BAYESIAN STATISTICS 6, pp. 1–16
J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (Eds.)

Bayesian Inference on Latent Structure in Time Series

Omar Aguilar, Gabriel Huerta, Raquel Prado & Mike West
ISDS, Duke University, Durham NC 27708-0251

SUMMARY

A range of developments in Bayesian time series modelling in recent years has focussed on issues of identifying latent structure in time series. This has led to new uses and interpretations of existing theory for latent process decompositions of dynamic models, and to new models for univariate and multivariate time series. This article draws together concepts and modelling approaches that are central to applications of time series decomposition methods, and reviews recent modelling and applied developments. Several applications in time series analyses in geology, climatology, psychiatry and finance are discussed, as are related modelling directions and current research frontiers.

Keywords: DYNAMIC FACTOR MODELS; DYNAMIC LINEAR MODELS; MULTIPLE TIME SERIES; MULTIVARIATE STOCHASTIC VOLATILITY; NON-STATIONARY TIME SERIES; TIME-VARYING AUTOREGRESSION; TIME SERIES DECOMPOSITION.

1. INTRODUCTION

The identification and interpretation of latent processes underlying observed time series is an old and essentially archetypal problem in time series analysis. In recent years applied interests in a variety of fields have stimulated Bayesian time series research focussed on latent structure analysis. This has led to theoretical developments of new decomposition methods that has generated new methodology and associated computational tools for model fitting and exploration. A collection of related research directions in this area are tied together and reviewed here, with reference to ranges of applications and new research directions.

We begin in Section 2 with discussion of univariate dynamic models and highlight concepts and methods of time series decomposition to infer characteristics of underlying latent component processes. These results are very general, arising in a broad class of dynamic models, and utilise new interpretations of existing theory of decompositions of dynamic models. The theory is exemplified in the simple but very important special case of a time series with a latent autoregressive component, and with a summary of an application in a climatological modelling and prediction problem. The section continues with more general time-varying autoregressive models as components of a time series, where the generality and broader utility of decomposition results becomes apparent. The section concludes with a summary of analysis of a model in this class applied to a geological time series study. Section 3 takes these models further, with applied motivation drawn from studies of EEG time series arising in clinical psychiatric studies. This begins with further discussion of the methodology of univariate decompositions for non-stationary time series, and moves into comparisons of analyses across several related time series. The section concludes with discussion of issues and theory associated with latent factor structure in multivariate time series, including commentary on the extensions of univariate decomposition results to multivariate dynamic models. Section 4 introduces a special class of multivariate latent factor models, namely dynamic factor models with multivariate stochastic volatility components, of interest in financial time series analysis. Section 5 concludes with summary remarks on additional multivariate dynamic modelling developments and current research frontiers.

2. UNIVARIATE TIME SERIES DECOMPOSITIONS AND LATENT STRUCTURE

2.1. *Introduction*

Much of the recent development in latent structure analysis is based on novel extensions and exploitation of the fundamental component structure of dynamic models (West and Harrison 1997, chapter 6). Begin with a general dynamic linear model (DLM) in which the scalar time series y_t, observed at equally spaced time points $t = 1, 2, \ldots,$ is modelled as

$$y_t = x_t + \nu_t, \quad x_t = \mathbf{F}_t' \boldsymbol{\theta}_t, \quad \boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \tag{2.1}$$

for each t. Here x_t is the latent signal process, ν_t is an observation error, $\boldsymbol{\theta}_t = (\theta_{t,1}, \ldots, \theta_{t,d})'$ is the $d \times 1$ state vector, $\mathbf{F}_t$ is the column regression $d-$vector, $\mathbf{G}_t$ is the $d \times d$ state evolution matrix, and $\boldsymbol{\omega}_t$ is the stochastic state evolution noise, or innovation. Often the $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$ sequences are mutually uncorrelated white noise, though more complex structures are possible and sometimes useful.

This is a very general class of models. Important special cases discussed below include constant models, characterised by constant **F** and **G** elements, and time-varying parameter regressions and autoregressions, among others. The central decomposition result is that the signal process in (2.1) has the representation

$$x_t = \sum_{j=1}^{d_z} z_{t,j} + \sum_{j=1}^{d_a} a_{t,j} \tag{2.2}$$

where, for each j, $z_{t,j}$ and $a_{t,j}$ are underlying latent processes with specific and relatively simple structure. Both the numbers (d_z, d_a) and structure of these latent processes are model dependent, and the utility of the decomposition result is evident through their definitions and interpretations in specific special cases. Some key special cases exemplify this in the following sections. Full background and mathematical details of this construction are given in West, Prado and Krystal (1997), and in special cases in West (1997c) and West and Harrison (1997, sections 9.5, 9.5 and 15.3).

2.2. *Latent Structure, Prior Specifications and Model Uncertainty in Autoregressions*

The simplest, and important, special case is that of an autoregressive signal in noise, in which $x_t = \sum_{j=1}^{d} \phi_j x_{t-j} + \omega_t$. This is a special case of (2.1) with

$$\mathbf{F}_t = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{G}_t = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_{d-1} & \phi_d \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & 0 \end{pmatrix}, \quad \boldsymbol{\omega}_t = \begin{pmatrix} \omega_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{2.3}$$

for all t. Note the constancy of **F** and **G** terms. In this case, the eigenvalues of **G** are the reciprocals of the usual characteristic roots of the AR(d) model, and the decomposition (2.2) is related to the standard partial fractions expansion of AR processes. Suppose the reciprocal eigenvalues occur as d_z pairs of complex conjugates $r_j \exp(\pm i\alpha_j)$, $(j = 1, \ldots, d_z)$, and d_a real values r_j, $(j = 1, \ldots, d_a)$, where $2d_z + d_a = d$. Then $z_{t,j}$ is a latent, quasi-periodic ARMA$(2, 1)$ process– a damped sinusoid of time-varying amplitude and phase, and fixed frequency α_j (wavelength or period $2\pi/\alpha_j$.) Each $a_{t,j}$ is a simple AR(1) process with AR parameter r_j. In application, it is often the case that some of the latent processes, and particularly the $z_{t,j}$ processes, have physical interpretation. The utility of the theory is in identifying and estimating these processes, and then investigating possible physical meaning. The quasi-periodic components are often primary, in connection with time-varying periodicities in the time series under study.

It is common to fit higher-order AR models to provide empirical approximations to lower order ARMA models or non-linear features in the series. This means that latent components with lower modulii r_j, and/or very high frequencies α_j, are induced in order to capture correlation structure in the series but do not represent physically meaningful components. Thus the prior perspective in a new application is that a higher-order model is likely needed, but that only a small number of the latent components will have higher modulii and lower frequencies, and that these will be of primary interest. This leads to a focus on component structure in assessing prior distributions for both model parameters and model order d. One of the major developments arising from this is a novel class of smoothness priors developed on the eigenvalues of **G**. In addition to this novel and practical focus in prior specification, this carries an immediate benefit in dealing with model order uncertainty. This is developed in Huerta (1998), and Huerta and West (1997a,b), which include a range of examples in component assessment, prediction and spectral analysis. Our work here provides a very different approach to prior modelling and model uncertainty than other recent approaches (e.g., Barnett et al 1996) in AR models. Among other things, the very practical focus on component structure leads to new classes of smoothness priors on the ϕ_j coefficients, permits models of possibly very high order, trivially concentrates on the AR stationary region or its boundaries, and allows unit roots to model persistent low frequency trends and sustained periodicities (related to alternative approaches in West 1995 and 1996). This latter feature helps to resolve difficulties in inference on "spikes" in spectral densities in using AR models (Huerta and West 1997b) that arise in standard approaches where unit roots are disallowed. Model fitting involves customised MCMC methods that are fully developed, explored and exemplified in the above references, with direct extensions to problems of missing observations and data analytic issues. We note further that, though our focus on component structure and assessment of latent processes is novel, others have explored inference on autoregressive root structure under more standard prior distributions (e.g., Geweke, 1988,89).

Finally, though not pursued here, the focus on component structure of state space models has also generated a quite novel theory of *continuous* time state space models (Huerta 1998), to be reported in the near future. This has been important, for example, in analysis of geological time series which can have very erratic spacings, and small numbers of very long gaps between consecutive observations. These developments also have potential uses in addressing issues of timing uncertainty (West 1996, 97a; Li 1997).

A currently very topical example, drawn from Huerta and West (1997a), concerns a series of observations on the *Southern Oscillation Index* (SOI) of interest in monitoring global climatic variability. This is a series of 540 monthly measurements (during 1950-1995) of the "difference of the departure from the long-term monthly mean sea level pressures" at Tahiti in the South Pacific and Darwin in Northern Australia (Trenberth and Hoar 1996) known as the El Niño southern oscillation series. The series, in the upper frame of Figure 1, oscillates about zero and the predominance of negative values in more recent years is related to a recent warming in the tropical Pacific. Of key climatological concern has been the long run of 22 months of consecutive negative values during the last two years of the series. Trenberth and Hoar (1996) conclude that, under a stationary ARMA model for the data selected from a set of possible models, this run is so unlikely as to cast serious doubt on stationarity and suggest structural changes in favour of a warming trend.

One analysis is summarised in Figure 1. The smoothness prior structure adopted here allows models of order up to $d = 40$. Figure 1 displays the resulting posterior for d (over the range plotted, the prior for d is essentially uniform). Uncertainty about d is high, though values in the 8-14 range are indicated. Figure 1 graphs posterior means of three identified dominant latent $z_{t,j}$ components over time, each on the same vertical scale as the original data. The posterior indicates that three quasi-periodic $z_{t,j}$ components with largest modulii r_j also have largest amplitudes; the trajectories displayed are of posterior means for these three components. The first is clearly dominant, and has a period of around 4 to 5 years, consistent with historical analyses of El Niño periodicities. Beyond the displayed components, subsidiary components all have lower amplitudes. It should be stressed that these inferences fully incorporate model order uncertainty as these are posterior estimates averaging over the posterior for d; we do not know how many components there are, but there are clearly at least three $z_{t,j}$ components.

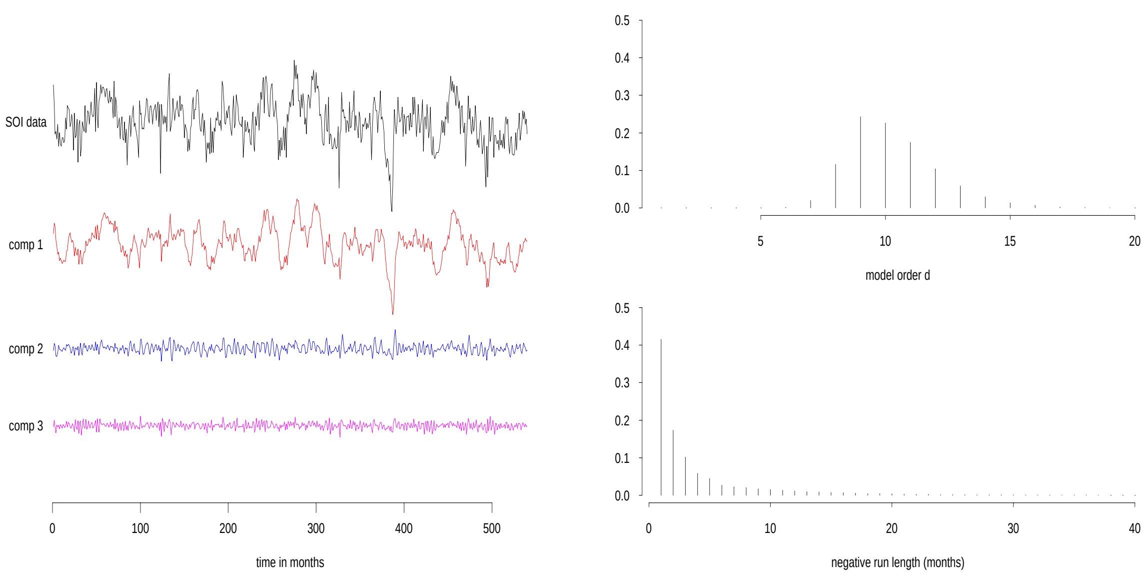

Figure 1. *Aspects of SOI analysis. Left frame: SOI series and three latent components. Upper right frame: Posterior distribution on model order d. Lower right frame: Predictive distribution for run length of negative values in the next 540 months.*

Our assessment of the recent run of 22 negative values is purely predictive. From the model we repeatedly simulate the next 540 months, generating "sample futures" over a time span equal to that of the data. This trivially delivers a Monte Carlo approximation to the predictive distribution of a run length of 22 negatives during that time period; see Figure 1. The probability of runs of length 22 or more is 0.02, which may be compared with the corresponding value of 0.00015 assessed by Trenberth and Hoar (1996). Our view is that these authors are wildly extreme in their assessment, partly due to ignoring both model and parameter uncertainties that are appropriately represented in our analysis. We conclude that the chance of such a run reoccurring in the next 45 years is only about 1 in 50, so the event is indeed unusual under a stationary linear model, though by no means as unusual as the earlier analysis suggested.

2.3. *Nonstationary Time Series Decompositions*

A range of recent developments and applications of decomposition analyses in *non-stationary* models have adapted the above framework to that of *time-varying autoregressions* (e.g., Kitagawa and Gersch 1996, chapter 11; West and Harrison 1997, section 9.6). These simply adopt the model (2.1, 2.3) but now with time-varying AR parameters $\boldsymbol{\phi}_t = (\phi_{t,1}, \ldots, \phi_{t,d})'$ in the first row of $\mathbf{G}_t$. An additional evolution model for the dynamic variation of $\boldsymbol{\phi}_t$ completes the specification, and this is often taken as a neutral random walk $\boldsymbol{\phi}_t = \boldsymbol{\phi}_{t-1} + \boldsymbol{\delta}_t$ for some zero-mean innovation $\boldsymbol{\delta}_t$ (e.g., West, Prado and Krystal 1997). Now the decomposition (2.2) still arises, though with subtle differences due to the time-variation in model parameters $\mathbf{G}_t$. At each time point $t,$ the latent processes $z_{t,j}$ and $a_{t,j}$ exist and have the instantaneous forms of ARMA$(2,1)$ and AR(1) processes, but now the defining modulii and frequencies are subject to change over time. There is also an element of linear mixing of the latent processes through time that slightly changes the theory, though this is generally completely negligible from a practical viewpoint (see discussion in West, Prado and Krystal 1997, and Prado 1998). In cases where the identified latent processes $z_{t,j}$ have physical interpretation, the changes over time in their frequencies $\alpha_{t,j}$ and modulii $r_{t,j}$ (both now indexed by t as they are time-varying) are often of primary interest. These changes represent patterns of time-variation in spectral characteristics of the signal x_t. Assessment of changes over time in these parameters can be viewed as a novel approach to exploring time-variation in spectral density functions, and the corresponding decomposition analysis represents a form of spectral decomposition in the time domain.

An example from geology concerns patterns of relative change in latent processes underlying variation in levels of geochemical indicators spanning the last few million years. Figure 2 graphs a series of oxygen isotope measures during the last 2.5million years (one of a collection of such series, another of which was presented and analysed with a slightly different, constant parameter model in West 1997a; see also West 1997b,c). The data measure relative abundance of δ^{18}O to δ^{16}O on a time scale of 3000 year (3kyr) intervals stretching back roughly 2.5Myr. The nature of changes over time in the structure of the evident time-varying periodicities in the data, induced by the forcing earth-orbital periodicities, are of interest in connection, particularly, with the changing nature of the so-called $\sim$ 100kyr "ice-age cycle" (Park and Maasch 1993). An extension of the time-varying AR model to incorporate a locally-constant, first-order polynomial trend component is needed to adequately model the apparent increasing levels over this time period, associated with generally increasing global temperatures in more recent times (note that the data is plotted with time reversed). That apart, the decomposition analysis applies for any specified model, and analysis here is based on the trend plus time-varying AR(20) model. The resulting posterior means of four dominant $z_{t,j}$ components are graphed in the left frame of Figure 2, on the same vertical scales for comparison; the residual components are negligible by comparison, so that the data series is essentially the sum of the trend plus four oscillatory components graphed. Corresponding to these four latent processes are their individual wavelengths (or periods), now time-varying. The posterior means of the time trajectories of the periods appear in the lower right frame of Figure 2, corresponding in order, from the top down, to the latent processes labelled 1 to 4. We see that the estimated periods are very stable, almost constant over the time interval, and the values correspond to the known ranges of cycle lengths of the main earth-orbital cycles: that is, around 110kyr for cycles induced by the *eccentricity* of the earth's rotation about its axis, around 41-43kyr for the *obliquity* of the Earth's orbit around the Sun, and two possible *precessionary* cycles induced by precession of the Earth about its axis, with periods of around 19 and 23kyr respectively. The changing form of the dominant eccentricity or "ice-age cycle" is of particular interest. The vertical lines in the figure indicate the point (and the only point) at which the estimates of the relative amplitudes of the two dominant latent cycles switched order, at about 1.1million year ago. Prior to that point, the ice-age cycle had lesser amplitude; characterising the timing of this switch-over is of current geological interest in connection with whether or not the increased significance of this cycle was gradual, or the result of abrupt and significant structural climatic change (Park and Maasch 1993). Our analysis does not model change points per se, of course, and this initial study suggests that an elaboration to include the possibility of abrupt changes in the AR parameters would be of interest to help resolve this issue. Nevertheless, this initial "routine" decomposition analysis provides interesting insights, and exemplifies the use of such methods for spectral decompositions in the time domain.

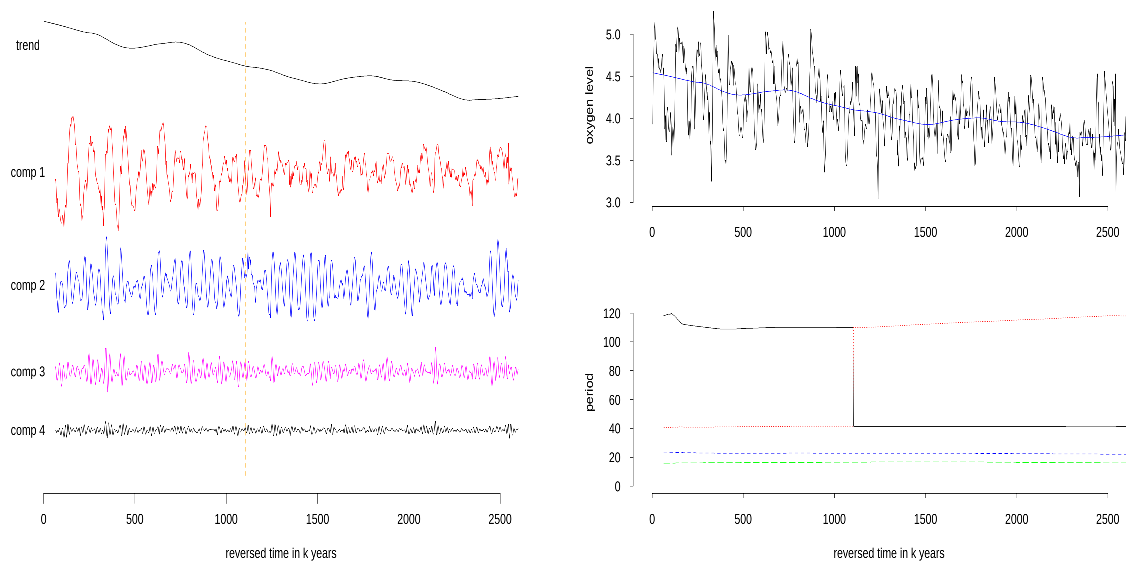

Figure 2. *Decomposition analysis of oxygen isotope series. Left frame (top down): Estimates of underlying trend and first four quasi-cyclical latent processes in isotope series. Upper right frame: Data with estimated trend. Lower right frame: Trajectories of the estimated periods (in thousands of years) of the four dominant quasi-cyclical components. The vertical line in the period graph indicates the point at which the relative amplitudes of the first two components switched.*

3. LATENT STRUCTURES IN MULTIPLE TIME SERIES

3.1. *Exploring Related Univariate Decompositions*

A range of recent applied studies has been generated in the area of EEG analysis (Prado and West 1997, West, Prado and Krystal 1997). In treatment of human subjects with critical neuro-psychiatric disorders, such as major depression, electroconvulsive (ECT) therapy is among the most effective clinical treatments (Weiner and Krystal 1993). In ECT a brain seizure is evoked, inducing major increases in amplitudes and frequencies of electrical potential fluctuations measuring neural communications. Effective seizures are characterised by longer lasting, high amplitude fluctuations, and clinical psychiatrists investigate treatment effects by exploring differing patterns of seizure stimulus. The resulting electroencephalographic (EEG) traces – long time series of potential fluctuations at various scalp locations – provide the main data on which to compare and assess the differences in characteristics of seizures. As illustrated in the above references, time series decomposition methods allow us to characterise the time-frequency structure of such records by isolating latent components representing fluctuations in various key frequency ranges. The comparison of the time courses of parameters defining these latent components (amplitudes, frequencies/wavelengths, modulii) provide accessible graphical displays for evaluation of the nature of the seizure effects in different wavebands, and for comparisons across seizures for the same patient but under differing ECT treatment controls.

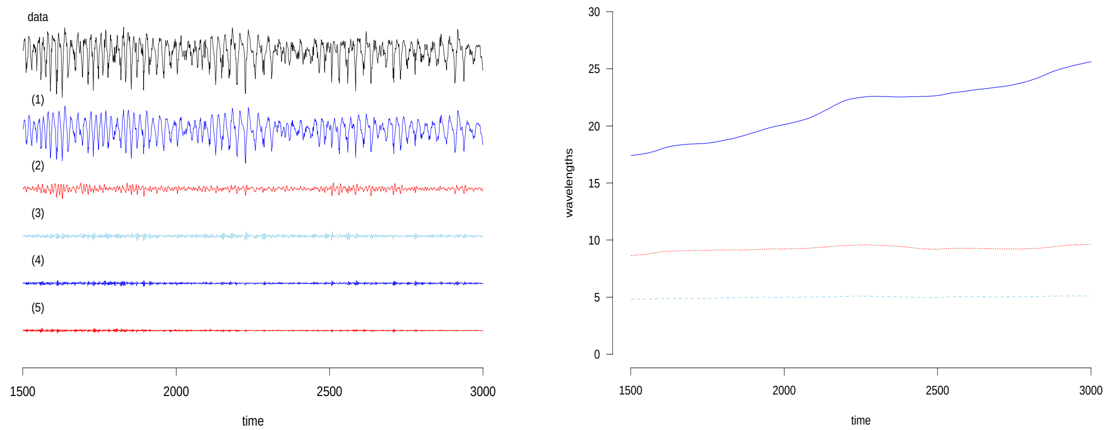

Figure 3. *Decomposition analysis of EEG series: channel 18. Left frame: Time series and estimated trajectories of five main latent components. Right frame: estimated trajectories of the time-varying wavelengths of the three dominant components.*

Figure 3 displays aspects of an analysis of an EEG series from a dataset in Prado and West (1997). The data in the first frame is a section of an EEG trace from one of 19 "channels" at different scalp locations on a patient in seizure. Analysis of EEG data using time-varying autoregressions has a history, with notable contributions in the work of Will Gersch and coauthors (see references in Kitagawa and Gersch 1996), and our developments of latent factor decompositions complement and extend such prior work in this applied field. The model assumes $d = 12$ and time variation is modelled using standard discount factor methods. The graph displays estimated trajectories of five latent, quasi-cyclical $z_{t,j}$ processes, displayed from the top down in order of increasing frequencies (decreasing wavelengths). These represent brain potential fluctuations in key wavebands. The data and components are graphed on the same scale for comparison. Trajectories of the estimated wavelengths of the first three, key components appear in the right frame of the figure. On the time scale here, the wavelength of the first component varies between 17-25 which, converting to the original time scale of the data corresponds to a frequency range of between 1.5 and 3 cycles per second (cps), the *delta* waveband of brain fluctuations. This is the *delta "slow wave"* that dominates seizure induced activity. Over the full time course of a seizure, this wave is initially negligible but then, as the seizure takes hold, increases rapidly in amplitude to become the dominant feature, then eventually decays as the seizure tails off. The section of seizure displayed here is in the central, active part; related analyses in Prado and West (1997) display similar graphs for other channels in which the early and late stages are included. A key feature here is the gradual increase in wavelength of the delta wave, corresponding to decreasing frequency of fluctuations as the seizure matures and begins to decay. Related graphs of time trajectories of the corresponding modulii and amplitudes indicate general stability of these parameters over time, but with a gradual decay of the amplitude of the delta wave consistent with the evolution of the seizure. Components 2 and 3 are in the *theta* $(4 - 8$cps) and *alpha* $(8 - 12$cps) wavebands respectively, and represent "normal" brain activity that is distorted by the seizure but whose characteristics nevertheless remain relatively stable over time as identified by the decomposition analysis. A similar comment applies to the lower amplitude components 4 and 5, the so-called *fast waves* in the *beta* waveband. This kind of analysis is a first in the sense of providing deconvolution of the EEG series into isolated and identifiable components corresponding to standard frequency ranges, and the resulting assessment of patterns of change in defining parameters over time that characterise the response to the specific ECT treatment. Further examples and discussion, including explicit comparisons of two treatments, appear in West, Prado and Krystal (1997).

Across the full set of 19 EEG channels, similar analyses yield similar inferences: each channel represents a noisy convolution of many underlying brain processes that interact with time-varying characteristics and that are evidenced in individual, univariate analyses. Comparisons of individual univariate decomposition analyses do provide insights into issues of variations across channels. For example, the left frame in Figure 4 plots the trajectory of the wavelength parameter of the delta waveform, estimated from each channel separately. The dominant delta wave is very stable across channels, and the obvious commonality indicates an underlying, "driving" mechanism, and suggests a multivariate model with all channels driven by a time-varying seizure process. Note that the time scale here now represents nearly all of the seizure record, and we see that the estimated wavelengths vary more widely in the later stages of the seizure record, primarily as a result of increasing uncertainty about this delta wave as it decays in amplitude. The right frame in Figure 4 displays related estimates of amplitudes of the dominant delta component across channels. For each individual channel analysis, we simply graph the estimated amplitude (in terms of voltage potential level) at four selected time points (t =1000, 2000, 3000 and 3650) during the seizure time course. Note the apparent cyclical form as a function of channel index; though the waveform is consistent across channels, there are small but significant lags between channels that induce this periodic graph and that can be related to the physical layout of the channels on the scalp, as discussed in Prado and West (1997). Spatially contiguous locations do tend to be more highly related in terms of their temporal patterns of change in amplitude. In addition, there is evidence that the lag structures across channels are evidently slowly varying in time, which suggests that very significant complications will be faced by multivariate models that attempt to isolate common underlying latent processes. Comparative evaluation of independent univariate decompositions is, in contrast, immediately accessible.

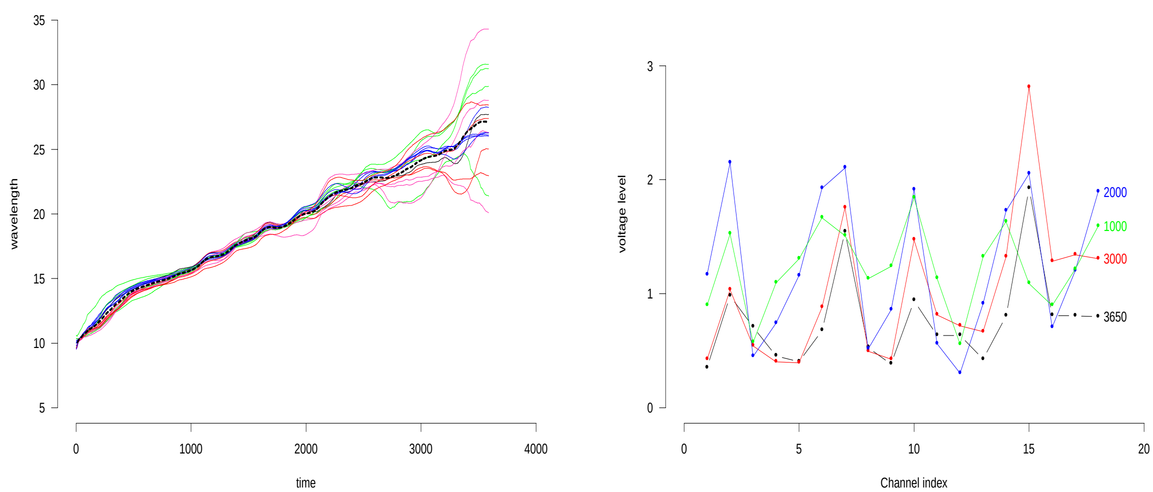

Figure 4. *Multiple EEG series: 19 channels. Left frame: Trajectories of wavelength of dominant latent component from independent univariate models. Right frame: Amplitudes of dominant latent component in independent univariate series at four time points during the seizure.*

3.2. *Multivariate Models and Decompositions*

The multiple EEG series framework is an example of a range of problems in which it is desirable to introduce multivariate models involving underlying latent factor processes. Various such models and approaches have been introduced during the last couple of decades, including, for example, the foundational work of Peña and Box (1987), contributions by Tiao and Tsay (1989) and Escribano and Peña (1994) related to cointegration and common component time series models, and, more specific to the Bayesian forecasting world, dynamic hierarchical models as developed by Gamerman and Migon (1993). We explore some basic ideas of factor modelling here.

Consider m parallel time series $y_{t,i}$, $(i = 1, \dots, m)$, that are driven by $k < m$ underlying latent processes $x_{t,j}$, $(j = 1, \dots, k)$, in a time-varying, dynamic linear model. Writing $\mathbf{y}_t$ for the $m-$vector time series and $\mathbf{x}_t$ for the $k-$vector latent process, a rather general framework is generated by the dynamic model

$$\mathbf{y}_t = \mathbf{B}_t \mathbf{x}_t + \boldsymbol{\nu}_t \tag{3.1}$$

over $t = 1, \dots, n$, where the $\boldsymbol{\nu}_t$ are zero-mean observation error vectors, and the $\mathbf{B}_t$ are dynamic regression matrices to be estimated. In factor analytic terminology, $\mathbf{x}_t$ represents a latent factor process and the $\mathbf{B}_t$ are time-varying factor loadings. The EEG context, for example, is suggestive of a model of the form (3.1) in which the elements of $\mathbf{x}_t$ are current and lagged values of the latent waveforms in various frequency bands, in which case they will be appropriately modelled via time-varying autoregressions, or possibly time-varying ARMA models. At a level of generality similar to (3.1), such models and others can be represented by a time-varying vector autoregression for the $\mathbf{x}_t$ process directly. Combined with (3.1), this generates a very rich and flexible structure for dynamic latent factor modelling, but one that is a long way from providing practical and implementable methodology. The issues of parametrisation and identification are key, raising questions about appropriate structuring and parametrisation of the $\mathbf{x}_t$ process model, and of strict parametric constraints on the $\mathbf{B}_t$ matrices. The framework is intriguing but the research agenda only just entered (Prado and West 1997).

Theoretical progress has been made in exploring and extending the time series decomposition results discussed earlier to multivariate settings. The results are perhaps surprisingly complete theoretically, though they have yet to be explored and exploited in practice. Consider equation (3.1) and suppose that $\mathbf{x}_t$ follows a time-varying parameter vector autoregressive model. Then, extending the foundational results of Peña and Box (1987) to time-varying parameter contexts, it follows that $\mathbf{y}_t$ itself follows a time-varying vector ARMA model. Often this will be adequately represented by a higher-order vector autoregression, again with time-varying parameters, and the focus for decomposition theory may therefore be restricted to this class. A general model has

$$\mathbf{y}_t = \sum_{j=1}^{d} \Phi_{t,j} \mathbf{y}_{t-j} + \boldsymbol{\omega}_t \tag{3.2}$$

where the $\Phi_{t,j}$ are time-varying coefficient matrices and $\boldsymbol{\omega}_t$ is a sequence of zero-mean vector innovations. It turns out that the univariate decomposition (2.2) has a direct extension to this multivariate context, obtained by casting (3.2) in state-space form. Full details can be found in Prado (1998). Key aspects of this result relate to the implied component structure of the univariate elements of $\mathbf{y}_t$. In particular, each $y_{t,i}$ has a decomposition of the form (2.2) and every characteristic component frequency and modulus appears in each of the $y_{t,i}$. This result is surprising, and suggestive of potential decomposition methodology arising from time-varying vector AR models. It is perhaps best appreciated in the case of constant parameters $\Phi_{t,j} = \Phi_j$ for all t. In such a case, each $y_{t,i}$ series has a decomposition as the sum of several AR(1) and

ARMA(2,1) processes. Focus on the ARMA(2,1) processes initially – it turns out that there are a subset of ARMA(2,1) component frequencies and modulii that are common across the $y_{t,i}$, though the observed amplitude and phase characteristics are specific to series i. A similar comment applies to the latent AR(1) processes. Hence we have a framework in which underlying latent phenomena "drive" the univariate output series $y_{t,i}$, but in which the effects of the latent processes are distorted and convoluted by factors specific to the individual output series. This encompasses, for example, an underlying waveform of a fixed or time-varying frequency arising as a component of each of the $y_{t,i}$ but with time-variation in lag structure across series i, or precisely the kind of phenomena observed with multiple EEG series. Future development of time-varying vector models is therefore suggested, and this link-up is also natural in view of existing studies of EEG and related series with vector autoregressions (Kitagawa and Gersch 1996), which may now be extended in practically important ways based on this novel perspective on latent components. Future applied work here will exploit this theory.

4. DYNAMIC FACTOR MODELS FOR MULTIPLE TIME SERIES

Recent methodological research in structured dynamic latent factor models has been motivated by financial forecasting problems where time-varying volatility plays a key role. Aguilar and West (1998a) develop and illustrate such models in forecasting and portfolio construction for multiple time series of international exchange rates, relating to and building on foundational work Quintana (1992), Putnam and Quintana (1994), and Quintana and Putnam (1996), and following earlier modelling developments as reported in Quintana and West (1987). These *dynamic factor models* are direct generalisations of univariate stochastic volatility models, as mentioned by Harvey, Ruiz and Shephard (1994), Jacquier, Polson and Rossi (1994, 95), and Kim, Shephard and Chib (1998). In particular, Aguilar and West (1998a) extend models and associated computational methods suggested in Kim, Shephard and Chib (1998), and their work is intimately related to the important and independent studies of Pitt and Shephard (1998).

At a general level, dynamic factor models are based on (3.1) with specific and rather structured models for the $\mathbf{B}_t$ and $\mathbf{x}_t$ processes. For an $m-$vector time series of financial returns $\mathbf{y}_t$, the primary model extends (3.1) to

$$\mathbf{y}_t = \boldsymbol{\theta}_t + \mathbf{B}_t\mathbf{x}_t + \boldsymbol{\nu}_t \tag{4.1}$$

where

- $\boldsymbol{\theta}_t$ is a dynamic level vector, possibly including regression on econometric predictors;
- $\mathbf{x}_t = (x_{t,1} \dots, x_{t,k})'$ is a $k-$vector latent factor process such that $\mathbf{x}_t \sim N(\mathbf{0}, \mathbf{H}_t)$ are conditionally independent with diagonal variance matrix $\mathbf{H}_t$ changing over time;
- $\boldsymbol{\nu}_t \sim N(\mathbf{0}, \Psi_t)$ with diagonal variance matrix Ψ_t; and
- $\boldsymbol{\nu}_t$ and $\mathbf{x}_s$ are mutually independent for all t, s.

The central notion here is that, beyond the potential to predict returns as modelled in $\boldsymbol{\theta}_t$, patterns of residual variation and, critically, correlation among the univariate series in $\mathbf{y}_t$ may potentially be explained by a small number of latent factor processes, so that k is often much less than m. Conditional on all quantities but the latent factor $\mathbf{x}_t$, the instantaneous variance matrix for $\mathbf{y}_t$ is

$$\Sigma_t = \mathbf{B}_t\mathbf{H}_t\mathbf{B}_t' + \Psi_t \tag{4.2}$$

which is of standard factor form but with possibly time-varying components. If volatility structure exhibited in the financial outcome series $\mathbf{y}_t$ is in fact largely determined by structural changes in volatility of a few underlying factor processes, and if this underlying volatility can be adequately modelled, there is resulting potential for meaningful gains: increased accuracy in forecasting changes in variance-covariance patterns in Σ_t in (4.2) that are critical in determining portfolio allocation decisions (Quintana 1992; Quintana and Putnam 1996; Aguilar and West 1998a); and increased control over resulting investment risks due to increased opportunities for informed interventions based on econometric interpretations of the underlying factor processes. These issues are key motivations behind recent work with these models.

Aguilar and West (1998a) use multivariate extensions of standard stochastic volatility models. Take $\mathbf{H}_t$ as diagonal with elements $\exp(\lambda_{t,j})$, where $\lambda_{t,j}$ is the log of the instantaneous variance of the $j^{\rm th}$ latent factor at time t. Write $\boldsymbol{\lambda}_t$ for the vector of the $\lambda_{t,j}$. The model adopts a stationary vector autoregression of order one, namely

$$\boldsymbol{\lambda}_t = \boldsymbol{\mu} + \Phi(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\omega}_t \tag{4.3}$$

with independent innovations $\boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{U})$ for some innovations variance matrix **U**. Note that our models explicitly allow contemporaneous dependencies between the innovations impacting the volatilities across factors through the general variance matrix **U**, a feature that is strongly supported in data analyses such as that summarised below. Model completion requires imposition of identifying parametric constraints to determine a specific model within this dynamic factor model class. Aguilar and West (1998a) study international exchange rates in a model that adopts a constant factor loading matrix $\mathbf{B}_t = \mathbf{B}$ whose upper triangular elements are zero, a parametrisation used, for example, in Geweke and Zhou (1996). They also constrain to constant level parameters $\boldsymbol{\theta}_t = \boldsymbol{\theta}$ and series-specific residual variances $\Psi_t = \Psi = \mathrm{diag}(\psi_1, \ldots, \psi_m)$ for that specific, preliminary analysis and illustration of the modelling approach. Model implementation requires, in addition, informed prior distributions and the development of numerical methods. It is beyond our scope here to review the computational algorithms used in Aguilar and West (1998a), but note that they they involve a range of customised MCMC components originating from Kim, Shephard and Chib (1998), drawing on Geweke and Zhou (1996) and using techniques in West and Aguilar (1997).

Figure 5 displays some aspects of an analysis of exchange rates, relative to the $US, of a collection of 12 major international currencies (extending the preliminary study of 6 currencies in Aguilar and West 1998a). In the specific model used here, we analysed the returns on daily spot rates over a period of several years, a total of 2,561 observations on each series. The factor model allowed $k = 6$ factors, and of these we focus on two for illustration: these are two factors related primarily to the Japanese Yen and the British Pound, respectively, but have the interpretations of factors for these currencies that are already "corrected" for movements in exchange rates of other key currencies, namely the German (Deutsch) Mark and Canadian Dollar, relative to the $US dollar. For our purposes here, the two latent factors of interest can be viewed as, and are referred to as, the *Japenese Yen factor* and the *British Pound factor* underlying co-movements in the full set of 12 exchange rate returns series. Further details of the model, analysis and its uses will be reported in a forthcoming article. In Figure 5 the left hand frames relate to the Japanse Yen, the right hand frames to the British pound. The daily spot rates appear in the first row of figures, and the daily returns (defined simply as $Spot_t/Spot_{t-1} - 1$) in the second row. Our strategy for data analysis and computation is fully discussed in Aguilar and West (1998a). Resulting posterior samples for all model parameters and latent processes may be explored and summarised in various ways for posterior and predictive inferences, and for model checking. Figure 5 displays some selected posterior summaries for key processes related to the Yen and Pound series from the current analysis.

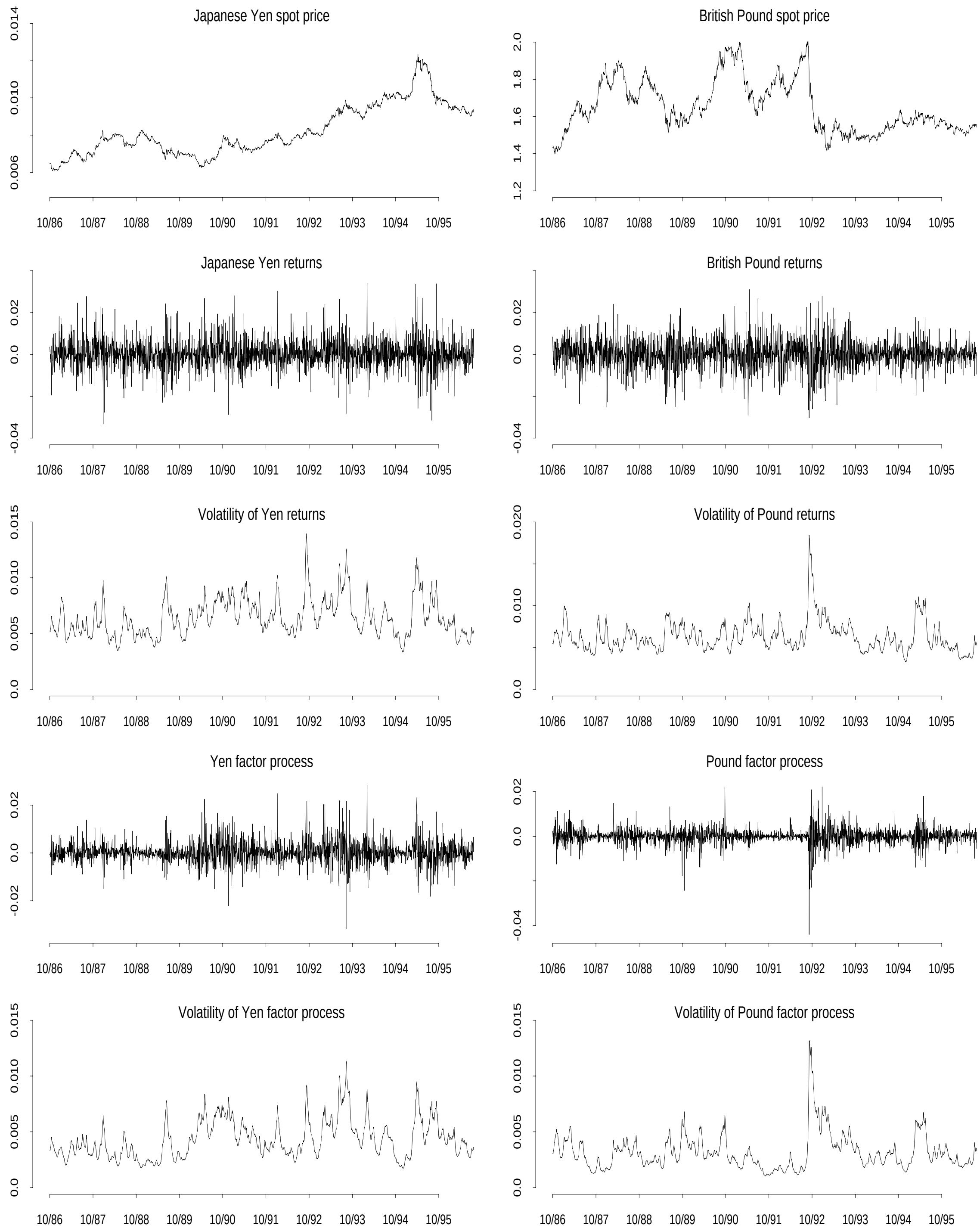

Figure 5. *Aspects of analysis of 12 international exchange rate series under a dynamic factor model. First row: Daily exchange rates of Yen and Pound in $US. Second row: Daily returns on exchange rates. Third row: Estimated volatility processes of Yen and Pound return series. Fourth row: Estimated Yen and Pound latent factor processes underlying co-movements in the full set of 12 exchange rate returns series. Fifth row: Estimated volatilities of Yen and Pound factor processes.*

The graphs in Figure 5 present estimated time trajectories, in terms of posterior means, for three key latent processes related to each of the two currencies. The third row of frames in Figure 5 displays such estimates for the conditional standard deviations of the Yen and Pound return series, respectively; at each time t, these are simply the square roots of the corresponding diagonal elements of the "current" conditional variance-covariance matrix Σ_t of the vector time series. To some extent these kinds of graphs provide for informal, subjective model checking; the estimated volatility series should appropriately "match" the observed fluctuations in the return time series, as we see is the case here. In addition, off-line parallel analysis of univariate return series using standard univariate stochastic volatility models (the special case of $m = k = 1$ here) can be routinely performed to provide comparisons that assist in validation of the MCMC analysis. A key feature of this time series is the major "event" associated with the forced withdrawal of Britain from the European Monetary System (EMS) in September 1992, resulting in increased volatility patterns in the series captured by this model. The associated imposition of target currency bands for several EU currencies in September 1993, that played a key role in breaking the EMS (Quintana and Putnam 1996, section 5), is reflected in additional spurts of increased volatility in the Yen series, though hardly impacts the Pound. The fourth row of frames displays the posterior means of the Yen and Pound factor processes $x_{t,\cdot}$ over time. The final row displays the time series of estimated conditional standard deviations of these two factors, namely the posterior means of $\exp(\lambda_{t,j}/2)$ for all t and $j = 1, 2$. On the log scale, the volatility AR processes are highly persistent, with posterior distributions for the diagonal matrix of AR coefficients Φ strongly supporting values in the $0.95 - 1$ range; this is indicative of the potential for improved short-term forecasting of changes in volatility under such models relative to the more standard "random walk" models underlying Bayesian variance matrix discounting (Aguilar and West 1998a; Quintana and West 1987) that are used in this area. Some initial indications of realisation of this potential are discussed in the assessment of portfolio allocation decisions in Aguilar and West (1998a), and full details of this, and of the application excerpted here, can be found in Aguilar (1998). Of course, a critical applied interest in these models, and variants of them, lies in the development of efficient and accurate methods for *sequential* analysis, as opposed to the "batch" analyses reported here. Some preliminary investigations of sequential analyses can be found in the closely related work of Pitt and Shephard (1998) using slightly different models and methods.

5. RELATED DEVELOPMENTS AND CURRENT RESEARCH

Related developments of non-linear and non-normal Bayesian models with multivariate latent structure are explored in Cargnoni, Muller and West (1997), Aguilar and West (1998b) and West and Aguilar (1997). These models fall under the general umbrella heading of multivariate, discrete dynamic models for longitudinal studies, but all share latent process structure closely linked to the developments discussed in this review. They also represent non-linear, multivariate extensions of dynamic generalised linear models (West 1985; West, Harrison and Migon 1985; West and Harrison 1997, chapter 14). The examples in Cargnoni, Muller and West (1997) involve collections of related time series of conditionally multinomial outcomes, with dynamic linear models providing the structure for variation over time, and cross-sectionally, in functions of the sets of multinomial probabilities. The developments in Aguilar and West (1998b) and West and Aguilar (1997) arise from motivating applications in institutional assessment and monitoring, and involve multivariate hierarchical models evolving in time, adding structured time series components to more standard hierarchical models used in this field (e.g., Christiansen and Morris 1997; Normand, Glickman and Gatsonis 1997). Here we are dealing with a large number of related time series of conditionally binomial counts, and adopt multivariate, latent factor dynamic models for vectors of parameters defining the sets of binomial outcome probabilities. Components of these dynamic models represent vectors of institution-specific random effects that are typically highly related over time, so that model components such as (4.3) arise as natural models for embedded latent factor processes. In addition to providing transfer of modelling ideas and concepts, the commonalities of mathematical structures across these various models has been important in aiding development of MCMC algorithms for model fitting and computation. Looking ahead, we anticipate broader application of these longitudinal dynamic models in various socio-economic contexts.

The above comments speak to the broader interest in dynamic latent process models and time series decomposition beyond the more traditional linear modelling framework of Section 2 and 3. In fact, the linear framework itself is deceptive; the fundamental equation (2.1) effectively covers many non-linear models too. To see this, note that the state matrices $\mathbf{G}_t$ and the distributions of the innovations $\boldsymbol{\omega}_t$ can be specified as we choose. In typical applied DLMs, the innovations are normally distributed, or conditionally normal based on hyperparameters that induce normal mixture models to allow for stochastic jumps in state vectors. However, we are free to specify these distributions, and may make them conditional on past state vectors $\boldsymbol{\theta}_s$ for $s < t$. Similarly, the $\mathbf{G}_t$ sequence may be functionally dependent on past state vectors, which means that the framework allows essentially arbitrary *state-dependent* models (Priestley 1980), and hence arbitrary non-linear structures (Tong 1990). Perhaps surprisingly, the decomposition results arise in such cases too. Hence, though these sections appear restricted to time-varying linear models, the generality is notable and indicative of potential for future work with latent structure in state-dependent, non-linear time series.

MCMC methods for model implementation are obviously quite fundamental to this field. In addition to building on and adapting general simulation methods from the growing toolboxes available to Bayesians, methods more specific to time series structures are critical in delivering tried, tested and efficient algorithms (e.g., Carter and Kohn 1994, de Jong and Shephard 1995). One key area of current research interest is focussed on adapting simulation methods to a sequential analysis context. In the dynamic factor framework of Section 4, for example, investment decisions are made sequentially in time as new data is processed and the model inferences adapt to a changing, possibly quite volatile, environment. In this area, in particular, the applied benefits are only beginning to be exploited, and advances in computational methods that enable serious real-time, analysis – that is both statistically and computationally efficient – are likely to have measurable applied impact. We are currently some way from this goal, in any but relatively "small" models, but recent work, that builds on imputation and adaptation methods of Berzuini *et al* (1998), Liu and Chen (1997), Pitt and Shephard (1997) and West (1993), has been encouraging. In their parallel development of dynamic factor models, Pitt and Shephard (1998) demonstrate impressive preliminary results in using these sequential methods. In addition to computation, there are very many issues of model structure, specification and choice to be explored in further work on dynamic factor models. We anticipate an active and exciting near future for this specific area, and for research more widely concerned with the utility of models and methods focussed on latent structure in time series.

ACKNOWLEDGEMENTS

Research reported here was partially supported by the National Science Foundation under grants DMS-9704432 and DMS-9707914. The authors acknowledge useful discussions with Merlise Clyde, Andrew Krystal, Hedibert Lopes, Peter Muller, Giovanni Parmigiani, Jose M Quintana and Neil Shephard. We are grateful to the providers of the various data sets used in this report, including Chris Wikle (the SOI data), Jeff Park (the Oxygen data), and Andrew Krystal (the EEG data). The financial time series were obtained from data vendor DataStream.

REFERENCES

Aguilar, O. (1998). Latent structure in Bayesian multivariate time series models. *PhD thesis,* Duke University.

Aguilar, O. and West, M. (1998a). Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. *ISDS Discussion Paper 98-03,* Duke University.

Aguilar, O. and West, M. (1998b). Analysis of hospital quality monitors using hierarchical time series models. *Bayesian Statistics in Science and Technology: Case Studies 4* (C. Gatsonis, R. Kass, B.P. Carlin, A.L. Carriquiry, A. Gelman. I. Verdinelli and M. West, eds.). New York: Springer-Verlag (to appear).

Berzuini, C., Best, N.G., Gilks, W.R. and Larizza, C. (1998). Dynamic conditional independence models and Markov chain Monte Carlo methods. *J. Amer. Statist. Assoc.* **92**, 1403-1412.

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

Cargnoni, C., Muller, P. and West, M. (1997). Bayesian forecasting of multinomial time series through conditionally Gaussian dynamic models. *J. Amer. Statist. Assoc.* **92**, 640-647.

Carter, C. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika* **81**, 541–553.

Christiansen, C.L. and Morris, C.N. (1997). Hierarchical Poisson regression modeling. *J. Amer. Statist. Assoc.* **92**, 618-632.

de Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. *Biometrika* **82**, 339-50.

Escribano, A. and Peña, D. (1994). Cointegration and common factors. *J. Time Series Anal.* **15**, 577-586.

Gamerman, D. and Migon H.S. (1993). Dynamic hierarchical models. *J. Roy. Statist. Soc. B* **55**, 629-642.

Geweke, J.F. (1988). The secular and cyclical behavior of real GDP in 19 OECD countries, 1957-1983. *J. Bus. and Econ. Statist.* **6**, 479–486.

Geweke, J.F. (1989). The posterior distribution of roots in multivariate autoregressions. *Proceedings of the ASA Section on Business and Economic Statistics.*

Geweke J. and Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Rev. Financial Studies* **9**, 557-587.

Harvey A.C., Ruiz E. and Shephard N. (1994). Multivariate stochastic variance models. *Rev. Economic Studies* **61**, 247-264.

Huerta, G. (1998). Bayesian analysis of latent structure in time series models. *PhD thesis,* Duke University.

Huerta, G. and West, M. (1997a). Priors and component structures in autoregressive time series. *J. Roy. Statist. Soc. B* (to appear).

Huerta, G. and West, M. (1997b). Bayesian inference on periodicities and component spectral structure in time series. *J. Time Series Anal.* (to appear).

Jacquier, E., Polson N.G. and Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *J. Bus. and Econ. Statist.* **12**, 371-388.

Jacquier, E., Polson N.G. and Rossi, P. (1995). Models and priors for multivariate stochastic volatility. *Discussion Paper,* Graduate School of Business, University of Chicago.

Kim S., Shephard N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev. Economic Studies* (to appear).

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

Li, F. (1997). Time deformation models: Theory and practice. *PhD thesis,* Duke University.

Liu, J.S. and Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *J. Amer. Statist. Assoc.* (to appear).

Normand, S.T., Glickman, M.E. and Gatsonis, C.A. (1997). Statistics methods for profiling providers of medical care: Issues and applications. *J. Amer. Statist. Assoc.* **92**, 803-814.

Park, J. and Maasch, K. (1993). Plio-Pleistocene time evolution of the 100-kyr cycle in marine paleoclimate records. *J. Geophys. Res.* **98**, 447–461.

Peña, D. and Box, G.E.P. (1987). Identifying a simplifying structure in time series. *J. Amer. Statist. Assoc.* **82**, 836-843.

Pitt, M.K. and Shephard, N. (1997). Filtering via simulation based on auxiliary particle filters. *Unpublished paper,* Nuffield College, Oxford.

Pitt, M.K. and Shephard, N. (1998). Time varying covariances: A factor stochastic volatility approach. (with discussion). *Bayesian Statistics 6* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: University Press.

Prado, R. (1998). Latent structure in non-stationary time series. *PhD thesis*, Duke University.

Prado, R. and West, M. (1997). Exploratory modelling of multiple non-stationary time series: Latent process structure and decompositions. *Modelling Longitudinal and Spatially Correlated Data* (T. Gregoire, ed.). New York: Springer-Verlag.

Priestley, M.B. (1980). State-dependent models: A general approach to non-linear time series analysis. *J. Time Series Anal.* **1**, 47-71.

Putnam, B.H. and Quintana, J. M. (1994). New Bayesian statistical approaches to estimating and evaluating models of exchange rates determination. *Proceedings of the ASA Section on Bayesian Statistical Science.*

Quintana, J.M. (1992). Optimal portfolios of forward currency contracts. *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: University Press.

Quintana, J.M. and West, M. (1987). An analysis of international exchange rates using multivariate DLMS. *The Statistician* **36**, 275-281.

Quintana, J. M. and Putnam, B. H. (1996). Debating currency markets efficiency using multiple-factor models. *Proceedings of the ASA Section on Bayesian Statistical Science.*

Tiao, G.C. and Tsay, R.S. (1989). Model specification in multivariate time series (with discussion). *J. Roy. Statist. Soc. B* **51**, 195-213.

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

Trenberth, K.E. and Hoar, T.J. (1996). The 1990-1995 El Niño-Southern oscillation event: Longest on record. *Geophys. Res. Letters* **23**, 57-60.

Weiner, R.D. and Krystal, A.D. (1993). EEG monitoring of ECT seizures. *The Clinical Science of Electroconvulsive Therapy*, American Psychiatric Press, Inc., 93-109.

West, M. (1985). Generalised linear models: scale parameters, outlier accommodation and prior distributions (with discussion). *Bayesian Statistics 2* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.), Amsterdam: North-Holland.

West, M., Harrison, P.J. and Migon, H.S. (1985). Dynamic generalised linear models and Bayesian forecasting (with discussion). *J. Amer. Statist. Assoc.* **80**, 73-97.

West, M. (1993). Mixture models, Monte Carlo, Bayesian updating and dynamic models. *Computing Science and Statistics* **24**, 325-333.

West, M. (1995). Bayesian inference in cyclical component dynamic linear models. *J. Amer. Statist. Assoc.* **90**, 1301-1312.

West, M. (1996). Some statistical issues in Palaeoclimatology (with discussion). *Bayesian Statistics 5* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: University Press.

West, M. (1997a). Bayesian time series: Models and computations for the analysis of time series in the physical sciences. *Maximum Entropy and Bayesian Methods 15* (K. Hanson and R. Silver, eds.), Kluwer.

West, M. (1997b). Modelling and robustness issues in Bayesian time series analysis (with discussion). *Bayesian Robustness 2* (J. Berger, F. Ruggeri and L. Wasserman, eds.), Hayward, CA: IMS Monographs.

West, M. (1997c). Time series decomposition. *Biometrika* **84**, 489-494.

West M. and Aguilar O. (1997). Studies of quality monitor time series: The V.A. hospital system. Report for the VA Management Science Group, Bedford, MA., and *ISDS Discussion Paper 97-22a*, Duke University.

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

West, M., Prado, R. and Krystal, A. (1997). Latent structure in non-stationary time series with application in studies of EEG traces. *ISDS Discussion Paper 97-14*, Duke University.