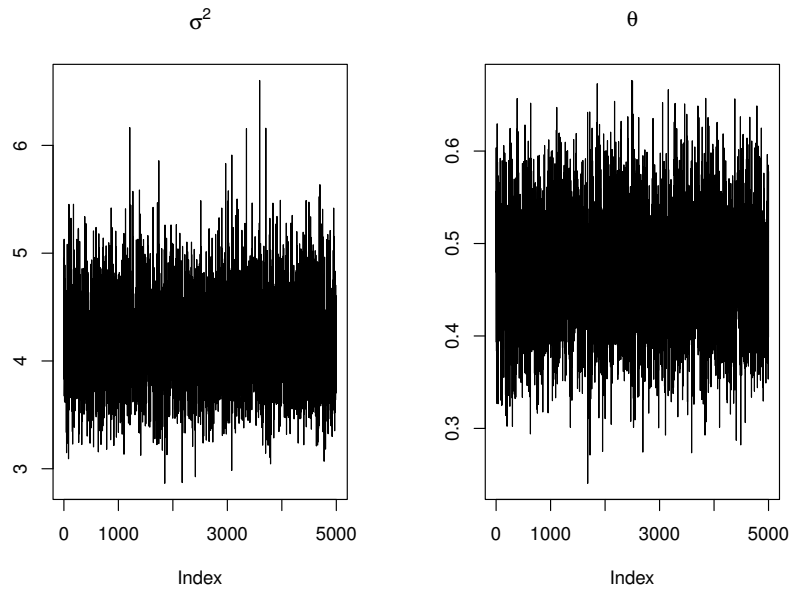


Figure 3: The trace plots of parameters of AR(1) model

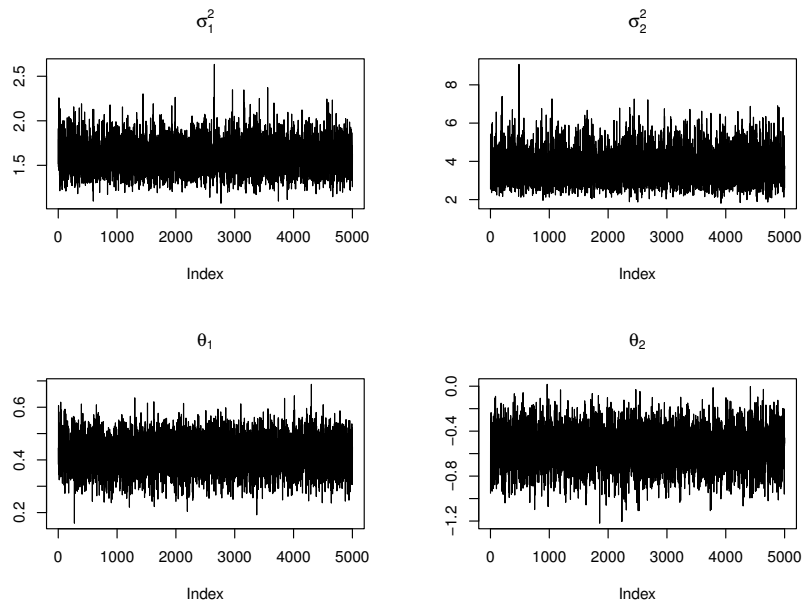


Figure 4: The trace plots of parameters of TAR(2,1,1) model

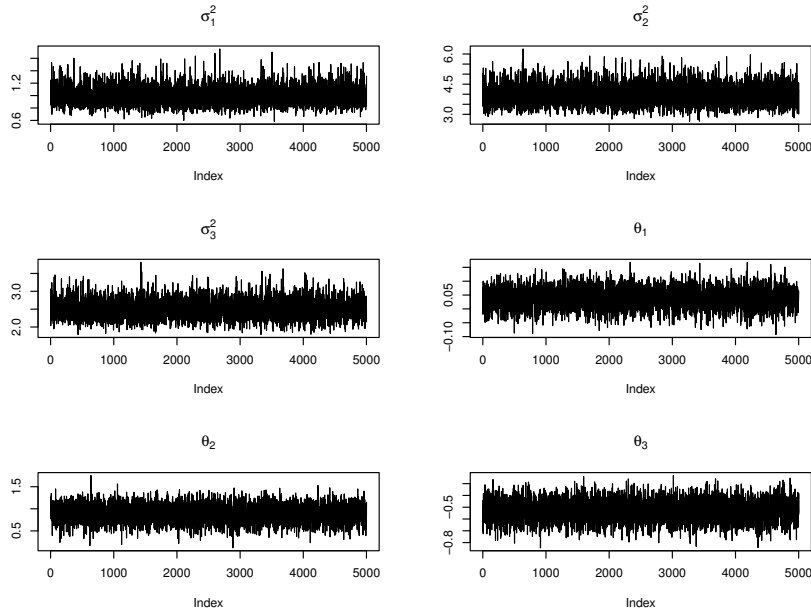


Figure 5: The trace plots of parameters of TAR(3,1,1,1) model

4.2 A Real Data Example

In this subsection, we illustrate our method by analyzing a real data example, i.e. sunspot number, which is yearly data of sunspot numbers from 1700 to 1979, given by Tong (1983). The series consists 280 observations and is known to exhibit asymmetric cyclic behavior. Fig.6 present the plots of the data with the regime cut-offs for the sunspot numbers.

Various linear and nonlinear models have been proposed for this series. In general, for this series it seems that different data spans would suggest different models. Among others, Tsay (1989) employed a three regimes TAR model to analyze the data of sunspot numbers. The AR orders of the three regimes refined by AIC criterion are 11, 10, and 10. In this work, without restricting the number of regimes, we employ the more general model $\text{TAR}(k; q_1, \dots, q_k)$ to fit the data of sunspot numbers. We use all the 280 observations in model building. The hyper-parameters are chosen as $\{q = 11, \lambda = 1/269, d_0 = 3, \theta_{i0} = 0, V_i = I_{12}, i \geq 1\}$.

It is seen from the trace plots of the Markov chain that simulation results change very little when we increased the length of the chain to 5000 cycles. This suggested that the Markov

chains have attained convergence. The estimates of all parameters are obtained by using 5000 iterations after 7000 burn-in cycles. Fig.7 presents the posterior probability plot of $\{P(\gamma_t = 1|Y), t = 1, \dots, n - q - 1\}$ and the marginal posterior distributions of K_γ and d . It shows that three threshold values are detected, i.e. $K_\gamma = 4$, and $d = 2$. For the MAP algorithm, the probability estimate of $\gamma = (0, \dots, 0, \gamma_{45} = 1, 0, \dots, 0, \gamma_{137} = 1, 0, \dots, 0, \gamma_{210} = 1, 0, \dots, 0)$ is 0.76. Thus, the threshold values r_1, r_2 and r_3 are detected in the intervals (10.2, 10.7), (40.0, 40.1) and (73.0, 74.0), respectively. The orders of each regime selected by AIC criterion are 7, 4, 11 and 10, respectively. In fact, the AIC value of our model is 1343.83 which is slightly less than 1379.4, the AIC value of Tsay (1989). Table 4 summarizes the results of estimated autoregressive coefficients in each regime and the estimated threshold parameters.

Table 4 Results of estimated parameters for sunspot numbers

Regime i	$\theta_0^{(i)}$	$\theta_1^{(i)}$	$\theta_2^{(i)}$	$\theta_3^{(i)}$	$\theta_4^{(i)}$	$\theta_5^{(i)}$	$\theta_6^{(i)}$
1	-4.8808 (1.8247)	2.0866 (1.0400)	-2.5091 (0.9585)	0.6104 (0.2396)	-0.8401 (0.3978)	0.8228 (0.4043)	-0.2238 (0.1233)
2	18.8633 (6.0323)	1.5922 (0.3480)	-0.8239 (0.3914)	-0.5061 (0.2224)	0.2357 (0.1241)	- -	- -
3	1.6252 (0.8482)	0.8585 (0.4143)	0.2014 (0.2922)	-0.5022 (0.2861)	0.2531 (0.1813)	-0.1679 (0.1223)	-0.1038 (0.1417)
4	1.4945 (0.1512)	0.6309 (0.1085)	0.1346 (0.1186)	-0.1265 (1855)	0.0178 (0.1721)	-0.1783 (0.1145)	0.1272 (0.1152)
Regime i	$\theta_7^{(i)}$	$\theta_8^{(i)}$	$\theta_9^{(i)}$	$\theta_{10}^{(i)}$	$\theta_{11}^{(i)}$	σ_i^2	r_i
1	0.3124 (0.1518)	- -	- -	- -	- -	130.3167 (23.82)	(10.2, 10.7) -
2	- -	- -	- -	- -	- -	166.1455 (28.04)	(40.0, 40.1) -
3	0.3033 (0.3825)	-0.1833 (0.2534)	-0.0211 (0.4167)	0.0554 (0.3825)	0.2077 (0.1127)	82.3864 (16.27)	(73.0, 74.0) -
4	0.3550 (0.2051)	-0.4292 (0.2051)	0.1555 (0.1325)	0.1951 (0.1017)	- -	88.883 (13.39)	

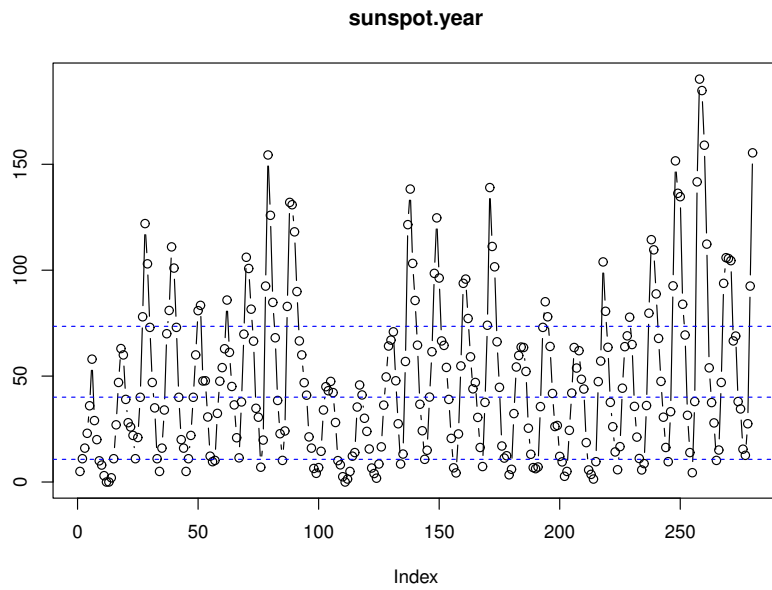


Figure 6: The time series of the yearly sunspot numbers

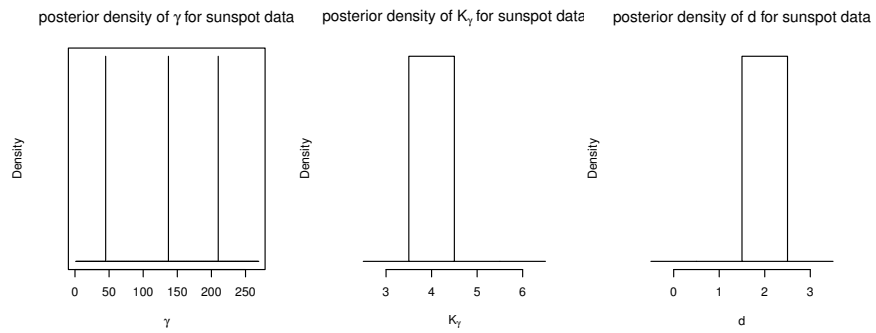


Figure 7: The marginal posterior distributions of γ , K_γ and d for sunspot numbers

5 Conclusion

In this paper, we consider Bayesian analysis of TAR model with possible multiple threshold values. Without assuming fixed number of the regimes, a method of Bayesian stochastic search selection is introduced for detecting threshold values of the model. For Bayesian inference, we derived the posterior distributions of the unknown parameters, particularly that of the threshold-dependent parameters. A hybrid MCMC method combining M-H algorithm and Gibbs sampler is established to compute the model parameters. The details of the procedure is presented for computing the MAP estimation via MCMC method. The major advantage of the methodology introduced here is that it avoids given the fixed number of thresholds, thus is flexible.

Numerical experiments examples show that the approach proposed here is effective in detecting the threshold values for various TAR models. It can handle multiple thresholds in a direct manner. The real data example analysis shows that our method is feasible in practice. For the sunspot data, three threshold values are detected by our method, two of them are close to those detected by Tsay (1989).

ACKNOWLEDGMENTS

The authors thank the editors and the referee for their constructive comments and valuable suggestions. This paper is partially supported by the EPSRC Bridging-the-Gaps project and the starter grant from the Faculty of Science, University of Strathclyde. The authors J. S. Liu and Q. Xia are supported by National Science Foundation of China (No.11171117), National Science Foundations of Guangdong Province of China (No.2016A030313414, No.S2011010002371) and National statistical plan for scientific research project of China (No.2015LZ48).

Appendix 1. The proof of Theorem 1

The conditional posterior distribution of (γ, d) given (Y, σ^2) is

$$\begin{aligned} f(\gamma, d|Y, \sigma^2) &\propto \int f(\Theta, \gamma, \sigma^2, d|Y)d\Theta \\ &\propto \pi(\sigma^2|\gamma, d)\pi(\gamma|d) \int L(\Theta, \gamma, \sigma^2, d|Y)\pi(\Theta|\gamma, d)d\Theta, \end{aligned} \quad (\text{A.1})$$

where $L(\Theta, \gamma, \sigma^2, d|Y)$ is defined by (2.4) and

$$\begin{aligned}
& L(\Theta, \gamma, d|Y)\pi(\Theta|\gamma, d) \\
& \propto \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-n_k/2} \exp \left\{ -\frac{1}{2\sigma_k^2} (Y_k^* - X_k^* \Theta_k)' (Y_k^* - X_k^* \Theta_k) \right\} \\
& \quad \times \prod_{k=1}^{K_\gamma} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \exp \left\{ -\frac{1}{2} (\Theta_k - \Theta_{k0})' V_k (\Theta_k - \Theta_{k0}) \right\}. \tag{A.2}
\end{aligned}$$

Let Θ_k^* , V_k^* and $\hat{\Theta}_k$ be defined as before, then we have

$$\begin{aligned}
& \frac{1}{\sigma_k^2} (Y_k^* - X_k^* \Theta_k)' (Y_k^* - X_k^* \Theta_k) + (\Theta_k - \Theta_{k0})' V_k (\Theta_k - \Theta_{k0}) \\
& = [Y_k^* - X_k^* \hat{\Theta}_k + X_k^* (\hat{\Theta}_k - \Theta_k)]' [Y_k^* - X_k^* \hat{\Theta}_k + X_k^* (\hat{\Theta}_k - \Theta_k)] \\
& \quad + (\Theta_k - \Theta_{k0})' V_k (\Theta_k - \Theta_{k0}) \\
& = \frac{1}{\sigma_k^2} [(Y_k^* - X_k^* \hat{\Theta}_k)' (Y_k^* - X_k^* \hat{\Theta}_k) + (\hat{\Theta}_k - \Theta_k)' (X_k^{*'} X_k^*) (\hat{\Theta}_k - \Theta_k)] \\
& \quad + (\Theta_k - \Theta_{k0})' V_k (\Theta_k - \Theta_{k0}) \\
& = \frac{1}{\sigma_k^2} [(Y_k^* - X_k^* \hat{\Theta}_k)' (Y_k^* - X_k^* \hat{\Theta}_k) + \hat{\Theta}_k' X_k^{*'} X_k^* \hat{\Theta}_k] + \Theta_{k0}' V_k \Theta_{k0} \\
& \quad + (\Theta_k - \Theta_k^*)' V_k^* (\Theta_k - \Theta_k^*) - \Theta_k^{*'} V_k^* \Theta_k^* \\
& = \frac{1}{\sigma_k^2} Y_k^{*'} Y_k^* - \Theta_k^{*'} V_k^* \Theta_k^* + \Theta_{k0}' V_k \Theta_{k0} + (\Theta_k - \Theta_k^*)' V_k^* (\Theta_k - \Theta_k^*) \\
& = S_k^{*2} + \Theta_{k0}' V_k \Theta_{k0} + (\Theta_k - \Theta_k^*)' V_k^* (\Theta_k - \Theta_k^*), \tag{A.3}
\end{aligned}$$

where S_k^{*2} is defined by (2.12). Thus the conditional posterior distribution (A.1) reduces to

$$\begin{aligned}
f(\gamma, d|Y, \sigma^2) & \propto \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-n_k/2} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \exp \left\{ -\frac{1}{2} (S_k^{*2} + \Theta_{k0}' V_k \Theta_{k0}) \right\} \\
& \quad \times \pi(\sigma^2|\gamma, d)\pi(\gamma|d) \\
& \propto \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-(n_k+v_k+2)/2} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \\
& \quad \times \exp \left\{ -\frac{1}{2} \sum_{k=1}^{K_\gamma} \left(S_k^{*2} + \Theta_{k0}' V_k \Theta_{k0} + \frac{v_k \lambda_k}{\sigma_k^2} \right) \right\} \left(\frac{\lambda}{1-\lambda} \right)^{K_\gamma} \\
& \propto \exp \{ -U(\gamma, d|Y, \sigma^2) \}, \tag{A.4}
\end{aligned}$$

where

$$U(\gamma, d|Y) = \frac{1}{2} \sum_{k=1}^{K_\gamma} (S_k^{*2} + \omega_k) + \beta K_\gamma,$$

in which ω_k and β are defined as in (2.14). The proof of Theorem 1 is completed.

Appendix 2. The proof of equation (2.17)

If $\sigma_1^2 = \dots = \sigma_k^2 = \sigma^2$ and the prior distribution of σ^2 is $IG(v/2, v\lambda/2)$, then conditional posterior distribution of (γ, d) given (Y, σ^2) is

$$\begin{aligned} f(\gamma, d|Y, \sigma^2) &\propto \int f(\Theta, \gamma, \sigma^2, d|Y) d\Theta \\ &\propto \pi(\gamma|d) \int L(\Theta, \gamma, \sigma^2, d|Y) \pi(\Theta|\gamma, d) d\Theta, \end{aligned} \quad (\text{A.5})$$

where $L(\Theta, \gamma, \sigma^2, d|Y)$ is defined by (2.15), $\pi(\Theta|\gamma, d)$ is as before. Then it follows from (A.3) the conditional posterior distribution (A.5) reduces to

$$\begin{aligned} f(\gamma, d|Y, \sigma^2) &\propto \prod_{k=1}^{K_\gamma} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \exp \left\{ -\frac{1}{2} (S_k^{*2} + \Theta'_{k0} V_k \Theta_{k0}) \right\} \pi(\gamma|d) \\ &\propto \prod_{k=1}^{K_\gamma} (2\pi)^{-(q_k+1)/2} |V_k|^{1/2} \exp \left\{ -\frac{1}{2} (S_k^{*2} + \Theta'_{k0} V_k \Theta_{k0}) \right\} \left(\frac{\lambda}{1-\lambda} \right)^{K_\gamma} \\ &\propto \exp \{ -U(\gamma, d|Y, \sigma^2) \}, \end{aligned} \quad (\text{A.6})$$

where

$$U(\gamma, d|Y) = \frac{1}{2} \sum_{k=1}^{K_\gamma} (S_k^{*2} + \omega_k) + \beta K_\gamma,$$

in which $S_k^{*2} = \frac{1}{\sigma^2} Y_k^{*'} Y_k^* - \Theta_k^{*'} V_k^* \Theta_k^*$, ω_k is defined by (2.17) and β is defined as in (2.14). The proof is completed.

References

- [1] Brooks, C. and Garrett, I. (2002) Can we explain the dynamics of the UK FTSE 100 stock and stock index futures markets. *Applied Financial Economics*, **12**, 25-31.
- [2] Chen, C. W. S. (1998). A Bayesian analysis of generalized threshold autoregressive models. *Statistics & Probability Letters* **40**, 15-22.
- [3] Chen, C.W.S., Gerlach, R., and Lin, A.M.H. (2010) Falling and explosive, dormant and rising markets via multiple-regime financial time series models. *Applied Stochastic Models in Business and Industry* **26**, 28-49.
- [4] Chen, C. W. S., Gerlach, R. and So, M. K. P. (2008). Bayesian model selection for heteroskedastic models. *Advances in Econometrics* **25**, 567-594.

- [5] Chen, C. W. S. and Lee, J. C. (1995). Bayesian inference of threshold autoregressive models. *Journal of Time Series Analysis* **16**, 483-492.
- [6] Chen, C.W.S., Lin, E.M.H., Liu, F.C., and Gerlach, R. (2008) Bayesian estimation for parsimonious threshold autoregressive models in R, *R Journal*, **8**, 26-33.
- [7] Chen, K. S. (1990). Testing for threshold autoregression. *Ann. Statist.* **18**, 1886-1894.
- [8] Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician* **49**, 327-335.
- [9] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-741.
- [10] Geweke, J. and Terui, N. (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis* **14**, 441-454.
- [11] Green, P.J. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika* **82**, 711-732.
- [12] Hansen, B.E. (2011). Threshold autoregression in economics. *Statistics and Its Interface* **4**, 123-127.
- [13] Hastings, W. K. (1970). Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97-109.
- [14] Ismail, M. A. and Charif, H. A. (2003). Bayesian Inference for threshold moving average models. *METRON* **1**, 119-132.
- [15] Lavielle, M. and Moulines, E. (2000). Least squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis* **21**(1), 33-59.
- [16] Ling, S. and Tong, H. (2005). Testing for a linear MA model against threshold MA models. *Ann. Statist.* **33**, 2529-2552.
- [17] Ling, S., Tong, H. and Li D. (2007). The ergodicity and invertibility of threshold moving-average models. *Bernoulli* **13**, 161-168.

- [18] Mcculloch, R. E. and Tsay, R. S. (1993a). Bayesian analysis of threshold autoregressive processes with a random number of regimes. *Computing Science and Statistics. Proceedings of the 25th Symposium on the Interface*, Fairfax Station, VA: Interface Foundation of North America, 253-262.
- [19] Mcculloch, R. E. and Tsay, R. S. (1993b). Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of the American Statistical Association* **88**, 968-978.
- [20] Sáfadi, T. and Morettin, P. A. (2000). Bayesian analysis of threshold autoregressive moving average models. *Sankyā* **62**, 353-371.
- [21] So, M. K. P., Chen, C. W. S. and Chen, M. T. (2005). A Bayesian threshold nonlinearity test in financial time series. *Journal of Forecasting*. **24**, 61-75.
- [22] Tong, H. (1978). On a threshold model, in C.H. Chen (eds). *Pattern Recognition and Signal Processing*, Amsterdam: Sijthoff and Noordhoff, 101-141.
- [23] Tong, H. and Lim, K. S. (1980). Threshold autoregressions, limit cycles, and data. *Journal of the Royal Statistical Society* **42**, 245-292.
- [24] Tong, H. (1983). *Threshold Model in Non-Linear Time Series Analysis*. Lecture Notes in Statistics, **21**, Springer, Berlin.
- [25] Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.
- [26] Tsay, R. S. (1989). Testing and modeling threshold autoregressive process. *Journal of the American Statistical Association* **84**, 231-240.
- [27] Tsay, R. S. (2005). *Analysis of Financial Time Series*. 2nd Edition, John Wiley & Sons.