Time-varying autoregressions with model order uncertainty

Raquel Prado* and Gabriel Huerta†

Abstract

We explore some aspects of the analysis of latent component structure in non-stationary time series based on time-varying autoregressive (TVAR) models that incorporate uncertainty on model order. Our modelling approach assumes that the AR coefficients evolve in time according to a random walk and that the model order may also change in time following a discrete random walk. In addition, we use a conjugate prior structure on the autoregressive coefficients and a discrete uniform prior on model order. Simulation from the posterior distribution of the model parameters can be obtained via standard *Forward Filtering Backward Simulation* algorithms. Aspects of implementation and inference on decompositions, latent structure and model order are discussed for a synthetic series and for an electroencephalogram (EEG) trace previously analysed using fixed order TVAR models.

Keywords: Dynamic linear models; Time-varying autoregressions; Model uncertainty; Time series decompositions; Markov chain Monte Carlo

*Departamento de Cómputo Científico y Estadística & Centro de Estadística y Software Matemático (CESMa), Universidad Simón Bolívar, Apartado 89000, Caracas, Venezuela. *E-mail:* raquel@cesma.usb.ve.

†Departamento de Probabilidad y Estadística, Centro de Investigación en Matemáticas, Apartado Postal 402, Guanajuato, Gto., 36000 México. *E-mail:* ghuerta@cimat.mx.

1 Introduction

From a Bayesian perspective, model uncertainty is relevant for parameter estimation and inference and it is summarised through the computation of posterior model probabilities. Additionally, in Bayesian time series, studying how inference on latent structure and prediction of future values is affected when model order uncertainty is considered becomes a key issue. In this paper, we study these aspects for a particular class of time series: the time-varying parameter autoregressive (TVAR) models.

A vast literature of time series models that incorporate model uncertainty via Markov chain Monte Carlo (MCMC) methods has flourished in recent years. For instance, when the class of models is restricted to the linear autoregressive (AR) process, Barnett *et al.* (1996) present a MCMC method, based on a stochastic variable search approach, to deal with model order uncertainty. The priors are specified on the partial autocorrelations with a support that restricts the AR process to be stationary. On the same line, Barbieri and O'Hagan (1997) used priors on the same parameterisation, and developed a MCMC reversible jump algorithm (Green, 1995) to obtain posterior inference on model order and model parameters. Troughton and Godsill (1997) also propose a reversible jump to obtain samples from the posterior distribution of the model order and parameters, but their priors are defined on the standard AR coefficients rather than on partial autocorrelations. More recently, Huerta and West (1999) incorporated model order uncertainty in the linear AR framework with emphasis on prior specification for latent structure. This leads to a novel class of prior distributions on the characteristic reciprocal roots of the AR process. For posterior simulation, they propose a MCMC method that uses a stochastic variable selection approach and reversible jump algorithms. The works cited above illustrate how model uncertainty may be considered in linear and/or stationary time series models using MCMC methods.

For non-stationary Dynamic Linear Models (DLMs), West and Harrison (1997, chapter 12), following Harrison and Stevens (1976), present an approach known as multi-process models where model uncertainty is addressed using mixtures of DLMs. In this scenario, the class of models are conjugate DLMs and the probability of model k, $p(k|D_t)$, where D_t denotes the information up to time t, has an explicit analytic form. When some of the DLMs in consideration are not conjugate but conditionally conjugate, the multi-process requires *Forward Filtering Backward Simulation* algorithms (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994) to obtain posterior model probabilities. If conditional conjugacy is suppressed, posterior model probabilities may be computed via *particle filtering* methods (Pitt and Shephard, 1999). The work by Andrieu *et al.* (1999) is a recent reference in this direction; model order uncertainty and sequential updating are considered for standard autoregressions. Their algorithm is based on *particle filters*, selected with Bayesian importance sampling and MCMC reversible jump steps.

In this paper, we deal with model order uncertainty restricting to the class of time-varying autoregressive models. This class of models has proven remarkably useful in studying the underlying structure of non-stationary signals like electroencephalographic (EEG) traces. References on TVAR modelling include Gersch (1987), Kitagawa and Gersch (1996), Prado and West (1997), West *et al.* (1999) and Krystal *et al.* (2000). These and other authors have demonstrated the flexibility of TVAR models in describing changes in the stochastic structure of non-stationary time series in various applied fields. In particular, West *et al.* (1999)

discusses the specification of TVAR models and decomposition theory of non-stationary time series based on flexible DLM representations. In Section 2, we briefly review the related methodology of decomposition and underlying structure analysis for time-varying autoregressions. In Section 3, we address model order uncertainty within the TVAR modelling framework and discuss issues of posterior inference and extensions of the decomposition results that consider model order uncertainty. Sections 4 and 5 illustrate how the proposed methodology is used to study latent structure in a simulated series and an EEG signal, respectively. Concluding remarks are presented in Section 6.

2 The class of TVAR models and decompositions

We begin by summarising the TVAR model specifications and decomposition results following West *et al.* (1999).

A univariate time series x_t, follows a time-varying autoregression of order p, or TVAR(p), if

$$x_t = \sum_{j=1}^{p} \phi_{t,j} x_{t-j} + \epsilon_t, \tag{1}$$

where $\boldsymbol{\phi}_t = (\phi_{t,1}, \ldots, \phi_{t,p})'$ is the time-varying vector of coefficients and ϵ_t are zero-mean independent innovations assumed Gaussian with possibly time-varying variances σ_t^2. No explicit stationarity constraints are imposed on the AR parameters at each time t. However, if such parameters lie in the stationary region, the series can be thought as locally stationary and the changes in the parameters over time represent global non-stationarities. The model is completed by specifying the evolution structure for $\boldsymbol{\phi}_t$ and σ_t^2. The AR parameters $\boldsymbol{\phi}_t$ evolve according to a random walk, $\boldsymbol{\phi}_t = \boldsymbol{\phi}_{t-1} + \boldsymbol{\xi}_t$, with zero mean innovations $\boldsymbol{\xi}_t$ that are uncorrelated and normal, $\boldsymbol{\xi}_t \sim N(0, \mathbf{W}_t)$. The degree of variation in time of $\boldsymbol{\phi}_t$ is controlled via standard discount factor methods (West and Harrison, 1997). A single discount factor $\beta \in (0, 1]$ leads to values of each $\mathbf{W}_t$ such that low values of β imply high variability of the $\boldsymbol{\phi}_t$ sequence, while high values, in the range 0.9-0.999, are typically relevant in practice. The unit value $\beta = 1$ implies no evolution of the parameters over time. Similarly, the changes in time of σ_t^2 are modelled with a multiplicative random walk $\sigma_t^2 = \sigma_{t-1}^2(\delta/\eta_t)$, where η_t are mutually independent and independent of ϵ_t and $\boldsymbol{\xi}_t$, and with $\eta_t \sim Be(a_t, b_t)$. The parameters a_t and b_t are defined at each t by a discount factor $\delta \in (0, 1]$ analogous to β. Suitable values of the discount factors and p may be assessed via marginal likelihoods as discussed deeply in West *et al.* (1999) and Prado (1998). Sequential updating and retrospective filtering/smoothing algorithms (West and Harrison, 1997) can be applied to obtain posterior model inference.

Once posterior inference is achieved, the focus is on exploring the latent time-frequency structure of the series. The TVAR model in (1) has a DLM form, $x_t = \mathbf{F}'\mathbf{x}_t$, $\mathbf{x}_t = \mathbf{G}_t\mathbf{x}_{t-1} + \boldsymbol{\omega}_t$, where $\mathbf{F} = (1, 0, \ldots, 0)'$, $\mathbf{x}_t = (x_t, x_{t-1}, \ldots, x_{t-p+1})'$, $\boldsymbol{\omega}_t = \epsilon_t\mathbf{F}$ and

$$\mathbf{G}_t \equiv \mathbf{G}(\boldsymbol{\phi}_t) = \begin{pmatrix} \phi_{t,1} & \phi_{t,2} & \ldots & \phi_{t,p-1} & \phi_{t,p} \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}.$$

The eigenvalues of $\mathbf{G}_t$ are the reciprocal roots of the autoregressive characteristic equation at time t. Suppose that, at each time t, $\mathbf{G}_t$ has p distinct eigenvalues, with c pairs of complex eigenvalues denoted by $r_{t,j}\exp(\pm i\omega_{t,j})$ for $j = 1, \ldots, c$, and $r = p - 2c$ real eigenvalues denoted by $r_{t,j}$ for $j = 2c+1, \ldots, p$. Then, the basic decomposition result for the class of TVAR models states that (West *et al.*, 1999),

$$x_t = \sum_{j=1}^{c} z_{t,j} + \sum_{j=2c+1}^{p} y_{t,j}, \tag{2}$$

where the $z_{t,j}$ processes are defined through the complex eigenvalues and the $y_{t,j}$ through the real eigenvalues. In particular, for the standard AR(p) process $\mathbf{G}_t = \mathbf{G}$ and the eigenvalues of $\mathbf{G}$ are the reciprocals roots of the AR characteristic equation. In this case $r_{t,j} = r_j$ for $j = 1, \ldots, p$ and $\omega_{t,j} = \omega_j$ for $j = 1, \ldots, c$. Furthermore, each $y_{t,j}$ follows a standard AR(1) process with AR parameter r_j, and each $z_{t,j}$ follows an ARMA(2,1) whose AR(2) component is quasi-periodic with constant characteristic frequency ω_j (or wavelength $2\pi/\omega_j$) and modulus r_j. In the general TVAR case, and under certain conditions discussed in Prado (1998), each $y_{t,j}$ is dominated by a TVAR(1) with time-varying AR parameter $r_{t,j}$ and each $z_{t,j}$ is dominated by a TVARMA(2,1) with time-varying characteristic frequency $\omega_{t,j}$ and modulus $r_{t,j}$.

In order to illustrate the above results, we present a TVAR decomposition of an EEG series recorded on a patient who received a moderate level of electroconvulsive therapy (ECT) stimulus intensity. Details about the EEG data, the conditions and implications of the therapy, as well as a complete statistical analysis based on TVAR models appears in West *et al.* (1999) and Krystal, Prado and West (1999). Figure 1 displays the data and estimated latent components in the series, based on a TVAR(12) model with constant observational variance $\sigma_t^2 = \sigma^2$, and discount factor $\beta = 0.997$ controlling the variability of $\boldsymbol{\phi}_t$. Components (1), (2), and (3) in the decomposition are the components with highest amplitudes, lying in the delta (0 to 4 Hz) and theta (4 to 8 Hz) frequency bands. These components are individual processes dominated by TVARMA(2,1) quasi-periodic structures. For instance, process (1) is dominated by a TVARMA(2,1) with time-varying characteristic frequency lying in a delta band that starts around 4 Hz and gradually decays in time. The time-varying characteristic modulus of process (1) is consistently high, with values higher than 0.95 over the seizure course. Components (4) to (6) are low amplitude components representing neural and experimental noise. The end of the seizure occurs around $t = 1800$. Clearly, the contribution of components (2) to (6) is practically negligible after $t = 1800$. The structure of the EEG signal is more complex while the seizure starts and matures than when it begins to decay and eventually dies off. Thus, it seems reasonable to consider a TVAR model with a higher model order during the beginning and middle parts of the seizure than towards the end.

3 Autoregressions with time-variation on the AR coefficients and model order

A time-varying autoregression with time-varying order p_t, or TVAR(p_t), is described by

$$x_t = \sum_{j=1}^{p_t} \phi_{t,j} x_{t-j} + \epsilon_t, \tag{3}$$

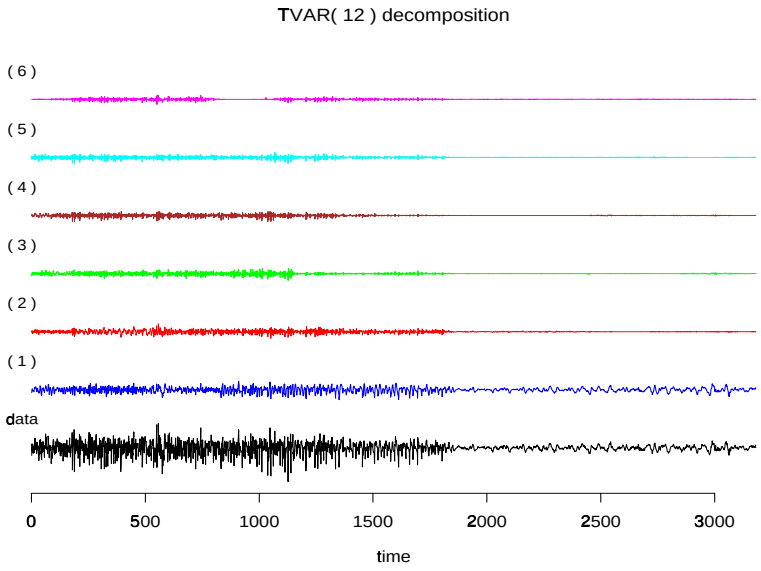

Figure 1: Data and estimated components in the decomposition of an EEG series based on a TVAR(12). From the bottom up, the graph displays the series followed by the estimated components in order of decreasing amplitude.

where the autoregressive coefficients change in time according to a random walk, as defined in Section 2 for a TVAR(p), and ϵ_t are zero-mean innovations, assumed Gaussian with constant variance σ^2. Additionally, we assume that p_t, the order of the autoregression at time t, is an integer that takes values between a fixed lower bound $p_{\min}$ and a fixed upper bound $p_{\max}$. The TVAR(p_t) model in (3) is a sub-model of the TVAR($p_{\max}$) described by

$$x_t = \sum_{j=1}^{p_{\max}} \phi_{t,j} x_{t-j} + \epsilon_t, \tag{4}$$

with a $p_{\max}$-dimensional vector of coefficients $\boldsymbol{\phi}_t = (\phi_{t,1}, \ldots, \phi_{t,p_t}, 0, \ldots, 0)'$. Model completion requires specification of an initial prior for $(\boldsymbol{\phi}_1, \sigma^2)$ and details concerned with the evolution of model parameters. For simplicity, we set $\boldsymbol{\phi}_1 = (\phi_{1,1}, \ldots, \phi_{1,p_{\max}})'$, with $\phi_{1,j} \neq 0$ for all j, so that $p_1 = p_{\max}$. In addition, we take relatively diffuse normal priors on $\boldsymbol{\phi}_1$, that is, $N(\boldsymbol{\phi}_1 | 0, \sigma^2 I_{p_{\max}})$ and vague inverse-gamma priors on σ^2. The evolution of p_t is defined through a first order discrete random walk with known transition probabilities denoted by $P[p_t = i | p_{t-1} = j]$ and with possible values for i and j ranging from $p_{\min}$ to $p_{\max}$.

For posterior inference of the TVAR(p_t), we propose a MCMC method that follows a two-stage Gibbs sampling format. Conditional on $\{p_1, \ldots, p_n\}$, the standard sequential updating and retrospective filtering/smoothing algorithms for DLMs apply. Based on all the observed information $D_n = \{D_0, x_1, x_2, \ldots, x_n\}$, where D_0 denotes the initial information, and the model orders, the sequences $\boldsymbol{\phi}_t$ and σ^2 are sampled from normal and inverse gamma distributions respectively. The second stage consists on sampling from the conditional posterior distribution of p_t, given $\boldsymbol{\phi}_t$ and σ^2, via the filtering/smoothing algorithm for discrete

random variables of Carter and Kohn (1994). A short description of the algorithm is made here with full details presented in the Appendix. To simplify the notation, let D_t include all the information up to time t, the conditioning sequences $\boldsymbol{\phi}_t$ and σ^2. We begin computing the conditional prior probabilities $P[p_t = i|D_{t-1}]$ for each t and $i = p_{\min}, \ldots, p_{\max}$, marginalising the joint probability distribution $P[p_t = i, p_{t-1} = j|D_{t-1}]$ over j. We compute the posterior probabilities $P[p_t = i|D_t]$ via Bayes theorem and save the prior and posterior probabilities for p_t, for all t. Then, we sample p_n^* from $P[p_n = i|D_n]$. Finally, for each $t = n-1, \ldots, 1$, we sample p_t^* from the distribution $P[p_t = i|p_{t+1}^*, D_t]$, where p_{t+1}^* is the sampled value for model order at time $t+1$. The vector $(p_1^*, \ldots, p_{n-1}^*, p_n^*)$ constitutes a sample from the joint conditional posterior distribution of model orders.

3.1 Decompositions for TVAR models with time-varying order

The TVAR(p_t) model defined above can be written in DLM form, $x_t = \mathbf{F}'\mathbf{x}_t$, $\mathbf{x}_t = \mathbf{G}_t\mathbf{x}_{t-1} + \boldsymbol{\omega}_t$, where $\mathbf{F}$ is a $p_{\max} \times 1$ vector $\mathbf{F} = (1, 0, \ldots, 0)'$, $\mathbf{x}_t = (x_t, x_{t-1}, \ldots, x_{t-p_{\max}+1})'$, $\boldsymbol{\omega}_t = \epsilon_t \mathbf{F}$ and $\mathbf{G}_t$ a $p_{\max} \times p_{\max}$ matrix,

$$\mathbf{G}_t \equiv \mathbf{G}(\boldsymbol{\phi}_t) = \begin{pmatrix} \phi_{t,1} & \phi_{t,2} & \cdots & \phi_{t,p_t-1} & \phi_{t,p_t} & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \ddots & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & & 1 & 0 \end{pmatrix}. \tag{5}$$

Assume that $\mathbf{G}_t$ has p_t distinct non-zero eigenvalues denoted by $\alpha_{1,t}, \ldots, \alpha_{p_t,t}$ and a zero eigenvalue, denoted by $\alpha = 0$, with multiplicity $p_{\max} - p_t$. The non-zero eigenvalues correspond to the reciprocal roots of the characteristic polynomial at time t, $\Phi_{p_t}(u) = (1 - \phi_{1,t}u - \ldots - \phi_{t,p_t}u^{p_t})$. Then, $\mathbf{G}_t = \mathbf{E}_t\mathbf{A}_t\mathbf{E}_t^{-1}$ with

$$\mathbf{A}_t = \text{block diag}[\mathbf{A}_{p_t}, \mathbf{J}_{(p_{\max}-p_t)}(0)], \quad \mathbf{E}_t = [\mathbf{e}_{1,t}, \ldots, \mathbf{e}_{p_t,t}, \mathbf{h}_{1,t}, \ldots, \mathbf{h}_{(p_{\max}-p_t),t}],$$

where $\mathbf{A}_{p_t} = \text{diag}(\alpha_{t,1}, \ldots, \alpha_{t,p_t})$ and $\mathbf{J}_{(p_{\max}-p_t)}(0)$ is the $(p_{\max} - p_t) \times (p_{\max} - p_t)$ Jordan block associated to the eigenvalue $\alpha = 0$,

$$\mathbf{J}_{(p_{\max}-p_t)}(0) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

$\mathbf{E}_t$ is a $p_{\max} \times p_{\max}$ matrix whose first p_t columns, $\mathbf{e}_{\cdot,t}$, correspond to the eigenvectors associated to the eigenvalues $\alpha_{1,t}, \ldots, \alpha_{p_t,t}$. The last $p_{\max} - p_t$ columns of $\mathbf{E}_t$ are the vectors $\mathbf{h}_{1,t}, \ldots, \mathbf{h}_{(p_{\max}-p_t),t}$, where each $\mathbf{h}_{j,t}$ is such that its first $p_{\max} - j$ components are zeroes and

its last j components are ones. Some of the non-zero eigenvalues of $\mathbf{G}_t$ could be complex and in such case they appear in conjugate pairs. Assume that at each time t, there are c_t pairs of complex eigenvalues denoted by $r_{t,j}\exp(\pm i\omega_{t,j})$, $j = 1, \ldots, c_t$ and $r_t = p_t - 2c_t$ non-zero real and distinct eigenvalues denoted by $r_{t,j}$, $j = 2c_t + 1, \ldots, p_t$. For each time t, define the matrix $\mathbf{H}_t = \mathbf{D}_t\mathbf{E}_t^{-1}$, with $\mathbf{D}_t = \text{diag}(\mathbf{E}_t'\mathbf{F})\mathbf{E}_t^{-1}$, and $\boldsymbol{\gamma}_t = \mathbf{H}_t\mathbf{x}_t$. Then, we may write $x_t = (1, \ldots, 1)'\boldsymbol{\gamma}_t$ or equivalently,

$$x_t = \sum_{j=1}^{c_t} z_{t,j} + \sum_{j=2c_t+1}^{p_t} y_{t,j}, \tag{6}$$

with $z_{t,j} = \gamma_{t,2j-1} + \gamma_{t,2j}$, for $j = 1, \ldots, c_t$, and $y_{t,j} = \gamma_{t,j}$, for $j = 2c_t + 1, \ldots, p_t$. Notice that the decomposition result is analogous to (2), but now the number of components depends on c_t and r_t which are time-varying.

To gain insight on the structure and interpretability of the latent processes $z_{t,j}$ and $y_{t,j}$, consider the fixed order TVAR(i) models $\mathcal{M}_i$ for each possible order $i = p_{\min}, \ldots, p_{\max}$,

$$\mathcal{M}_i: \quad x_t = \mathbf{F}'\mathbf{x}_t, \quad \mathbf{x}_t = \mathbf{G}_t^{(i)}\mathbf{x}_{t-1} + \boldsymbol{\omega}_t$$

with

$$\mathbf{G}_t^{(i)} = \begin{pmatrix} \phi_{t,1} & \phi_{t,2} & \ldots & \phi_{t,i-1} & \phi_{t,i} & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots & & \ddots & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & & 1 & 0 \end{pmatrix}.$$

For each $\mathcal{M}_i$, reparameterise $\mathbf{x}_t$ and $\boldsymbol{\omega}_t$ via $\boldsymbol{\gamma}_t^{(i)} = \mathbf{H}_t^{(i)}\mathbf{x}_t$ and $\boldsymbol{\delta}_t^{(i)} = \mathbf{H}_t^{(i)}\boldsymbol{\omega}_t$ with $\mathbf{H}_t^{(i)} = \mathbf{D}_t^{(i)}(\mathbf{E}_t^{(i)})^{-1}$. Then we have,

$$x_t = \mathbf{1}'\boldsymbol{\gamma}_t^{(i)} \quad \text{and} \quad \boldsymbol{\gamma}_t^{(i)} = \mathbf{A}_t^{(i)}\mathbf{K}_t^{(i)}\boldsymbol{\gamma}_{t-1}^{(i)} + \boldsymbol{\delta}_t^{(i)},$$

with $\mathbf{K}_t^{(i)} = \mathbf{D}_t^{(i)}(\mathbf{E}_t^{(i)})^{-1}\mathbf{E}_{t-1}^{(i)}\mathbf{D}_{t-1}^{(i)*} = \mathbf{H}_t^{(i)}\mathbf{H}_{t-1}^{(i)*}$. $\mathbf{D}_t^{(i)*}$ and $\mathbf{H}_{t-1}^{(i)*}$ are generalised inverse matrices of $\mathbf{D}_t^{(i)}$ and $\mathbf{H}_{t-1}^{(i)}$ respectively. Given known values of $\mathbf{G}_t^{(i)}$ and $\mathbf{x}_t$ for each t, we can obtain a decomposition for x_t based on $\mathcal{M}_i$,

$$x_t = \sum_{j=1}^{c^{(i)}} z_{t,j}^{(i)} + \sum_{j=2c^{(i)}+1}^{i} y_{t,j}^{(i)},$$

for each $i = p_{\min}, \ldots, p_{\max}$ where $i = r^{(i)} + 2c^{(i)}$. The value $c^{(i)}$ is the number of conjugate pairs of complex non-zero eigenvalues of $\mathbf{G}_t^{(i)}$, denoted by $r_{t,j}^{(i)}\exp(\pm i\omega_{t,j}^{(i)})$, $j = 1, \ldots, c^{(i)}$, and $r^{(i)}$ is the number of real eigenvalues. As in the TVAR(p) case, we are assuming that $c^{(i)}$

and $r^{(i)}$ are fixed in time (see West *et al.*, 1999). If the AR coefficients change slowly in time, as is often the case in practice, $\mathbf{K}_t^{(i)} \approx \text{blockdiag}[\mathbf{I}_{i\times i}, \mathbf{0}_{(p_{\max}-i)\times(p_{\max}-i)}]$. Then, each $z_{t,j}^{(i)}$ is dominated by a TVARMA(2,1) process with time-varying modulus $r_{t,j}^{(i)}$ and frequency $\omega_{t,j}^{(i)}$, and each $y_{t,j}^{(i)}$ is dominated by a TVAR(1) with time-varying modulus $r_{t,j}^{(i)}$. The general decomposition result (6) is such that at a given time t, $z_{t,j} = z_{t,j}^{(p_t)}$, $y_{t,j} = y_{t,j}^{(p_t)}$, $r_{t,j} = r_{t,j}^{(p_t)}$ and $\omega_{t,j} = \omega_{t,j}^{(p_t)}$. Thus, assuming that the AR coefficients change smoothly in time, each $z_{t,j}$ in (6) coincides at time t, with a process dominated by a TVARMA(2,1) with instantaneous characteristic modulus and frequency $r_{t,j}^{(p_t)}$ and $\omega_{t,j}^{(p_t)}$. Similarly, each $y_{t,j}$ coincides at time t, with a process dominated by a TVAR(1) with time-varying modulus $r_{t,j}^{(p_t)}$.

4 Study of synthetic data

In this section we compare estimations of the latent structure of a synthetic series obtained using a TVAR(p_t) and a fixed order TVAR(p). Figure 2 (a) displays the synthetic data at the bottom. The first 550 data points were generated from a TVAR(12) with five quasi-periodic latent components and two real components, while the last 550 observations were generated from a purely quasi-periodic TVAR(4). The latent processes shown in the figure appear in order of decreasing wavelength therefore, component (1) corresponds to the highest wavelength component and component (5) to the lowest wavelength component for $t = 1, \ldots, 550$. Component (6) is the result of adding the latent processes associated to the two real roots for $t = 1, \ldots, 550$. Similarly, component (1) is the highest wavelength component and component (4) is the lowest wavelength component for $t = 551, \ldots, 1100$. Note that processes (2), (3), (5) and (6) have no contribution to the decomposition of the series for $t = 551, \ldots, 1100$ since the series was generated from a quasi-periodic TVAR(4) during this time period. The trajectories in time of the characteristic wavelengths, moduli and amplitudes associated to the actual time-varying AR parameters, also ordered by wavelength, are displayed in Figures 2 (b), (c) and (d). Note that the ordering of the latent processes in the decomposition is a key issue, since there is no inherent mathematical identification of the characteristic roots.

We fitted a TVAR(p_t) model to the series. Discount factors in the range of $0.99 - 0.998$ were considered to control the evolution of the AR coefficients in time. Such values impose smoothness restrictions on the changes of the model parameters in time that are typical in practice (West and Harrison, 1997; West *et al.*, 1999) and guarantee the interpretation of the latent component structure described in Section 3. The transition probabilities that model the evolution of p_t in time are specified as follows

$$P[p_t = i | p_{t-1} = j] = \begin{cases} q_{ii} & j = i \\ q_{i-1,i} & j = i-1, \quad i \leq p_{\max} \\ q_{i-2,i} & j = i-2, \quad i \leq p_{\max} \\ q_{i+1,i} & j = i+1, \quad i \geq p_{\min} \\ q_{i+2,i} & j = i+2, \quad i \geq p_{\min} \\ 0 & \text{otherwise,} \end{cases}$$

with q_{ii} usually in the range $0.9 - 0.9999$. Note that this structure allows to increase or decrease the current model order by only one or two units at each time t. Therefore, we

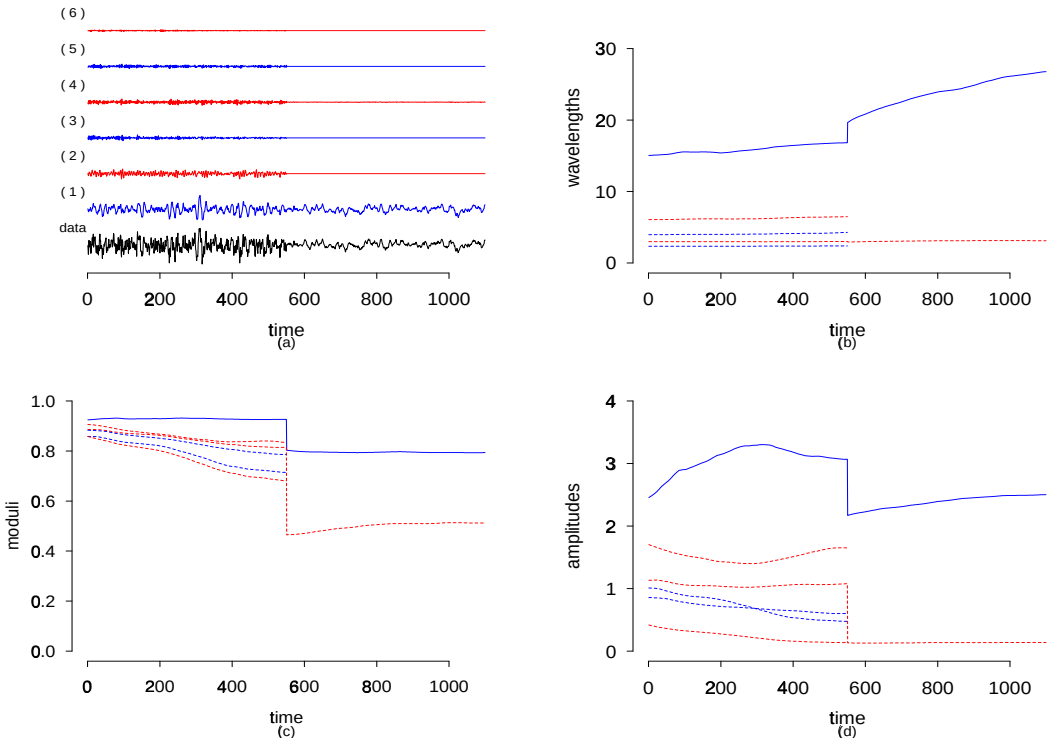

Figure 2: (a) Simulated series and latent processes ordered by wavelength. (b) Trajectories of the estimated wavelengths. (c) Trajectories of the estimated moduli ordered by wavelengths. The ordering of the components in terms of decreasing moduli is (1), (2), (4), (3), (5) and (6) for $t = 1, \ldots, 550$ and (1), (4) for $t = 551, \ldots, 1100$. (d) Estimated amplitude trajectories ordered by wavelengths. The ordering of the components in terms of decreasing amplitude is (1), (2), (4), (3)/(5) (switching) and (6) for $t = 1, \ldots, 550$ and (1), (4) for $t = 551, \ldots, 1000$.

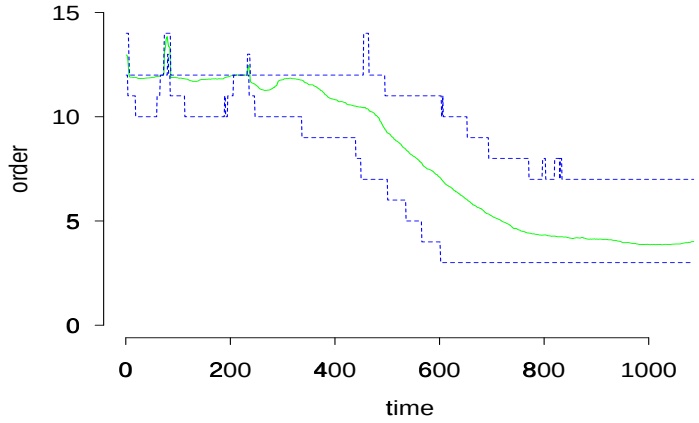

Figure 3: Simulated series: estimated posterior mean for model order at each time t (solid line) with 95% posterior probability bands (dotted lines).

are also imposing smoothness conditions on the changes of p_t, allowing to include or delete only one characteristic root, complex or real, at each time t. A uniform prior $P(p_1 = i) = 1/(p_{\max} - p_{\min} + 1)$, $i = p_{\min}, \ldots, p_{\max}$ is used on the model order, while relatively diffuse normal/inverse-gamma priors are considered for the AR coefficients and σ^2 respectively.

Figure 3 shows the estimated posterior mean (solid line) and 95% posterior bands (dotted lines) for the model order at each time t. Posterior summaries were obtained using a model with $p_{\min} = 2$, $p_{\max} = 15$, $\beta = 0.995$ and a transition probability matrix defined as follows. We took $q_{ii} = 0.99$ for all i; $q_{2,3} = 0.002$; $q_{2,4} = 0.008$; $q_{3,2} = q_{3,4} = 0.004$; $q_{3,5} = 0.002$; $q_{15,14} = q_{15,13} = 0.005$; $q_{14,15} = q_{14,13} = 0.001$; $q_{14,12} = 0.008$; $q_{i,i-1} = q_{i,i+1} = 0.001$ and $q_{i,i-2} = q_{i,i+2} = 0.004$ for $i = 4, 6, 8, 10, 12$ and $q_{i,i-1} = q_{i,i+1} = q_{i,i-2} = q_{i,i+2} = 0.0025$ for $i = 5, 7, 9, 11, 13$. The results presented here are based on a posterior sample of 3000 draws taken from 10000 iterations of the Gibbs sampler after a burn-in of 4000 iterations for MCMC convergence. The graph shows that the model order oscillates between 10 and 14 for, roughly, the first 400 observations, with a posterior mean of $\tilde{p}_t = 12$ for most of this initial period. After $t = 400$ the model order decreases and the uncertainty on the model order increases. At about $t = 750$ and up to $t = 1100$, $\tilde{p}_t$ oscillates around 4.

Figures 4 (a) and (b) show the latent processes and the trajectories of the characteristic wavelengths estimated using a *fixed order* TVAR(12) model. The estimated latent components were computed at the estimated posterior means of the AR coefficients and innovations variance. A discount factor value of $\beta = 0.995$, chosen by marginal likelihood maximisation, was used to fit the model. Figure 4 (a) displays, from the bottom up, the data and estimated latent processes in the decomposition ordered by decreasing wavelengths. Figure 4 (b) graphs the time trajectories of the estimated characteristic wavelengths. The solid line corresponds to the highest wavelength component. Note that from about $t = 550$ and up to $t = 1100$, the estimated highest wavelength component takes lower values than the actual wavelength component (see Figure 2 (b)). The estimated wavelength trajectories associated to the latent processes (2), (3) and (5) are constant over time after $t = 550$, while such components do not arise in the original data. In addition, from approximately $t = 700$ and up to $t = 1100$,

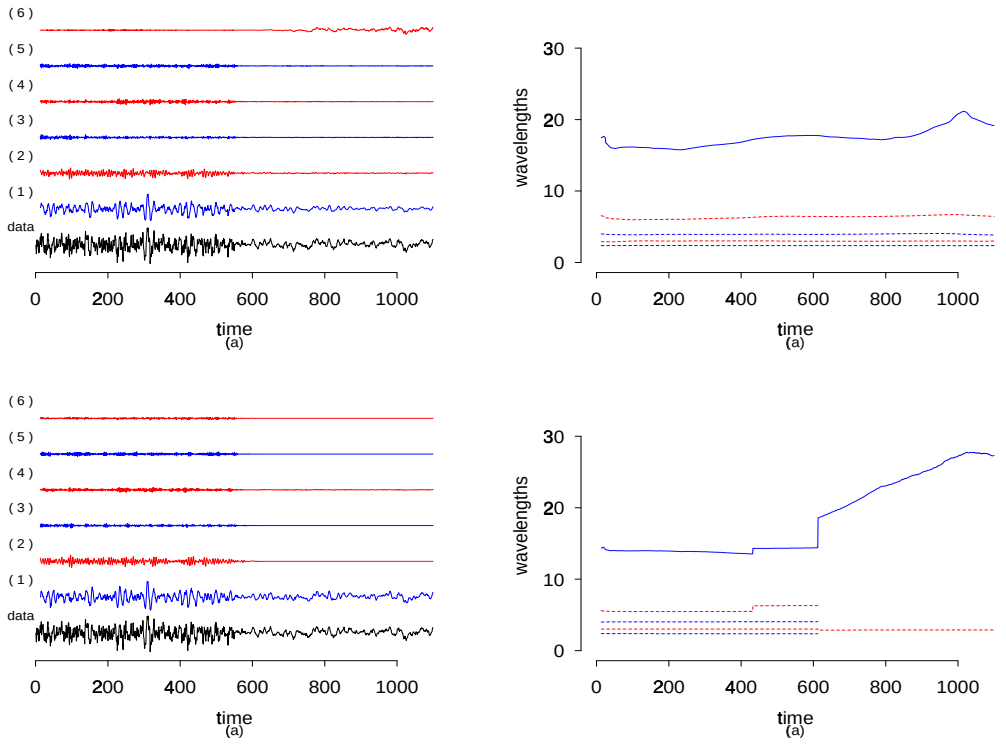

Figure 4: (a) Fixed order TVAR: simulated process and estimated components ordered by decreasing wavelength from the bottom up. (b) Fixed order TVAR: estimated wavelength trajectories. (c) TVAR(p_t): simulated process and estimated components ordered by decreasing wavelength from the bottom up. (d) TVAR(p_t): trajectories of the estimated wavelengths.

the estimated component (6), which is the sum of two components associated to a couple of real roots, shows a pattern that that should have been captured by component (1). Therefore, the TVAR model is reproducing spurious features during some time periods and is not adequately estimating the latent structure of the series.

Figures 4 (c) and (d) show the estimated latent processes in the decomposition and the trajectories of the characteristic wavelengths for the TVAR(p_t) fitted with the discount factor and transition probabilities values specified above. Again, the estimated latent processes are ordered by decreasing wavelength from the bottom up. The estimated latent components and wavelengths were computed at the posterior mode for model order and the estimated posterior means for the AR coefficients and innovations variance for each time t. The trajectory of the highest estimated wavelength resembles the actual wavelength trajectory for $t > 600$. The wavelength trajectories for components (2), (3) and (5) are roughly constant up to $t = 600$ and zero from this point until the end of the series, in agreement with the behaviour of the actual components. Then, the TVAR(p_t) model is correctly estimating the latent component structure of the synthetic series. This example illustrates the relevance of considering model order uncertainty when estimating the latent structure with TVAR models.

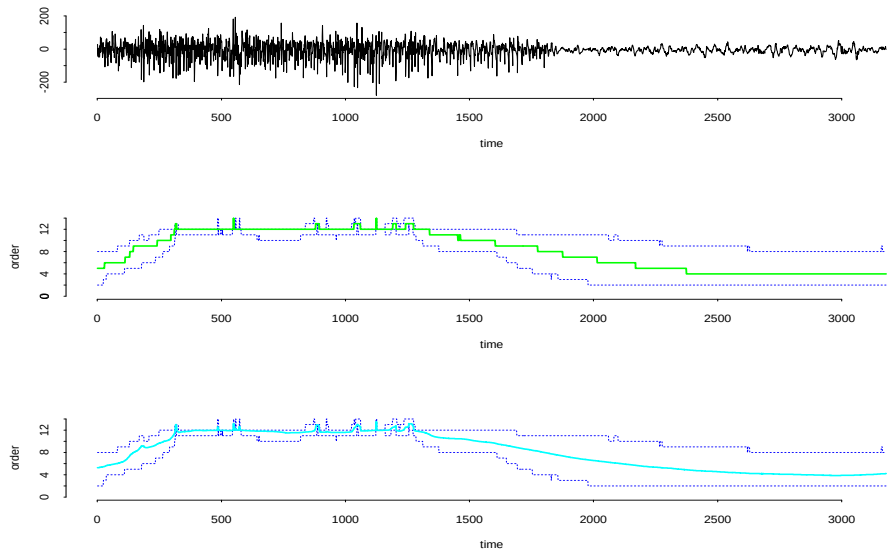

Figure 5: From the top down we have the EEG data, the estimated posterior median for model order and each time t with 95% posterior bands and the estimated posterior mean for model order at each time t with 95% posterior bands.

5 An application: analysis of an EEG trace via TVAR models with uncertainty on the model order

Consider again the EEG series displayed at the bottom of Figure 1. The latent components shown in the graph were computed using estimated posterior means for the AR coefficients and the innovations variance of a TVAR(12) model. In this section we model the same series with a TVAR(p_t) whose model order at time t, p_t, may take values from $p_{\min} = 0$ up to $p_{\max} = 14$. Different values of $p_{\min}$ and $p_{\max}$ were also considered, leading to similar inferences in terms of the latent structure. The transition probability structure used in this example to model the evolution of p_t, is similar to the structure described in the previous section. In particular, values of q_{ii} in the range of 0.9 - 0.9999 were taken for $i = 0, \ldots, 14$, so abrupt changes in the model order from $t-1$ to t are not permitted. Similarly, we use values of the discount factor for the AR coefficients in the range of 0.99 - 0.999. These values were chosen based on exploration of marginal likelihood functions that result from analysing the series via TVAR models with fixed model order. A discrete uniform prior on model order, $P(p_1 = i) = 1/15$, and relatively diffuse conjugate normal/inverse-gamma priors were used for the AR coefficients and the innovations variance.

Figure 5 displays from the top down, the EEG data, the trajectory in time of the estimated posterior median for model order with 95% posterior probability bands (center panel) and the trajectory of the estimated posterior mean for model order with 95% posterior probability bands. The graphs were obtained with a discount factor $\beta = 0.997$ for the AR coefficients, and a transition probability matrix defined by, $q_{ii} = 0.99$ for all i; $q_{i,i+1} = q_{i,i-1} = 0.004$ and $q_{i,i+2} = q_{i,i-2} = 0.001$ for $2 \leq i \leq 12$; $q_{0,1} = q_{0,2} = q_{14,13} = q_{14,12} = 0.005$; $q_{1,0} = q_{1,2} =$

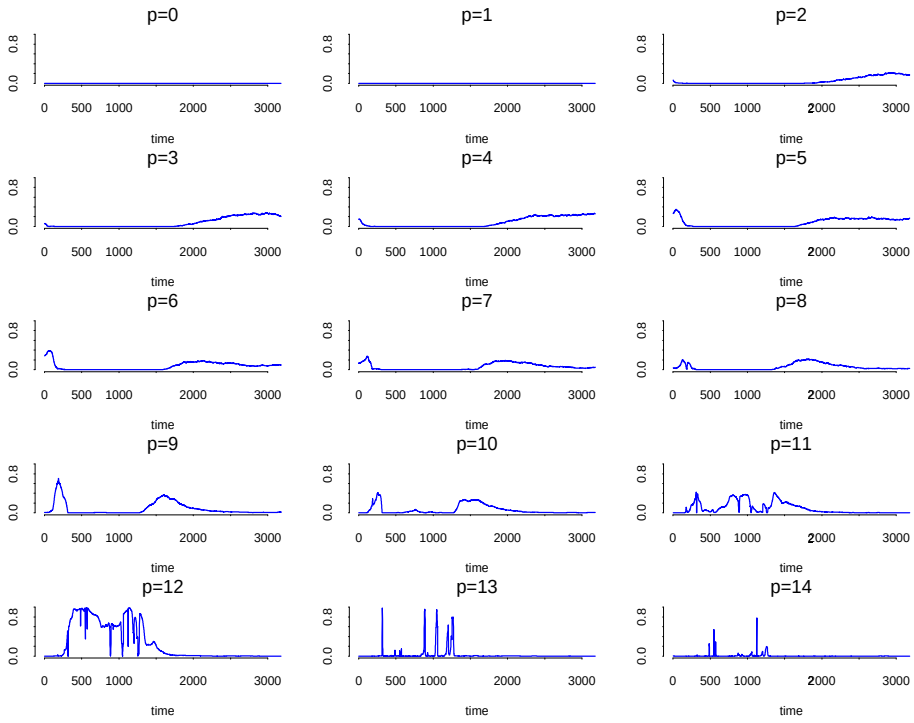

Figure 6: Estimated posterior probabilities for model order at each time t.

$q_{13,14} = q_{13,12} = 0.004$ and $q_{1,3} = q_{13,11} = 0.002$. The instantaneous posterior means, medians and 95% posterior probability bands for model order displayed in Figure 5, are based on 5000 samples taken from 15000 iterations of the Gibbs sampler after a burn-in of 3000 iterations for MCMC convergence. The graphs show that the model order is higher in the middle sections of the seizure - roughly from $t = 350$ until $t = 1300$ - than at the beginning and towards the end of the seizure. The posterior median for p_t increases from $\hat{p}_t = 4$ up to $\hat{p}_t = 12$, at the beginning of the seizure, decreasing from $\hat{p}_t = 12$ to $\hat{p}_t = 4$ between $t = 1300$ and $t = 2400$. The patterns observed in the trajectories of the estimated model order posterior mean and median over time, indicate that the complexity of the latent structure is higher at middle parts of the seizure than at the beginning and once the seizure is over. Furthermore, the complexity of the data structure measured as a function of model order starts to decrease prior to the seizure dissipation, just before $t = 1500$. Alternative transition probability values, always imposing smoothness restrictions on the changes of model order in time, were taken into account, leading to similar results in terms of inference on time-varying model order and latent structure of the EEG series.

Figure 6 displays estimated posterior probabilities for model order at each time t. Model orders 11 and 12 are favoured for $300 < t < 1300$ while lower order models are preferred for starting and late periods of the series. Figure 7 shows the decomposition of the EEG series based on estimated posterior means of the AR coefficients, the estimated posterior mean of the innovations variance and posterior medians for model order at each time t. From the bottom up, the graph displays the data followed by the estimated latent processes. Component (1), which has the highest amplitude, is related with the highest characteristic wavelength for

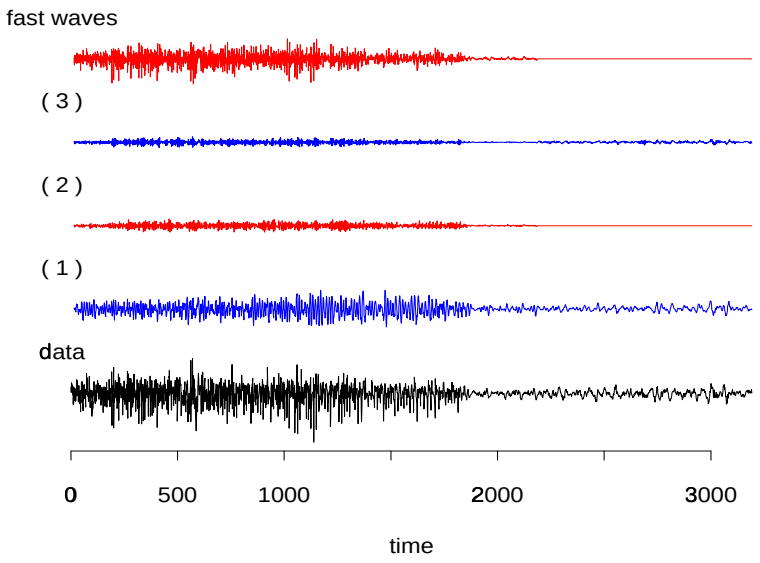

Figure 7: Data and estimated components in the decomposition of an EEG series based on a TVAR(p_t).

almost all t from $t = 1$ to roughly $t = 2100$. After $t = 2100$ the latent process (1) is the sum of two components, a component associated to the highest characteristic wavelength after $t = 2100$ and a component associated to a real root with relatively high time-varying modulus. Component (2) in the decomposition is a quasi-periodic process related to the second highest wavelength from $t = 1$ to $t = 2100$ that dissipates for $t > 2100$, while component (3) is related to a real root. Component (4), denoted in the Figure as "fast waves", is the sum of all remaining low amplitude processes that correspond to high frequency and noise components. Figure 8 (a) sketches the trajectories of the wavelength components underlying the data taking into account uncertainty on the number of such components over time. These trajectories display jumps and artifacts that are simply the result of using the first $\hat{p}_t$ elements of the estimated $\boldsymbol{\phi}_t$ vector at each time t, where $\hat{p}_t$ is the estimated posterior median for model order. Such artifacts make interpretation of the latent processes extremely difficult in practice. A graph of the approximate wavelengths trajectories can be obtained by considering the maximum order model for all time t. Figure 8 (b) displays the wavelengths trajectories of the first four quasi-periodic components computed at the estimated posterior mean of $\boldsymbol{\phi}_t$ with $p_t = p_{\max} = 14$ for all t. The highest wavelength also corresponds to the highest amplitude component in the decomposition for $t = 1$ to roughly $t = 2100$, characterising the frequency structure of process (1) in the decomposition of the series. This is indicated in the Figure by the dark segment on the trajectory of the highest wavelength component. In terms of frequency in the original sampling scale (see discussion in West *et al.*, 1999 about subsampling of the original data series), we have a dominant frequency of about 4 to 5 Hz at the beginning of the seizure that gradually decays to approximately 3 Hz at $t = 2100$. These findings are consistent with the results obtained for a fixed order model presented in West *et*

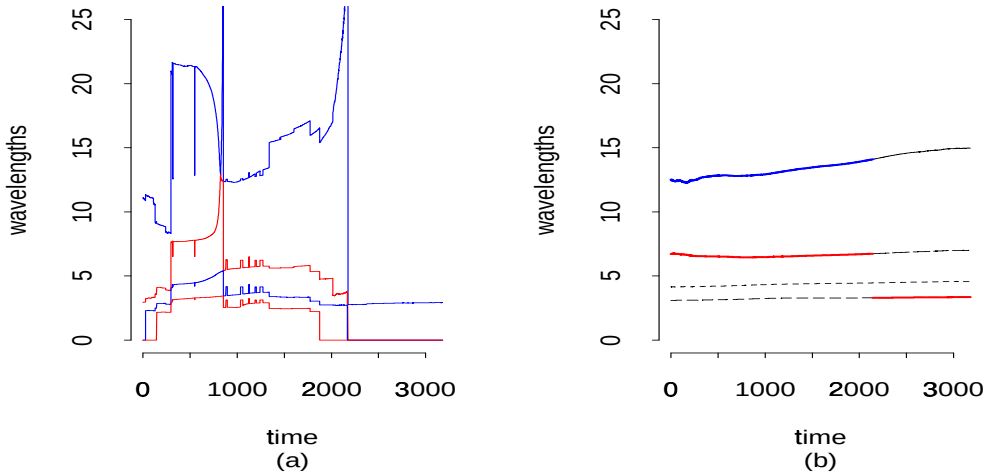

Figure 8: (a) Traces of wavelengths corresponding to the first four quasi-cyclical components computed at instantaneous model order medians. (b) Traces of wavelengths corresponding to the first four quasi-cyclical components ordered by wavelength computed at the maximum order for all t.

al. (1999). The second highest wavelengths in the Figure corresponds to the second highest amplitude component for $1 < t < 2100$, while the lowest wavelength component in the Figure corresponds to the highest amplitude component for $t > 2100$.

Inferences obtained by including model order uncertainty as a modelling component are consistent with the results obtained in previous analyses of the same series (West *et al.*, 1999; Krystal *et al.*, 2000). Additional insight in terms of complexity of the latent structure is gained using a TVAR(p_t) model approach. For instance, the decrease in model order observed just prior to the end of the seizure might be relevant in connection with assessing the clinical efficacy of the treatment.

6 Remarks, conclusions and future directions

In this paper we present a TVAR model that fully incorporates model order uncertainty using a first order discrete random walk to describe the evolution of model order in time. Our modelling approach allows decomposing the data in terms of latent processes, perhaps of quasi-cyclical nature, generalising the decomposition results for fixed order TVAR models presented in West *et al.* (1999). Based on MCMC methods, we explored the performance of the model for a synthetic series and an EEG series. In both cases, we studied the impact of discount factor selection and specification of transition probabilities in terms of inference on model order and latent structure. Since model fitting relies on simulation, the TVAR(p_t) requires a higher computational demand than a fixed order TVAR. Additionally, and as seen in the EEG series analysis of Section 5, the TVAR(p_t) may need more thoughtful interpretations for the estimated latent components and trajectories of wavelengths, moduli and amplitudes

corresponding to the different characteristic roots of the process. These issues can be resolved by dealing with model order uncertainty directly on the number of instantaneous characteristic reciprocal roots, following a structure similar to the one developed in Huerta and West (1999) and Huerta (1998) for standard autoregressions. Alternatively, TVAR model order uncertainty can be handled in the partial correlation (PARCOR) domain (see Barnett *et al.*, 1996; Kitagawa and Gersch, 1996). However, implementation of such models needs to incorporate *particle filter* methods to obtain updating and smoothing posterior distributions of the model parameters (Pitt and Shephard, 1999; Godsill *et al.*, 2000). Such extensions, surely leading to very interesting results and challenging methodological issues, will be considered in the future.

In connection to mixture models, the TVAR(p_t) can be seen as a multi-process class II mixture model (West and Harrison, 1997, pp. 444-445), where the mixing components are determined by the possible values of p_t. The limitation of this approach is that it introduces a mixture of DLMs where each component has state vectors of different dimension. In consequence, expressions for the posterior distributions of the state vectors and model order are not available in closed form and the multi-process requires MCMC methods as presented in this paper. In contrast, a mixture of TVARs, each with a fixed order and assuming that one of the models holds for all time t, could be considered. This structure introduces a multi-process class I model with posterior distributions available in closed form. Although, this multi-process model can be handled more easily, it lacks flexibility compared to the TVAR(p_t), since it does not allow instantaneous transitions on model order.

Acknowledgments

The authors are grateful to the editor and the referee for their valuable comments and suggestions. The first author was partially funded by Conicit-Venezuela. The second author was partially funded by Sistema Nacional de Investigadores and Conacyt-Mexico.

Appendix: posterior sampling algorithm

We describe the details to obtain samples from the full posterior distribution of $(\boldsymbol{\phi}_t, \sigma^2, p_t)$ given D_n for $t = 1, \ldots, n$ with the model specifications described in Section 3. Define $\Phi = \{\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n\}$, with $\boldsymbol{\phi}_t = (\phi_{t,1}, \ldots, \phi_{t,p_{\max}})'$, $\mathbf{X} = \{x_1, \ldots, x_n\}$, and $\mathbf{P} = \{p_1, \ldots, p_n\}$. We follow a Gibbs sampling format defined in two stages.

Stage 1. *Sampling from the full conditional distribution of* Φ *and* σ^2. This can be done by sampling Φ from $p(\Phi|\mathbf{X}, \mathbf{P}, \sigma^2)$ and sampling σ^2 from $p(\sigma^2|\mathbf{X}, \Phi, \mathbf{P})$. Conditional on $\mathbf{P}$, we have the DLM structure

$$\begin{aligned} x_t &= \mathbf{F}_t'\boldsymbol{\phi}_t + \epsilon_t \\ \boldsymbol{\phi}_t &= \boldsymbol{\phi}_{t-1} + \boldsymbol{\xi}_t \end{aligned}$$

with $\mathbf{F}_t = (x_{t-1}, \ldots, x_{t-p_t})'$. Assuming that the set of system covariance matrices $\{\mathbf{U}_t; t = 1, \ldots, n\}$ is assumed known or specified by a single discount factor (West and Harrison, 1997), efficient generation can be performed via *Forward Filtering-Backward Sampling* algorithms (Carter and Kohn, 1994). Now, sampling from $p(\sigma^2|\mathbf{X}, \Phi, \mathbf{P})$ reduces to computing

$$e = \sum_{t=1}^{n}(x_t - \sum_{j=1}^{p_t}\phi_{t,j}x_{t-j})^2,$$

and sample a distribution proportional to $p(\sigma^2)(\sigma^2)^{-(\alpha+1)}\exp\{-\beta/\sigma^2\}$ with $\alpha = n/2 - 1$ and $\beta = e/2$. Inverse gamma priors on σ^2 are conditionally conjugate.

Stage 2. *Sampling from* $p(\mathbf{P}|\mathbf{X}, \Phi, \sigma^2)$. Let D_t be the information up to time t, i.e., $D_t = \{D_0, \mathbf{X}_t, \Phi_t, \sigma^2\}$ with $\mathbf{X}_t = \{x_1, \ldots, x_t\}$, $\Phi_t = \{\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_t\}$. $P[p_t = i|p_{t-1} = j]$ defines the transition probabilities for model order between times $t-1$ and t. The filtering part of the algorithm requires computation and storage of $P[p_t = i|D_t]$ and $P[p_t = i|D_{t-1}]$, for $i = p_{\min}, \ldots, p_{\max}$ and $t = 1, \ldots, n$. By Bayes theorem,

$$P[p_t = i|D_t] \propto f(x_t|\phi_t, \sigma^2, p_t)P[p_t = i|D_{t-1}],$$

where $f(x_t|\phi_t, \sigma^2, p_t)$ is the likelihood function for the observation x_t that is easily obtained with the model definition of the TVAR(p_t). Additionally,

$$P[p_t = i|D_{t-1}] = \sum_{j=p_{\min}}^{p_{\max}} P[p_t = i|p_{t-1} = j]P[p_{t-1} = j|D_{t-1}],$$

with $P[p_{t-1} = j|D_{t-1}]$ the posterior probability evaluated at time $t-1$. Now, we apply the *Backward Simulation Sampling* algorithm step for discrete random variables presented in Carter and Kohn (1994). First, we generate a value p_n^* from $P[p_n = i|D_n]$. Then, for $t = n-1, n-2, \ldots, 1$, we compute $P[p_t = i|p_{t+1} = p_{t+1}^*, D_t]$ using

$$P[p_t = i|p_{t+1} = p_{t+1}^*, D_t] = \frac{P[p_{t+1} = p_{t+1}^*|p_t = i]P[p_t = i|D_t]}{P[p_{t+1} = p_{t+1}^*|D_t]},$$

where p_{t+1}^* is a generated value of the distribution $P[p_{t+1} = i|p_{t+2} = p_{t+2}^*, D_t]$. We sample a value p_t^* from $P[p_t = i|p_{t+1} = p_{t+1}^*, D_t]$ and continue until $t = 1$. The values $p_1^*, p_2^*, \ldots, p_n^*$ constitute a sample from the conditional posterior distribution of $\mathbf{P}$.

Bibliography

Andrieu, C., Frietas, N. De and Doucet, A. (1999) Sequential MCMC for Bayesian model selection. In *Proc. Work. IEEE HOS'99.*

Barbieri, M.M. and O'Hagan, A. (1997) A reversible jump MCMC sampler for Bayesian analysis of ARMA time series. Technical Report. Department of Statistics Universita "La Sapienza".

Barnett, G., Kohn, R. and Sheather, S. (1996) Bayesian estimation of an autoregressive model using Markov Chain Monte Carlo. *Journal of Econometrics*, **74,** 237–254.

Carter, C.K. and Kohn, R. (1994) Gibbs sampling for state space models. *Biometrika*, **81,** 541–53.

Frühwirth-Schnatter, S. (1994) Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15,** 183–202.

Gersch, W. (1987) Non-stationary multichannel time series analysis. In *EEG Handbook, Revised Series* (ed. A. Gevins), vol. 1, pp. 261–96. New York: Academic Press.

Godsill, S. J., A., Doucet and M., West (2000) Maximum *a posteriori* sequence estimation using Monte Carlo particle filters. *Annals of the Institute of Statistical Mathematics.*

Green, P.J. (1995) Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82,** 711–32.

Harrison, P.J. and Stevens, C.F. (1976) Bayesian Forecasting (with discussion). *Journal of the Royal Statistical Society-Series* B, **38,** 205–247.

Huerta, G. (1998) Bayesian analysis of latent structure in time series. Ph.D. Thesis. Duke University, Durham, NC.

Huerta, G. and West, M. (1999) Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society-Series* B, **61,** 881–899.

Kitagawa, G. and Gersch, W. (1996) *Smoothness Priors Analysis of Time Series,* Lecture Notes in Statistics, vol. 116. New-York: Springer-Verlag.

Krystal, A.D., Prado, R. and West, M. (1999) New methods of time series analysis of non-stationary EEG data: Eigenstructure decomposition of time varying autoregressions. *Clinical Neurophysiology*, **110,** 1–10.

Krystal, A.D., West, M., Prado, R., Greenside, H.S., Zoldi, S. and Weiner, R.D. (2000) The EEG effects of ECT: Implications for rTMS. *Depression and Anxiety* (to be published).

Pitt, M. and Shephard, N. (1999) Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, **94,** 590–599.

Prado, R. (1998) Latent structure in non-stationary time series. Ph.D. Thesis. Duke University, Durham, NC.

Prado, R. and West, M. (1997) Exploratory modelling of multiple non-stationary time series: Latent process structure and decompositions. In *Modelling Longitudinal and Spatially Correlated Data. Methods, Applications, and Future Directions*, *New York* (eds T.G. Greogoire, D.R. Brillinger, P.J. Diggle, E. Russek-Cohen, W.G. Warren and R.D. Wolfinger), Lecture Notes in Statistics 12. Springer Verlag.

Troughton, P.T. and Godsill, S.J. (1997) Bayesian model selection for time series using Markov Chain Monte Carlo. Technical Report. Signal Processing and Communications Laboratory. Department of Engineering, University of Cambridge.

West, M. and Harrison, J. (1997) *Bayesian Forecasting and Dynamic Linear Models (2nd Edn.)*. New York: Springer-Verlag.

West, M., Prado, R. and Krystal, A. (1999) Evaluation and comparison of EEG traces: Latent structure in non-stationary time series. *Journal of the American Statistical Association*, **94,** 1083–1095.