ANALYZING THRESHOLD AUTOREGRESSIONS WITH A BAYESIAN APPROACH

Peyton Cook and Lyle D. Broemeling

ABSTRACT

A Bayesian approach to the analysis of threshold autoregressive models is presented based on the conditional likelihood function. The posterior density of the threshold is obtained numerically. The marginal posterior distributions for the remaining parameters are mixtures of familiar distributions. Posterior means and variances for the parameters can easily be obtained with simple formulas relevant to mixture distributions. Credible intervals can easily be constructed using a Monte Carlo approach, which uses exact random number generators. The Monte Carlo approach also allows for the estimation of nonlinear functions of the autoregressive coefficients. A test for the existence of a threshold is developed based on a nonlinear function of the autoregressive coefficients. A numerical example is presented that illustrates the approach.

Advances in Econometrics, Volume 11, Part B, pages 89–107.
Copyright © 1996 by JAI Press Inc.
All rights of reproduction in any form reserved.
ISBN: 1-55938-974-5

I. INTRODUCTION

Considerable development of the theory of linear time series models took place prior to the mid 1980s. Since that time, interest has shifted to the development of practical approaches to the statistical analysis of nonlinear time series. Jones (1978) sparked considerable interest in nonlinear time series models. Books on nonlinear time series models include Tong (1983, 1990) and Priestley (1988).

One type of nonlinear time series model is the threshold autoregressive process, which was first studied by Tong and Lim (1980). The two books by Tong (1983, 1990) provide a thorough treatment of the properties of threshold autoregressive processes. These processes can exhibit several types of behavior not present in linear time series models: amplitude-frequency dependence, limit cycle behavior, and jump resonance. Tong (1982) developed an approach to prediction for threshold autoregressive processes. Moeanaddin and Tong (1990) studied marginal and conditional distributions for threshold autoregressive processes. Chan and Tong (1986) used a smooth transition function from one regime to another. Tests for the existence of thresholds were presented in Chan (1990), Tong and Yeung (1990), Tong and Yeung (1991), and Tsay (1989). Threshold autoregressive models for irregularly spaced data were considered in Tong and Yeung (1990). Stationarity and stability of threshold autoregressive models were considered in Brockwell, Liu, and Tweedie (1992) and Lim (1992). The consistency of least squares estimators for threshold autoregressive models was addressed in Petruccelli (1986). Interesting alternatives to threshold autoregressive models can be found in Martin (1992) and Lye and Martin (1994).

There is great similarity between threshold time series models and change point time series models. The literature on change point models is quite large, and includes Ferreira (1975), Poirier (1976), Broemeling and Tsurumi (1987), and Krishnaiah and Miao (1988).

The statistical analyses for threshold autoregressive models listed above have primarily been based on the approach of maximum likelihood, which is justified through asymptotic considerations. Several researchers have developed feasible Bayesian approaches that do not require asymptotic approximations for the posterior distributions. Pole and Smith (1985) compared the maximum likelihood approach to the Bayesian approach for various non–time series threshold models as well as for a first-order autoregressive model. Their switching mechanism allowed for the use of an exogenous variable. Broemeling and Cook

(1992) used a Bayesian approach to estimate parameters in threshold autoregressive processes where each regime was of pth order. Their approach allowed for testing the existence of a threshold. The posterior distribution for the threshold was obtained numerically. The posterior distributions for the other parameters were presented as mixtures of familiar distributions. Geweke and Terui (1993) presented an analysis of threshold autoregressive models that focused on the prediction of future values of the time series. They used a Monte Carlo approach to obtain posterior and predictive distributions.

Broemeling and Cook (1992) presented exact posterior moments for each of the parameters. Posterior credible intervals were constructed for the parameters using Chebychev's inequality. The goal of this chapter is to compare credible intervals constructed with Chebychev's inequality with credible intervals constructed with a Monte Carlo analysis.

Section 2 specifies the threshold autoregressive model used. Section 3 specifies the likelihood function and prior distributions, and develops the various conditional and marginal posterior distributions. The various conditional and marginal posterior distributions allow one to construct exact random number generators for the posterior distributions of the parameters. Section 4 describes the estimation of linear functions of the autoregressive coefficients. Section 5 describes the one-step-ahead predictive distribution for a new measurement. Section 6 presents a numerical example illustrating the procedures. Exact posterior moments and posterior moments computed from a Monte Carlo approach are compared.

II. THE TAR(2;*p*;*d*) MODEL

This section describes a two-regime threshold autoregressive process where the order of each process is p and the delay parameter is d time units. We will indicate such a process by TAR(2;p;d).

A threshold autoregressive model with two regimes each of which has pth order is specified by

$$x(t) = \begin{cases} \theta_{10} + \Sigma_{j=1}^{p}\theta_{1j}x(t-j) + e(t), & x(t-d) \le r, \\ \theta_{20} + \Sigma_{j=1}^{p}\theta_{2j}x(t-j) + e(t), & x(t-d) > r, \end{cases} \tag{1}$$

where $x(t)$ is an observable random variable, $t = 0, \pm 1, \pm 2, \ldots, r \in \Re$ is an unobservable threshold, $\theta_1 = [\theta_{10}, \theta_{11}, \ldots, \theta_{1p}]$ and $\theta_2 = [\theta_{20}, \theta_{21}, \ldots, \theta_{2p}]$ are unobservable coefficient vectors corresponding to the two possible

autoregressive processes, d is an unobservable delay parameter, and $e(t)$ is an unobservable innovation sequence with mean zero and variance σ^2. We will assume that $e(t)$ is Gaussian. Further, we assume that the delay parameter d is known. Geweke and Terui (1993) treat the case where d is an unknown parameter whose posterior distribution is desired. The precision of the innovations will be denoted by $\tau = \sigma^{-2}$. If $x(t-d) \le r$, then $x(t)$ is determined by the first autoregressive model having coefficient vector θ_1. Otherwise $x(t)$ is determined by the second autoregressive model having coefficient vector θ_2. Both processes are constrained to have the same innovation variance. Generalizations are possible. One might include more than one threshold. Also, the different processes might be allowed to have different innovation variances. These more general models are described in Priestley (1988:73–83) and in Tong (1990:99–101). Additionally, we assume that the order of each regime, p, is known. Priestley (1988:79) and Tong (1990:281–292) discuss the order identification problem.

We wish to develop a somewhat simpler notation for handling TAR models. It will be convenient to let $\theta = [\theta_1, \theta_2]$. We wish to define an indicator function

$$i_1(t,r) = \begin{cases} 1, & x(t-d) \le r, \\ 0, & x(t-d) > r, \end{cases} \tag{2}$$

and $i_2(t,r) = 1 - i_1(t,r)$. Since we assume that d is known, we will suppress notational dependence on d for simplification. Also define square matrices that are either the identity or all zeroes, $I_j(t,r) = i_j(t,r)I_{p+1}, j = 1, 2$, where I_{p+1} is a $p+1$ by $p+1$ identity matrix. Let $I(t,r) = [I_1(t,r), I_2(t,r)]$, which is a $p+1$ by $2(p+1)$ matrix. Finally, let $z'(t) = [1, x(t-1), x(t-2), \ldots, x(t-p)]$. If it is desired to have each process have zero mean, then the leading one in $z(t)$ can be deleted. The TAR(2;p;d) model can be written as

$$x(t) = z'(t)I(t,r)\theta + e(t), \tag{3}$$

where $I(t,r)$ simply picks out θ_1 or θ_2 as needed. The notation in (3) is easily generalized to include more than one threshold.

Section 3 discusses the domain for the threshold and the creation of the likelihood function, choice of prior distributions, and posterior distributions for the parameters.

III. PRIOR AND POSTERIOR ANALYSIS

This section presents the conditional likelihood function for the TAR(2;p;d) parameters, the prior distributions for the parameters, and the various posterior distributions for the parameters.

Using (3), it is easy to write a matrix expression for a sequence of observations of the TAR(2;p;d) process. We will condition on the first p observations. Let

$$\begin{bmatrix} x(p+1) \\ x(p+2) \\ \vdots \\ x(n) \end{bmatrix} = \begin{bmatrix} z'(p+1)I(p+1,r) \\ z'(p+2)I(p+2,r) \\ \vdots \\ z'(n)I(n,r) \end{bmatrix} \theta + \begin{bmatrix} e(p+1) \\ e(p+2) \\ \vdots \\ e(n) \end{bmatrix}, \tag{4}$$

or just

$$x = H(r)\theta + e. \tag{5}$$

The conditional likelihood function is

$$f(x \mid z(p), \theta, \tau, r) \propto \tau^{(n-p)/2} \exp\left[-\frac{\tau}{2}(x - H(r)\theta)'(x - H(r)\theta)\right], \tag{6}$$

where $\tau > 0$, $\theta \in \Re^{2p+2}$, and $r \in \Re$. However, from (2) it is clear that the values of r where the likelihood function changes its values are the actual data values. We can indicate these data values by their order statistics $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$. In order to have observed data from both processes it is necessary that $x_{(1)} < r < x_{(n)}$. So, the values of r for which the conditional likelihood should be evaluated are contained in $D_r = \{r: r = x_{(i)}, i = 2,3, \ldots, n-1\}$.

We will take the approach of Box and Tiao (1973:25–60) and use locally uniform or noninformative prior distributions for the parameters. In particular,

$$\begin{aligned} f(\theta,\tau,r) &= f(\theta \mid \tau) f(\tau) f(r) \\ &\propto \tau^{-1}, \quad \theta \in \Re^{2p+2}, \quad \tau > 0, \quad r \in D_r, \end{aligned} \tag{7}$$

where we have chosen to model r as being a priori independent of (θ,τ).

By combining the likelihood function (6) and the prior probability density function (7) for the parameters, the joint posterior probability density function is

$$f(\theta,\tau,r \mid x,z(p)) \propto \tau^{(n-p)/2-1} \exp\left[-\frac{\tau}{2}(x - H(r)\theta)'(x - H(r)\theta)\right], \quad (8)$$

where $\theta \in \Re^{2(p+1)}$, $\tau > 0$, and $r \in D_r$. The structure of the density function can be seen more easily if we use a substitution. Let

$$\theta(r) = [H'(r)H(r)]^{-1}H'(r)x,$$

$$\text{wss}(r) = x'x - \theta'(r)H'(r)H(r)\theta(r),$$

$$\nu = n - 3p - 2,$$

$$P(r) = \frac{\nu}{\text{wss}(r)}H'(r)H(r). \quad (9)$$

Then, with some rearrangement, the joint probability density function for all of the parameters can be written as

$$f(\theta,\tau,r \mid x,z(p)) \propto \tau^{(n-p)/2-1}\exp\left\{\frac{\tau}{2}[\text{wss}(r) + (\theta - \theta(r))'H'(r)H'(r)(\theta - \theta(r))]\right\}, \quad (10)$$

where $\theta \in \Re^{2(p+1)}$, $\tau > 0$, and $r \in D_r$. Conditional upon r, (10) is seen to be proportional to a normal-gamma probability density function. Using results in Box and Tiao (1973:113–117) one finds that

$$\begin{aligned} f(\theta,\tau,r \mid x,z(p)) &\propto \left\{\tau^{p+1} \mid H'(r)H(r) \mid^{1/2} \times \exp\left[-\frac{\tau}{2}(\theta - \theta(r))'H'(r)H(r)(\theta - \theta(r))\right]\right\} \\ &\times \left\{[\text{wss}(r)]^{\nu/2}\tau^{\nu/2-1}\exp\left[-\frac{\tau}{2}\text{wss}(r)\right]\right\} \\ &\times \{\mid H'(r)H(r) \mid^{-1/2}[\text{wss}(r)]^{-\nu/2}\}, \\ &\propto f(\theta \mid \tau,r,x,z(p))f(\tau \mid r,x,z(p))f(r \mid x,z(p)), \quad (11) \end{aligned}$$

where $\theta \in \Re^{2(p+1)}$, $\tau > 0$, and $r \in D_r$. It should be pointed out that for some values of r, $H'(r)H(r)$ may be singular. In such cases, one simply eliminates such a value of r. This circumstance can happen when, after accounting for initial values, all observations are classified as belonging to one of the models.

From (11) it can be seen that

$$(\theta \mid \tau,r,x,z(p)) \sim \text{normal}(\theta(r), [\tau H'(r)H(r)]^{-1}), \quad (12)$$

$$(\tau \mid r,x,z(p)) \sim \text{gamma}(\nu/2, \text{wss}(r)/2), \quad (13)$$

$$f(r \mid x,z(p)) \propto \mid H'(r)H(r) \mid^{-1/2}[\text{wss}(r)]^{-\nu/2}, \qquad r \in D_r. \quad (14)$$

Further, since $(\theta,\tau \mid r,x,z(p))$ has a normal-gamma distribution, we also conclude that

$$(\theta \mid r,x,z(p)) \sim T_{2(p+1)}(\nu, \theta(r), P(r)). \quad (15)$$

The posterior distributions listed above can be found in Broemeling and Cook (1992:2466–2467), where $d = 1$ was assumed.

The posterior density for r must be analyzed numerically. Since (14) is a discrete probability mass function, one can easily compute the normalizing constant as well as moments. Moreover, it is possible to construct an exact random number generator from this discrete distribution. Geweke and Terui (1993) were first to view (14) from the point of view of generating random numbers.

Once the exact posterior mass function for r is known, it is possible to construct expressions for the posterior distributions for the remaining parameters. These posterior distributions are mixtures that are generally difficult to analyze. However, since the mixtures involve familiar distributions, it is possible to compute exact posterior moments for θ and for τ. In particular, the posterior density for τ is a mixture of gamma densities with

$$f(\tau \mid x,z(p)) = \sum_{r \in D_r} f(\tau \mid r,x,z(p))f(r \mid x,z(p)),$$

$$E[\tau \mid x,z,(p)] = \sum_{r \in D_r} \frac{\nu}{\text{wss}(r)}f(r \mid x,z(p)),$$

$$\mathrm{Var}(\tau \mid x,z(p)) = \sum_{r\in D_r} \frac{2\nu}{\mathrm{wss}(r)} f(r \mid x,z(p))$$

$$+ \sum_{r\in D_r} \left(\frac{\nu}{\mathrm{wss}(r)} - E[\tau \mid x,z(p)] \right)^2 f(r,x,z(p)). \tag{16}$$

In addition, the posterior distribution for θ is a mixture of multivariate-t densities with

$$f(\theta \mid x,z(p)) = \sum_{r\in D_r} f(\theta \mid r,x,z(p)) f(r \mid x,z(p)),$$

$$E[\theta \mid x,z(p)] = \sum_{r\in D_r} \theta(r) f(r \mid x,z(p)),$$

$$\mathrm{Var}(\theta \mid x,z(p) = \sum_{r\in D_r} \frac{\nu}{\nu-2} P^{-1}(r) f(r \mid x,z(p))$$

$$+ \sum_{r\in D_r} (\theta(r) - E[\theta \mid x,z(p)])(\theta(r) - E[\theta \mid x,z(p)])'$$

$$\times f(r \mid x,z(p)). \tag{17}$$

Moments computed with the formulas in (16) and (17) will be described as being computed with a *mixture* approach. Also, though it is an abuse of the terminology, evaluation of $f(r \mid r/x, z(p))$ and the posterior moments of r will also be described as being computed from a mixture approach. Using (14), (13), and (12) one can generate (θ,τ,r) from its joint posterior distribution using familiar random number generators. This Monte Carlo approach can be used to compute approximate posterior moments, quantiles, and credible intervals. The posterior moments can be compared to their exact values computed with (16) and (17). In addition, credible intervals computed from Chebychev's inequality can be compared to the credible intervals computed from the Monte Carlo approach.

Often one is interested in linear functions of parameters. Section 4 presents a posterior analysis for linear functions of parameters. In particular, by looking at the posterior distribution for $\theta_1 - \theta_2$ one can make inferences about the existence of a threshold.

IV. ESTIMATING LINEAR FUNCTIONS OF θ

In order to make inferences about the existence of a threshold, one may wish to estimate $\theta_1 - \theta_2$. If $\theta_1 - \theta_2 = 0$, then there is no threshold. One way to test for the existence of a threshold is to determine whether or not 0 is contained in a credible region for $\theta_1 - \theta_2$. Since $\theta_1 - \theta_2$ is simply a linear function of the autoregressive coefficients, it will be interesting to look at a slightly more general formulation. Let $\gamma = L\theta$, where L is a k by $2(p+1)$ matrix whose rank equals k. Further define

$$\gamma(r) = L\theta(r),$$

$$B(r) = L[H'(r)H(r)]^{-1}L',$$

$$P_\gamma(r) = [LP^{-1}(r)L']^{-1},$$

$$m_\gamma = LE[\theta \mid x,z(p)],$$

$$V_\gamma = L\,\mathrm{Var}(\theta \mid x,z(p))L'. \tag{18}$$

Then one easily sees that

$$\gamma \mid \tau,r,x,z(p) \sim \mathrm{normal}_k(\gamma(r), \tau^{-1}B(r)),$$

$$\gamma \mid r,x,z(p) \sim T_k(\nu, \gamma(r), P_\gamma(r)). \tag{19}$$

One can make inferences about components of γ using the result above. However, we would also like to construct a joint credible region for γ. It would seem to be difficult to construct a highest posterior density (HPD) region for γ. However, it is possible to construct a joint credible region that is not HPD. For motivation, notice that the quadratic form $(\gamma - \gamma(r))'P_\gamma^{-1}(r)(\gamma - \gamma(r))/k$ has a posterior distribution that, conditional on r, is an F distribution with k numerator and ν denominator degrees of freedom. Conditional on r, this quadratic form could be used to check whether or not γ_0 is inside a joint HPD region for γ.

In order to eliminate the conditioning on r, we wish to define the quadratic form

$$Q(\gamma) = (\gamma - m_\gamma)' V_\gamma^{-1} (\gamma - m_\gamma)/k. \tag{20}$$

Values of $Q(\gamma)$ can be simulated with a Monte Carlo approach and the quantile $q(1-\alpha)$ of order $1-\alpha$ can be determined. In particular, let

$$P(Q(\gamma) \le q(1-\alpha) \mid x,z(p)) = 1 - \alpha. \tag{21}$$

Then, γ_0 is inside a $1-\alpha$ credible region for γ if $Q(\gamma_0) \le q(1-\alpha)$. Alternatively, one might compute a posterior p-value by using the Monte Carlo approach to compute $P(Q(\gamma_0) > Q(\gamma) \mid x,z(p))$. If the posterior p-value is less than α, then γ_0 is outside a $1-\alpha$ credible region for γ.

We should remark that for the special case where $k = 1$, the use of quadratic forms is not necessary. The use of quadratic forms is important only if $k > 1$.

The next section discusses prediction of future measurements. In particular, the predictive density for an $x(n+1)$ is shown to be a mixture of t-distributions.

V. PREDICTION OF $x(n+1)$

This section discusses the prediction of the next measurement $x(n+1)$ given the past data. We will derive the one-step-ahead predictive density function and its mean and variance. As described in Geweke and Terui (1993), a Monte Carlo approach can be used to obtain a random sample from the predictive distribution. Geweke and Terui also used the Monte Carlo approach to obtain multistep ahead predictions.

The next observation is

$$x(n+1) = z'(n+1)I(n+1,r)\theta + e(n+1). \tag{22}$$

Conditional on r, this is simply in the form of the usual normal-theory linear model. The predictive distribution for the usual normal-theory linear model is given in Zellner (1971:72–74). We will mention the details only briefly. Let

$$\begin{aligned}
\hat{x}(r) &= E[x(n+1) \mid r,x,z(p)] \\
&= z'(n+1)I(n+1,r)\theta(r), \\
v_x(r) &= \mathrm{Var}(x(n+1) \mid r,x,z(p)) \\
&= \mathrm{Var}(z'(n+1)I(n+1)\theta + e(n+1) \mid r,x,z(p)) \\
&= z'(n+1)I(n+1,r)\left[\frac{\nu}{\nu-2}P^{-1}(r)\right]I'(n+1,r)z(n+1) + \frac{\mathrm{wss}(r)}{\nu-2},
\end{aligned}$$

$$p_x(r) = \frac{\nu}{\nu-2}\, v_x^{-1}(r). \tag{23}$$

Then

$$x(n+1) \mid r,x,z(p) \sim T_x(\nu, \hat{x}(r), p_x(r)), \tag{24}$$

and let the predictive density for $x(n+1)$, conditional upon r, be denoted by $f(x(n+1) \mid r,x,z(p))$. Then, $x(n+1) \mid x,z(p)$ has a distribution that is a mixture of univariate t-distributions. In particular,

$$\begin{aligned}
f(x(n+1) \mid x,z(p)) &= \sum_{r \in D_r} f(x(n+1) \mid r,x,z(p)) f(r \mid x,z(p)), \\
E[x(n+1) \mid x,z(p)] &= \sum_{r \in D_r} \hat{x}(r) f(r \mid x,z(p)), \\
\mathrm{Var}(x(n+1) \mid x,z(p)) &= E\,\mathrm{Var}(x(n+1), \mid r,x,z(p)) \\
&\quad + \mathrm{Var}(E[x(n+1) \mid r,x,z(p)]) \\
&= \sum_{r \in D_r} v_x(r) f(r \mid x,z(p)) \\
&\quad + \sum_{r \in D_r} (\hat{x}(r) - E[x(n+1) \mid x,z(p)])^2 f(r \mid x,z(p)).
\end{aligned} \tag{25}$$

Using the one-step-ahead predictive mean and variance, one can use Chebychev's inequality to construct a prediction interval, or one can generate predictions with a Monte Carlo approach. First, randomly select $r \mid x,z(p)$. Second, randomly select $\tau \mid r,x,z(p)$. Third, randomly select $\theta \mid \tau,r,x,z(p)$. Fourth, randomly select $e(n+1) \mid \tau$. Finally, one can construct the simulated value of the prediction from (22). Finally, one can then repeat this random selection process a large number of times and evaluate various posterior quantities such as moments and quantiles.

The next section describes a numerical example with generated data comparing the quadrature and Monte Carlo approaches.

VI. A NUMERICAL EXAMPLE

This section presents a numerical example comparing the quadrature and Monte Carlo approaches to analyzing the various posterior distributions. The computations were performed using the *Mathematica* programming language. The example involves a two-regime process where each regime is first order and has a nonzero mean.

Data were generated from the following model:

$$x(t) = \begin{cases} \theta_{10} + \theta_{11}x(t-j) + e(t), & x(t-d) \le r, \\ \theta_{20} + \theta_{21}x(t-j) + e(t), & x(t-d) > r, \end{cases} \tag{26}$$

where $\theta_{10} = 1$, $\theta_{11} = 0.5$, $\theta_{20} = -1$, $\theta_{21} = -0.3$, $r = 0$, $d = 1$, and $e(t) \sim$ normal(0, 1). The startup value used to begin the data generation was the mean of the first process, which is 2. The initial value introduces a transient. In order to eliminate the transient, 200 data values were generated and the first 100 values were discarded. The resulting data set contained 100 data values. Figure 1 displays a plot of the data with the points connected by a straight line. Approximately 50% of the data values are above the threshold of $r = 0$ for this example. Four separate analyses were performed. One method of analysis used the mixture approach to compute posterior moments. Using the posterior means and variances, approximate credible intervals with *at least* 80% coverage were computed using Chebychev's inequality. The other method of analysis used a Monte Carlo approach to approximate the posterior moments and quantiles for the various posterior distributions. The Monte Carlo approach used 5000 random selections of $\theta, \tau, r \mid x, z(1)$. That is, (i) a value of r was randomly selected from $f(r \mid x, z(1))$, (ii) a value of $\tau \mid r, x, z(1)$ was randomly selected from $f(\tau \mid r, x, z(1))$, (iii) a value of $\theta \mid \tau, x, z(1)$ was randomly selected from $f(\theta \mid \tau, r, x, z(1))$. This was repeated a total of 5000 times. Coverage of 80% for the credible intervals was chosen in order to avoid using the Monte Carlo approach to approximate extreme quantiles. Further, each of the two methods of analysis was applied to the first $n = 20$ observations (after discarding the first 100 observations) as well as all $n = 100$ of the observations (after discarding the first 100 observations). Finally, a posterior p-value was computed in an attempt to determine whether or not a threshold is in effect.

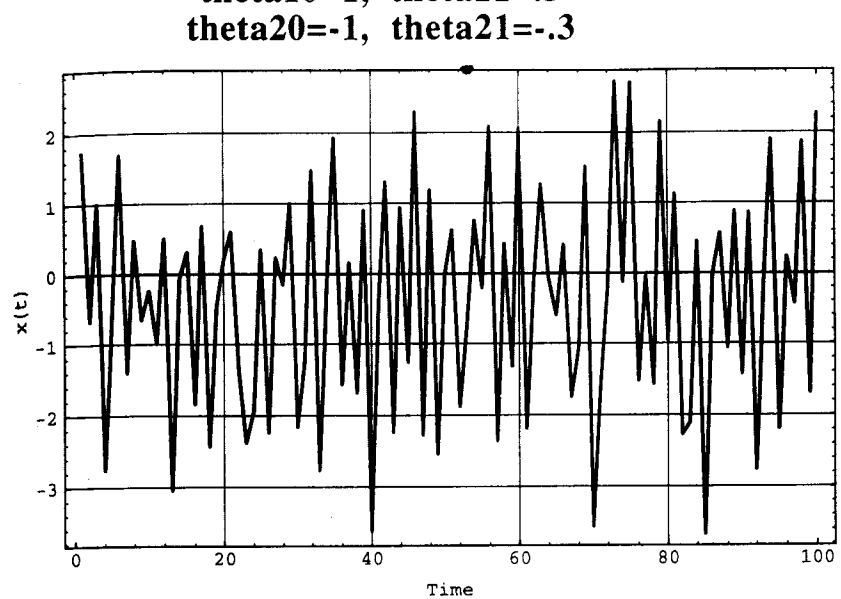

Figure 1. Simulated data for a TAR(2; 1; 1) model.

Figure 2 displays the posterior pdf for the threshold r for each of the two sample sizes of $n = 20$ and $n = 100$. Values of the posterior pdf for the threshold based on the smaller sample size of $n = 20$ are plotted as large dots. With the smaller sample size $f(r \mid x, z(1))$ seems to be spread out from about –0.5 to 0.5. The posterior pdf for the threshold r based on $n = 100$ points is displayed by connecting the points. As one would expect, the posterior pdf for r is much more concentrated for the larger sample size of $n = 100$. One can see that the pdfs are skewed to the left.

Table 1 displays various descriptive measures of the posterior pdf for r. The second two columns used the first $n = 20$ observations, and the last two columns used all $n = 100$ observations. The columns headed by M contain values computed with the mixture approach. The quantiles computed by the mixture approach are only approximate since they were computed with Chebychev's inequality as $E[r \mid x, z(1)] \pm 2.24\text{sd}(r \mid x, z(1))$, where $\text{sd}(\cdot \mid x, z(1))$ indicates a posterior standard deviation. The posterior median was not computed by the mixture ap-

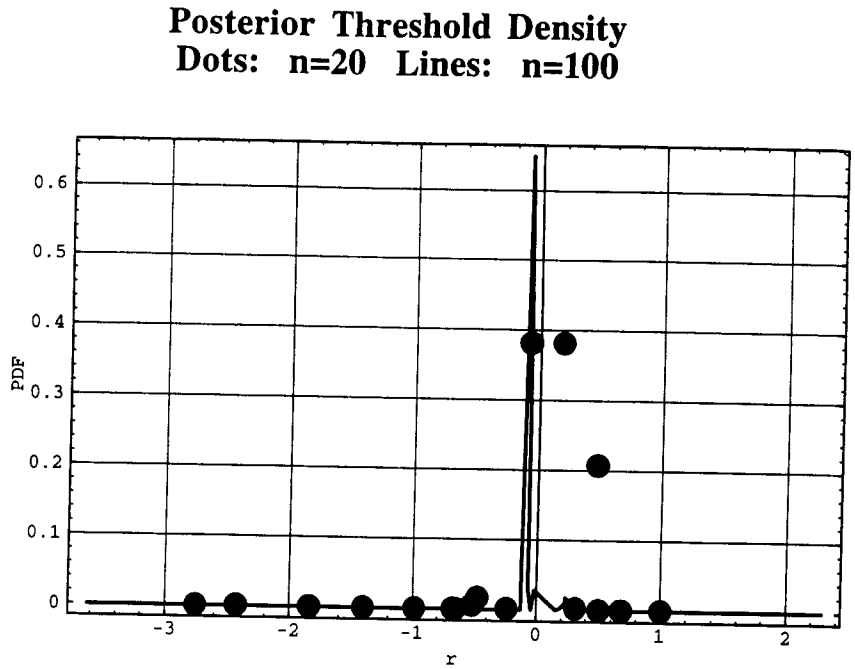

Figure 2. Posterior probability density for the threshold r. The true value of the threshold is 0.

We emphasize that since r is discrete and we have an explicit formula for its posterior pdf, the posterior moments for r are exact up to round-off error. The columns headed by MC contain values computed by Monte Carlo with 5000 replications. Posterior quantiles computed by Chebychev's inequality are conservative. One expects that the credible interval determined by the 0.10 and 0.90 posterior quantiles using Chebychev's inequality will contain the corresponding interval determined by the 0.10 and 0.90 posterior quantiles computed by the Monte Carlo approach. This anticipated result is present in Table 1 for each of the two sample sizes. (This is not true for some of the other parameters.) Three out of the four credible intervals contain the true value of $r = 0$. When $n = 100$ the Monte Carlo approximations for the quantiles, based on 5000 replications, did not contain 0. (We note that a credible interval with posterior content of 0.95 may have contained 0.)

Table 2 displays posterior results for θ_{10} whose true value is 1. The format for Table 2 is similar to that of Table 1. The second two columns contain the mixture and Monte Carlo analyses for a sample size of $n =$

Table 1. Posterior Results for the Threshold r^a

Sample size	*n = 20*		*n = 100*	
Method	*M*	*MC*	*M*	*MC*
Mean	.127	.125	–.073	–.073
Variance	.061	.057	.003	.003
Standard deviation	.248	.238	.056	.056
0.10 quantile	–.428	–.075	–.199	–.108
0.90 quantile	.682	.476	.053	–.075
Median	NA	.186	NA	–.075

Note: [a]The true value of r is 0.

20 and the last two columns contain mixture and Monte Carlo analyses for $n = 100$. There were 5000 replications in the Monte Carlo analysis. Generally, the mixture and Monte Carlo results were similar. However, the posterior variances (and standard deviations) when $n = 20$ were quite noticeably different. Since the quantiles computed from the mixture and the Monte Carlo approaches were similar, we suspect that the simulated distribution had outliers that affected the Monte Carlo approximation for the variance and standard deviation. Each of the 80% credible intervals defined by the 0.10 and 0.90 quantiles contained the true value of $\theta_{10} = 1$, except for the case with $n = 100$ using the Monte Carlo approach. The posterior means and medians are quite similar, indicating that the posterior pdf for θ_{10} is not noticeably skewed.

Table 3 displays posterior results for θ_{11} whose true value is 0.5. For all entries, the 80% credible intervals all contained the true value of 0.5. The results for mixture and Monte Carlo were similar. The posterior

Table 2. Posterior Results for $\theta_{10}{}^a$

Sample size	*n = 20*		*n = 100*	
Method	*M*	*MC*	*M*	*MC*
Mean	.358	.368	1.466	1.468
Variance	.345	1.575	.084	.085
Standard deviation	.588	1.255	.290	.291
0.10 quantile	–.958	–1.022	.817	1.095
0.90 quantile	1.674	1.783	2.11	1.834
Median	NA	.346	NA	1.475

Note: [a]The true value of θ_{10} is 1.

Table 3. Posterior Results for θ_{11}[a]

Sample size	n = 20		n = 100	
Method	M	MC	M	MC
Mean	.143	.151	.645	.646
Variance	.098	.558	.025	.025
Standard deviation	.313	.747	.157	.158
0.10 quantile	–.559	–.737	.294	.443
0.90 quantile	.846	1.050	.997	.843
Median	NA	.158	NA	.650

Note: [a]The true value of θ_{11} is 0.5.

means and medians are quite similar indicating that there is very little skew.

Table 4 contains the posterior results for θ_{20} whose true value is –1. The mixture and Monte Carlo approaches are similar except for the posterior variance and standard deviation with the smaller sample size of $n = 20$. Each of the 80% credible intervals contains the true parameter value of –1. The posterior means and medians are quite similar, indicating that there is very little skew.

Table 5 contains posterior results for θ_{21} whose true value is –0.3. The posterior variance for θ_{21} computed from mixture and Monte Carlo with 5000 replications are quite different for $n = 20$. However, the credible intervals are similar.

Table 6 displays posterior results for the precision τ whose true value is 1. The mixture and Monte Carlo results are generally similar. All of the 80% credible intervals contain the true value of 1.

Table 4. Posterior Results for θ_{20}[a]

Sample size	n = 20		n = 100	
Method	M	MC	M	MC
Mean	–2.687	–2.673	–1.295	–1.305
Variance	3.501	34.151	.098	.100
Standard deviation	1.871	5.844	.314	.316
0.10 quantile	–6.888	–5.371	–1.998	–1.672
0.90 quantile	1.494	–.319	–.592	–.929
Median	NA	–2.613	NA	–1.285

Note: [a]The true value of θ_{20} is –1.

Table 5. Posterior Results for θ_{21}[a]

Sample size	n = 20		n = 100	
Method	M	MC	M	MC
Mean	.824	.827	–.300	–.298
Variance	1.497	13.385	.055	.056
Standard deviation	1.223	3.658	.234	.237
0.10 quantile	–1.916	–1.375	–.825	–.583
0.90 quantile	3.565	3.220	.225	–.004
Median	NA	.780	NA	–.310

Note: [a]The true value of θ_{21} is –0.3.

For $n = 20$ and for $n = 100$, a test of the existence of a threshold was conducted. This posterior p-value was based on 5000 Monte Carlo replications. As indicated earlier, this test examines the similarity of the two estimated processes. The null hypotheses is H_0: $\theta_1 - \theta_2 = 0$ versus H_a: $\theta_1 - \theta_2 \neq 0$. One interpretation of the posterior p-value is as follows: Construct a credible region for $\theta_1 - \theta_2$ with 0 on the boundary. The posterior content outside the credible region is the p-value. For $n = 20$, the posterior p-value was 0.465. The p-value was computed as $=P(Q(\gamma) > Q(0) \mid y)$, where $\gamma = \theta_1 - \theta_2$. Such a large posterior p-value would seem to suggest that no threshold process was in operation. Of course, such a conclusion would be in error for this example. For $n = 100$, the posterior p-value was 0. None of the 5000 randomly selected values of $\theta_1 - \theta_2$ fell outside the credible region for $\theta_1 - \theta_2$ having 0 on the boundary. Such a small p-value would seem to indicate that a threshold process was in fact in operation.

Table 6. Posterior Results for the Precision τ[a]

Sample size	n = 20		n = 100	
Method	M	MC	M	MC
Mean	1.447	1.443	.906	.903
Variance	.284	.274	.017	.017
Standard deviation	.533	.524	.132	.131
0.10 quantile	.254	.821	.610	.740
0.90 quantile	2.640	2.151	1.20	1.076
Median	NA	1.386	NA	.901

Note: [a]The true value for the precision τ is 1.

VII. SUMMARY

We presented a Bayesian approach to estimating the parameters of a threshold process with two regimes. The posterior pdf for the threshold must be calculated numerically. However, this is straightforward because the threshold is treated as discrete. The exact posterior distributions of the other parameters were presented as mixtures of familiar distributions: normal-gamma, t, and gamma. The mixing constants were the probabilities from the posterior pdf of the threshold. While exact values of the posterior moments are fairly easy to compute, quantiles are much more difficult to compute for mixture distributions. However, a Monte Carlo approach was used in order to approximate the posterior quantiles. The one-step-ahead predictive density for a new observation was also presented as a mixture distribution. Formulas were given for its posterior moments. A numerical example was presented that illustrated the approach. The example used 5000 Monte Carlo replications, which seemed adequate for the most part. However, there were instances when 5000 replications produced posterior variances very different from the exact or Mixture approach. On the other hand, in all cases the posterior quantiles were comparable. Finally, a test for the existence of a threshold process was conducted. The uncertainty of parameter estimation was large when the sample size was small and the threshold was not detected. However, the threshold was readily detected for the larger sample size.

We conclude that the Monte Carlo approach to analyzing mixtures works quite well for our threshold processes. Care must always be used in selecting the number of replications when using Monte Carlo.

REFERENCES

Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Reading, MA: Addison-Wesley.

Brockwell, P. J., Liu, J., and R. L. Tweedie (1992). "On the Existence of Stationary Autoregressive Moving-Average Processes," *Journal of Time Series Analysis* 13(2), 95–107.

Broemeling, L. D. and P. Cook (1992). "Bayesian Analysis of Threshold Autoregressions," *Communications in Statistics—Theory and Methods* 21(9), 2459–2482.

Broemeling, L. D. and H. Tsurumi (1987). *Econometrics and Structural Change*. New York: Dekker.

Chan, K. S. (1990). "Testing for Threshold Autoregression," *Annals of Statistics* 18(4), 1886–1894.

Chan, K. S. and H. Tong (1986). "On Estimating Thresholds in Autoregressive Models," *Journal of Time Series Analysis* 7(3), 179–190.

Ferreira, P. E. (1975). "A Bayesian Analysis of a Switching Regression Model: Known Number of Regime," *Journal of The American Statistical Association* 70, 370–374.

Geweke, J. and N. Terui (1993). "Bayesian Threshold Autoregressive Models for Nonlinear Time Series," *Journal of Time Series Analysis* 14(5), 441–454.

Jones, D. A. (1978). "Non-Linear Autoregressive Processes," *Proceedings of the Royal Society* A360, 711–95.

Krishnaiah, P. R. and B. Q. Miao (1988). "Review About Estimation of Change Points." Pp. 375–402 in *Handbook of Statistics*, edited by P. R. Krishnaiah and C. R. Rao. Amsterdam: Elsevier.

Lim, K. S. (1992). "On the Stability of a Threshold AR(1) Without Intercepts," *Journal of Time Series Analysis* 139(2), 119–132.

Lye, J. N. and V. L. Martin (1994). "Non-Linear Time Series Modeling and Distributional Flexibility," *Journal of Time Series Analysis* 15(1), 65–84.

Martin, V. L. (1992). "Threshold Time Series Models as Multimodal Distribution Jump Processes," *Journal of Time Series Analysis* 13(1), 80–94.

Moeanaddin, R. and H. Tong (1990). "Numerical Evaluation of Distribution in Non-Linear Autoregression," *Journal of Time Series Analysis* 11, 33–48.

Petruccelli, J. D. (1986). "On the Consistency of Least Squares Estimators for a Threshold AR(1) Model," *Journal of Time Series Analysis* 7(4), 269–278.

Pole, A. M. and A. F. M. Smith (1985). "A Bayesian Analysis of Some Threshold Switching Models," *Journal of Econometrics* 29, 97–119.

Poirier, D. J. (1976). *The Econometrics of Structural Change*. Amsterdam: North-Holland.

Priestley, M. B. (1988). *Non-Linear and Non-Stationary Time Series Analysis*. London: Academic Press.

Tong, H. (1982). "A Note on Using Threshold Autoregressive Models for Multi-Step-Ahead Prediction of Cyclical Data," *Journal of Time Series Analysis* 3, 137–140.

Tong, H. (1983). *Threshold Models in Non-Linear Time Series*. New York: Springer-Verlag.

Tong, H. (1990). *Non-Linear Time Series, A Dynamical Approach*. Oxford: Oxford Science.

Tong, H. and K. S. Lim (1980). "Threshold Autoregression, Limit Cycles, and Cyclical Data" (with discussion), *Journal of the Royal Statistical Society* B42, 245–292.

Tong, H. and I. Yeung (1990). "On Tests for Threshold-Type Nonlinearity in Irregularly Spaced Time Series," *Journal of Statistical Computing and Simulation* 34, 177–194.

Tong, H. and I. Yeung (1991). "On Tests for Self-Exciting Threshold Autoregressive-Type Non-Linearity in Partially Observed Time Series," *Statistician* 1, 43–62.

Tsay, R. S. (1989). "Testing and Modeling Threshold Autoregressive Processes," *Journal of the American Statistical Association* 84(405), 231–240.

Zellner, A. (1971). *Bayesian Inference in Econometrics*. New York: Wiley.