

CONTINUOUS TIME THRESHOLD AUTOREGRESSIVE MODELS

Peter J. Brockwell, Rob J. Hyndman and Gary K. Grunwald

*Colorado State University, University of Melbourne and
University of Melbourne*

Abstract: The importance of non-linear models in time series analysis has been recognized increasingly over the past ten years. A number of discrete time non-linear processes have been introduced and found valuable for the modelling of observed series. Among these processes are the discrete time threshold models, discussed extensively in the book of Tong (1983). The purpose of this paper is to define a continuous time analogue of the threshold AR(p) process and to discuss some of its properties. For the continuous time threshold AR(1) process (henceforth denoted CTAR(1)) we derive the stationary distribution (under appropriate assumptions) and consider problems of prediction and inference. The techniques developed apply equally well both to regularly and to irregularly spaced observations.

Key words and phrases: Non-linear model, stationary distribution, prediction, Gaussian likelihood.

1. Introduction

A time series $\{X_t\}$ is a discrete self-exciting threshold autoregressive (SE TAR) process with delay parameter d if it satisfies the equations

$$X_t = a_0^{(i)} + \sum_{j=1}^p a_j^{(i)} X_{t-j} + \sigma^{(i)} e_t, \quad r_{i-1} \leq X_{t-d} < r_i, \quad (1.1)$$

where $-\infty = r_0 < r_1 < \dots < r_l = \infty$, $a_j^{(i)}$ and $\sigma^{(i)} (> 0)$, $i = 1, \dots, l$, are constants, and $\{e_t\}$ is a white noise sequence with unit variance. The thresholds are the levels r_1, \dots, r_{l-1} . Thus, the real line is partitioned into l intervals, and X_t satisfies one of l autoregressive equations depending on the interval in which X_{t-d} falls. When $l = 1$, $\{X_t\}$ is an AR process.

The model (1.1) is capable of reproducing features which are frequently observed in naturally occurring time series but which cannot be reproduced by linear Gaussian models, e.g. time-irreversibility and the occurrence of occasional bursts of outlying observations. For a comprehensive account of the use of threshold models see Tong (1983) and Tong and Lim (1980). The threshold

AR(1) process is discussed in considerable detail by Chan et al. (1985), who derive necessary and sufficient conditions for existence of a stationary distribution.

There are several good reasons for investigating continuous time analogues of (1.1). For a continuously evolving process observed at discrete time intervals it is natural to allow the possibility of a model change at *any* time, possibly between observations. It would also be useful to allow a delay parameter, d , which is not necessarily an integer multiple of the observation spacing, although in this paper we shall confine ourselves to the case of zero delay (which means that the state vector $S(t)$ of Equation (2.4) is Markovian). As in Jones (1981), where linear continuous time autoregressive processes are used for modelling observations in discrete time, the continuous time model is particularly well suited to dealing with missing or irregularly spaced observations. This is demonstrated in Section 4 where we show how to compute the Gaussian likelihood for a finite set of arbitrarily spaced observations.

As we shall see, the analysis of continuous time threshold autoregressive models can be reduced to the study of diffusion processes with piecewise linear coefficients. Stationary distributions and conditional expectations can be computed (numerically at least) from the forward and backward Kolmogorov equations with appropriate boundary conditions. The calculations are illustrated in the examples.

2. Continuous Time Autoregressive Models

Before studying continuous time threshold autoregressive (CTAR) processes, we shall quickly review the corresponding linear CAR models.

We define a zero-mean CAR(1) process to be a stationary solution of the stochastic differential equation,

$$dX(t) = aX(t)dt + \sigma dW(t), \quad -\infty < t < \infty, \quad (2.1)$$

where $\{W(t)\}$ denotes standard Brownian motion and $a < 0$. Under the latter assumption, Equation (2.1) has the unique stationary solution

$$X(t) = \sigma \int_{-\infty}^t \exp[a(t-y)]dW(y). \quad (2.2)$$

Equivalently, a zero-mean CAR(1) process can be defined as a stationary diffusion process with drift coefficient $\alpha(x) = ax$ and diffusion coefficient $\beta(x) = \sigma^2/2$, i.e. a stationary Ornstein-Uhlenbeck process.

A CAR(1) process with mean $\mu = -b/a$ is a stationary solution of

$$dY(t) = (aY(t) + b)dt + \sigma dW(t), \quad -\infty < t < \infty, \quad (2.3)$$

where again we assume that $a < 0$. Equivalently we can say that $\{Y_t\}$ is a stationary diffusion process with drift coefficient $\alpha(x) = ax + b$ and diffusion coefficient $\beta(x) = \sigma^2/2$.

In order to define a zero-mean CAR(p) process we introduce the stationary p -variate process $\{S(t)\}$ satisfying the stochastic differential equation,

$$dS(t) = AS(t)dt + e\sigma dW(t), \tag{2.4}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \dots & -a_1 \end{bmatrix}, \quad e = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

and, to ensure stability, we assume that the zeros of $a(z) = z^p + a_1z^{p-1} + \dots + a_pz^0$ all have negative real parts. The zero-mean CAR(p) process $\{X_t\}$ is then defined to be the first component of the process $\{S(t)\}$, i.e.

$$X(t) = [1 \ 0 \ \dots \ 0]S(t). \tag{2.5}$$

Equation (2.4) is simply a state-space version of the equation

$$\frac{d^p X(t)}{dt^p} + a_1 \frac{d^{p-1} X(t)}{dt^{p-1}} + \dots + a_p X(t) = \sigma \frac{dW(t)}{dt},$$

in which the j th component of $S(t)$ is $\frac{d^{j-1} X(t)}{dt^{j-1}}$, $j = 1, \dots, p$. Even though $\frac{dW(t)}{dt}$ is not defined in the usual sense, the p th-order differential equation for $\{X(t)\}$ acquires a precise interpretation when it is written in the state-space form (2.4).

We can also specify $\{X(t)\}$ as the first component of the stationary multivariate diffusion process $\{S(t)\}$ with drift coefficient vector $\mathbf{a}(\mathbf{x}) = A\mathbf{x}$ and diffusion coefficient matrix $B(\mathbf{x}) = \sigma^2 \mathbf{e}\mathbf{e}'/2$. A CAR(p) process with mean $\mu = b/a_p$ is defined analogously, by substituting $b dt + \sigma dW(t)$ for $\sigma dW(t)$ on the right of Equation (2.4).

3. The Continuous Time Threshold AR(1) Process

We define the CTAR(1) process to be a stationary solution of the stochastic differential equations

$$dX(t) = (a^{(i)}X(t) + b^{(i)})dt + \sigma^{(i)}dW(t), \quad r_{i-1} < X(t) < r_i, \quad i = 1, \dots, l, \tag{3.1}$$

where $-\infty = r_0 < r_1 < \dots < r_l = \infty$, $a^{(1)} < 0$, $a^{(i)} < 0$, each $\sigma^{(i)} > 0$ and $b^{(1)}, \dots, b^{(l)}$ are constants. The thresholds are the levels r_1, \dots, r_{l-1} . If $l = 1$

then $\{X(t)\}$ is a CAR(1) process. The process defined by (3.1) is a diffusion process with drift and diffusion coefficients

$$\alpha(x) = \sum_{i=1}^l (a^{(i)}x + b^{(i)})I_{(r_{i-1}, r_i)}(x), \tag{3.2}$$

and

$$\beta(x) = \sum_{i=1}^l \frac{\sigma^{(i)2}}{2} I_{(r_{i-1}, r_i)}(x), \tag{3.3}$$

respectively, where I_A denotes the indicator function of the set A .

The transition function of $\{X(t)\}$ cannot be determined until we specify the boundary behaviour at the thresholds r_1, \dots, r_{l-1} . We do this by specifying that the functions in the domain $\mathcal{D}(\mathcal{G})$ of the generator \mathcal{G} of $\{X(t)\}$ satisfy the condition,

$$\sqrt{\beta(r_i-)}f'(r_i-) = \sqrt{\beta(r_i+)}f'(r_i+), \quad i = 1, \dots, l-1, \quad \text{for all } f \in \mathcal{D}(\mathcal{G}). \tag{3.4}$$

For $f \in \mathcal{D}(\mathcal{G})$ the generator of $\{X(t)\}$ is defined by

$$\mathcal{G}f(x) = \alpha(x)f'(x) + \beta(x)f''(x), \quad x \notin \{r_1, \dots, r_{l-1}\}, \tag{3.5}$$

with $\mathcal{G}f(x)$, $x \in \{r_1, \dots, r_{l-1}\}$, determined by continuity of $\mathcal{G}f(\cdot)$.

In the constant variance case, $\sigma^{(i)} = \sigma$, $i = 1, \dots, l$, the condition (3.4) reduces to continuity of f' . This is the condition used by Atkinson and Caughey (1968) in their determination of the spectral density for this case. The more general condition is obtained from the solution $\{X^{(n)}(t), t \geq 0\}$ of Equation (3.1) with $\{W(t)\}$ replaced by $I^{(n)}(t) = \int_0^t Y^{(n)}(u)du$, where $\{Y^{(n)}(t)\}$ is the continuous time Markov process with state-space $\{-\sqrt{n}, \sqrt{n}\}$ and generator $\begin{bmatrix} -n & n \\ n & -n \end{bmatrix}$. It is known (see Brockwell, Resnick and Pacheco-Santiago (1982)) that $\{I^{(n)}(t)\}$ converges weakly to $\{W(t)\}$ as $n \rightarrow \infty$. Consideration of the generator of the process $\{(X^{(n)}(t), Y^{(n)}(t))\}$ as $n \rightarrow \infty$ leads to the condition (3.4). (Note that a different process is obtained if (3.4) is replaced by continuity of f' .)

Proposition 3.1. *Suppose that $\sigma^{(i)} > 0$, $i = 1, \dots, l$. Then the process $\{X(t)\}$ defined by (3.1) and (3.4) has a stationary distribution if and only if $\lim_{x \rightarrow -\infty} (a^{(1)}x^2 + 2b^{(1)}x) < 0$ and $\lim_{x \rightarrow \infty} (a^{(l)}x^2 + 2b^{(l)}x) < 0$.*

Proof. If either of the two conditions is violated, if $[a, b]$ is any finite interval and if $0 < \epsilon < (b - a)/2$, then either the expected passage time of $X(t)$ from $b + \epsilon$ to $b - \epsilon$ or the expected passage time from $a - \epsilon$ to $a + \epsilon$ is infinite. On the other hand the expected time, starting from any state $x \in [a + \epsilon, b - \epsilon]$, for

$X(t)$ to reach either $a - \epsilon$ or $b + \epsilon$ is bounded. A simple renewal theory argument therefore shows that there can be no stationary distribution for $X(t)$, since such a distribution would necessarily assign probability zero to the interval $[a + \epsilon, b - \epsilon]$ and hence to every finite interval.

If both conditions are satisfied we show that there is a stationary distribution by computing it explicitly. To do this we note (see, e.g., Breiman (1968), p.346) that $\pi(x)$ is a stationary density for $\{X(t)\}$ if and only if

$$\int_{-\infty}^{\infty} \mathcal{G}f(x)\pi(x)dx = 0 \quad \text{for all } f \in \mathcal{D}(\mathcal{G}). \tag{3.6}$$

Substituting from (3.5), integrating by parts, using (3.4) and assuming that π has continuous second derivatives except at the thresholds, we find that π satisfies the forward Kolmogorov equation,

$$\alpha(x)\pi'(x) + \alpha'(x)\pi(x) - \beta(x)\pi''(x) = 0, \quad x \notin \{r_1, \dots, r_{l-1}\}, \tag{3.7}$$

with boundary conditions (for $i = 1, \dots, l - 1$),

$$\sqrt{\beta(r_i-)}\pi(r_i-) = \sqrt{\beta(r_i+)}\pi(r_i+), \tag{3.8}$$

and

$$\alpha(r_i-)\pi(r_i-) - \beta(r_i-)\pi'(r_i-) = \alpha(r_i+)\pi(r_i+) - \beta(r_i+)\pi'(r_i+), \tag{3.9}$$

where α and β are defined in (3.2) and (3.3). Integrating (3.7) we find that

$$(a^{(i)}x + b^{(i)})\pi(x) - \sigma^{(i)2}\pi'(x)/2 = c^{(i)}, \quad r_{i-1} < x < r_i.$$

Integrability of π implies that $c^{(1)} = 0$ and (3.9) then implies that $c^{(i)} = 0$ for all i . Integrating again we find that

$$\pi(x) = k_i \exp[(a^{(i)}x^2 + 2b^{(i)}x)/\sigma^{(i)2}], \quad r_{i-1} < x < r_i, \quad i = 1, \dots, l, \tag{3.10}$$

where k_1, k_2, \dots, k_l are determined by (3.8) and the fact that $\int_{-\infty}^{\infty} \pi(x)dx = 1$.

Example 3.1. Consider the CTAR(1) process,

$$\begin{aligned} dX(t) + 0.50X(t)dt &= 0.50dW(t), & \text{if } X(t) < 0; \\ dX(t) + 1.00X(t)dt &= dW(t), & \text{if } X(t) \geq 0. \end{aligned}$$

From (3.10) we immediately find that the stationary density of $\{X(t)\}$ is

$$\pi(x) = \begin{cases} 2ce^{-2x^2}, & x < 0, \\ ce^{-x^2}, & x > 0, \end{cases}$$

where $c = (2\sqrt{2} - 2)/\sqrt{\pi} \approx 0.46739$.

Example 3.2. Consider the two-threshold CTAR(1) process,

$$\begin{aligned} dX(t) + 0.18 X(t)dt &= 1.2 dW(t), & \text{if } X(t) < -0.5; \\ dX(t) + 0.50 X(t)dt &= dW(t), & \text{if } -0.5 \leq X(t) < 0.5; \\ dX(t) + 0.80 X(t)dt &= 0.4 dW(t), & \text{if } X(t) \geq 0.5. \end{aligned}$$

From (3.10) the stationary density of $\{X(t)\}$ is

$$\pi(x) = \begin{cases} ce^{-0.09375 - .125x^2}, & x < -0.5, \\ 1.2ce^{-0.5x^2}, & -0.5 < x < 0.5, \\ 3ce^{1.125 - 5x^2}, & x > 0.5, \end{cases}$$

where $c \approx 0.2940$.

4. Prediction and Inference for CTAR(1) Processes

We first consider the problem of finding the minimum mean squared error predictor $\hat{X}_n(n+h)$ of $X(n+h)$ based on $\{X(s), s \leq n\}$. By the Markov property of $\{X(t)\}$ this is simply the conditional expectation $E(X(n+h)|X(n))$. (This is also the best predictor based on any finite set of observations of which the last is $X(n)$.) We can express $E(X(n+h)|X(n))$ in the form $m(X(n), h)$ where

$$m(x, t) = E(X(t)|X(0) = x).$$

In the same way we can express the mean squared error of the predictor, conditional on $X(n)$, as $v(X(n), h)$ where

$$v(x, t) = s(x, t) - m(x, t)^2 \quad \text{and} \quad s(x, t) = E(X(t)^2|X(0) = x).$$

In order to determine the functions m and s numerically, we solve the backward Kolmogorov equations for the conditional characteristic function, $\phi(x, t) = E_x e^{i\theta X(t)}$, namely

$$\frac{\partial \phi}{\partial t} = \alpha(x) \frac{\partial \phi}{\partial x} + \beta(x) \frac{\partial^2 \phi}{\partial x^2}, \quad x \notin \{r_1, \dots, r_{l-1}\},$$

with initial condition $\phi(x, 0) = e^{i\theta x}$ and boundary conditions

$$|\phi| \leq 1 \quad \text{and} \quad \sqrt{\beta(r_i-)} \frac{\partial \phi}{\partial x}(r_i-) = \sqrt{\beta(r_i+)} \frac{\partial \phi}{\partial x}(r_i+), \quad i = 1, \dots, l-1.$$

The functions m and s are determined from the first and second partial derivatives of ϕ with respect to θ at $\theta = 0$.

Example 4.1. For the CTAR(1) process of Example 3.1 the functions $m(x, t)$ and $v(x, t)$ for $t = 1, 2$ and 5 are shown in Figures 1 and 2 respectively. As expected, the values of $m(x, t)$ and $v(x, t)$ for $t > 10$ are very close to the mean, 0, and variance, .35355, of the stationary distribution in Example 3.1. Notice also that for small t , $m(x, t)$ is reasonably well approximated by the solution $m^*(x, t) = x \exp(-t)$, $x > 0$, $m^*(x, t) = x \exp(-.5t)$, $x < 0$, of the corresponding defining equations with the white noise terms set equal to zero. In general however, this approximation will not be good for large values of the lead time, t .

Maximum Gaussian likelihood estimation of parameters for discrete time ARMA processes based on the observations $\{X_1, \dots, X_n\}$ is equivalent (see e.g. Brockwell and Davis (1987)) to maximization of the Gaussian likelihood of the linear innovations, i.e. of

$$L = (2\pi)^{-n/2} (v_0 v_1 \dots v_{n-1})^{-1/2} \exp \left\{ - \sum_{i=1}^n \frac{(X_i - \hat{X}_i)^2}{2v_{i-1}} \right\},$$

where $\hat{X}_1 = EX_1$, \hat{X}_i , $i \geq 2$, is the minimum mean-squared error *linear* predictor of X_i in terms of $1, X_1, \dots, X_{i-1}$, $v_0 = \text{Var}(X_1)$ and $v_j = E(X_{j+1} - \hat{X}_j)^2$, $j \geq 1$. A natural analogue of this procedure for the CTAR(1) process observed at a finite set of times $\{t_1, \dots, t_n\}$ is maximization of the Gaussian likelihood defined as follows. Letting $\hat{X}(t_1) = EX(t_1)$, $\hat{X}(t_i) = m(X(t_{i-1}), t_i - t_{i-1})$, $i \geq 2$, $v_0 = \text{Var}(X(t_1))$ and $v_i = v(X(t_i), t_{i+1} - t_i)$, $i \geq 1$, we define the Gaussian likelihood L^* by supposing the distribution of $X(t_1)$ and the transition densities to be Gaussian. Thus

$$L^* = (2\pi)^{-n/2} (v_0 v_1 \dots v_{n-1})^{-1/2} \exp \left\{ - \sum_{i=1}^n \frac{(X(t_i) - \hat{X}(t_i))^2}{2v_{i-1}} \right\}.$$

For a given CTAR(1) model and a given data set, the function L^* is easily calculated from the stationary distribution and the functions m and v discussed above. Properties of the estimators obtained by maximization of L^* will be investigated in a subsequent paper. We note however that the expression for L^* is valid both for regularly and irregularly spaced observations. (Calculation of the true likelihood of the observations is also possible from the marginal (stationary) distribution and transition probabilities of $\{X(t)\}$, however the computation of the latter is considerably more difficult.)

5. The CTAR(p) Process

Analysis of the CTAR(p) process is more complicated when $p > 1$ since it involves the analysis of a p -dimensional diffusion process. We shall therefore restrict ourselves, here, to specification of the process. A detailed study of its

properties will be given in a subsequent paper. A CTAR(p) process is defined (cf. the CAR(p) process) as the first component of a stationary p -dimensional series $\mathbf{S}(t)$ satisfying the equations,

$$d\mathbf{S}(t) = A^{(i)}\mathbf{S}(t)dt + \mathbf{e}(b^{(i)}dt + \sigma^{(i)}dW(t)), \quad r_{i-1} < S_1(t) < r_i,$$

where $-\infty = r_0 < r_1 < \dots < r_l = \infty$, $S_1(t)$ denotes the first component of $\mathbf{S}(t)$ and each $\sigma^{(i)} > 0$. The matrices $A^{(i)}$ and vector \mathbf{e} have the form,

$$A^{(i)} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_p^{(i)} & -a_{p-1}^{(i)} & -a_{p-2}^{(i)} & \dots & -a_1^{(i)} \end{bmatrix} \quad \text{and} \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

The process $\{\mathbf{S}(t)\}$ is a diffusion process with drift and diffusion coefficients

$$\mathbf{a}(\mathbf{x}) = \sum_{i=0}^{l-1} (A^{(i)}\mathbf{x} + \mathbf{e}b^{(i)})I_{(r_{i-1}, r_i)}(x_1),$$

and

$$B(\mathbf{x}) = \sum_{i=0}^{l-1} \frac{\sigma^{(i)2}}{2} \mathbf{e}\mathbf{e}' I_{(r_{i-1}, r_i)}(x_1).$$

(We shall write x_j , $j = 1, \dots, p$, to denote the j th component of any p -vector \mathbf{x} .) In order to determine the transition function of $\mathbf{S}(t)$, or equivalently the conditional characteristic function, $\phi(\mathbf{x}, t) = E(\exp(i\theta'\mathbf{S}(t)) | \mathbf{S}(0) = \mathbf{x})$, we need to solve the backward equation,

$$\frac{\partial \phi}{\partial t} = \frac{\sigma^2(x_1)}{2} \frac{\partial^2 \phi}{\partial x_1^2} + \mathbf{a}(\mathbf{x}) \cdot \nabla \phi,$$

where $\sigma^2(x_1) = \sum_{i=0}^{l-1} \sigma^{(i)2} I_{(r_{i-1}, r_i)}(x_1)$, $\phi(\mathbf{x}, 0) = e^{i\theta'\mathbf{x}}$, $|\phi| \leq 1$ and $\partial^2 \phi / \partial x_1^2$ is continuous.

Acknowledgments

We are greatly indebted to Fima Klebaner and Claes Jogleus for valuable discussions in connection with this work. The work of the first author was partially supported by Research Grants A68830996 from the Australian Research Council and DMS-9006422 from the National Science Foundation.

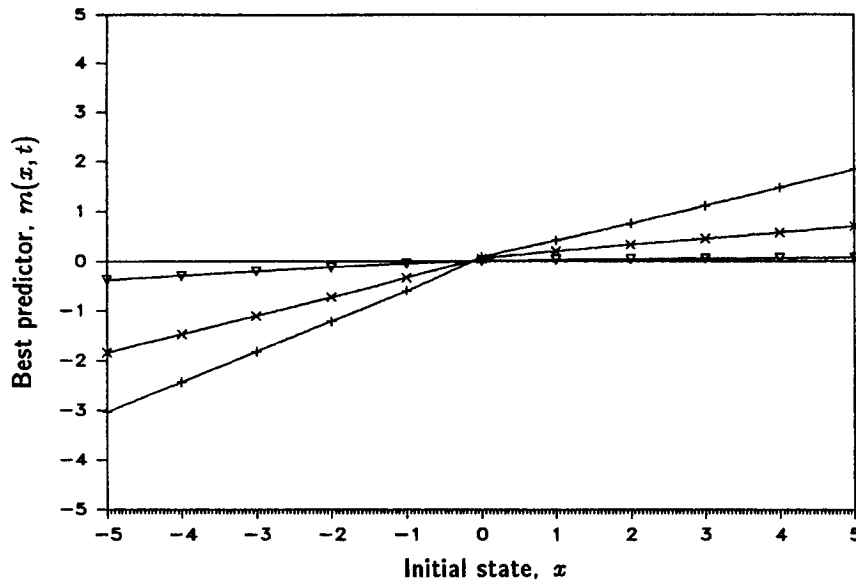


Figure 1. The best predictors $m(x, t)$, $-5 \leq x \leq 5$, for Example 3.1 with lead times $t = 1, 2$ and 5 . Legend: $+ t = 1$, $\times t = 2$, $\nabla t = 5$.

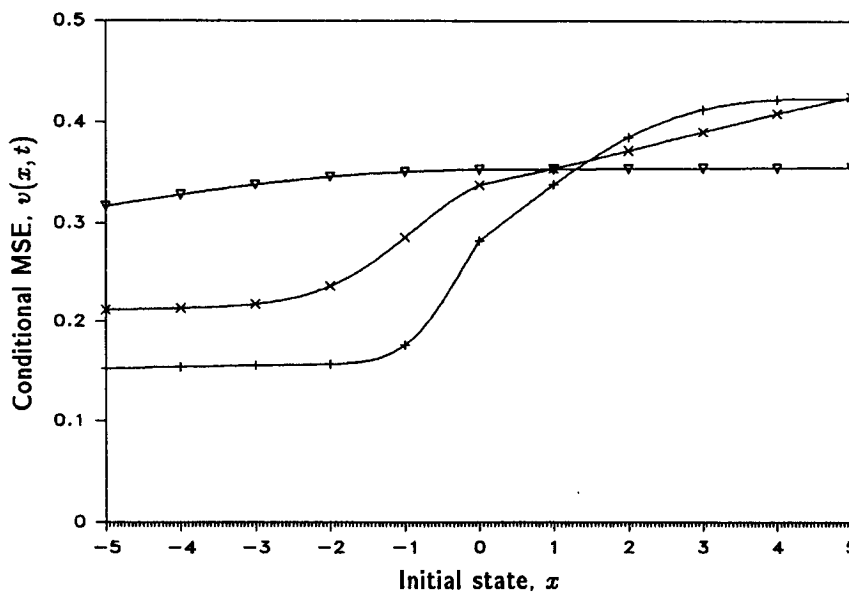


Figure 2. Conditional mean squared errors $v(x, t)$, $-5 \leq x \leq 5$, of the predictors in Figure 1. Legend: $+ t = 1$, $\times t = 2$, $\nabla t = 5$.

References

- Atkinson, J. D. and Caughey, T. K. (1968). Spectral density of piecewise linear first order systems excited by white noise. *J. Non-Linear Mech.* **3**, 137-156.
- Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, Massachusetts.
- Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*. Springer-Verlag, New York.
- Brockwell, P. J., Resnick, S. I. and Pacheco-Santiago, N. (1982). Extreme values, range and weak convergence of integrals of Markov chains. *J. Appl. Probab.* **19**, 272-288.
- Chan, K. S., Petrucci, Joseph D., Tong, H. and Woolford, Samuel W. (1985). A multiple-threshold AR(1) model. *J. Appl. Probab.* **22**, 267-279.
- Jones, Richard H. (1981). Fitting a continuous time autoregression to discrete data. In *Applied Time Series Analysis II* (Edited by D. F. Findley), 651-682. Academic Press, New York.
- Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis*. Lecture Notes in Statistics **21**, Springer-Verlag, New York.
- Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data (with discussion). *J. Roy. Statist. Soc. Ser. B* **42**, 245-292.

Department of Statistics, Colorado State University, Fort Collins, CO 80523, U.S.A.

Department of Statistics, University of Melbourne, Parkville, VIC 3052, Australia.

Department of Statistics, University of Melbourne, Parkville, VIC 3052, Australia.

(Received February 1990; accepted January 1991)