Bayesian Time Series Analysis

Mark Steel, University of Warwick*

Abstract

This article describes the use of Bayesian methods in the statistical analysis of time series. The use of Markov chain Monte Carlo methods has made even the more complex time series models amenable to Bayesian analysis. Models discussed in some detail are ARIMA models and their fractionally integrated counterparts, state-space models, Markov switching and mixture models, and models allowing for time-varying volatility. A final section reviews some recent approaches to nonparametric Bayesian modelling of time series.

1 Bayesian methods

The importance of Bayesian methods in econometrics has increased rapidly over the last decade. This is, no doubt, fuelled by an increasing appreciation of the advantages that Bayesian inference entails. In particular, it provides us with a formal way to incorporate the prior information we often possess before seeing the data, it fits perfectly with sequential learning and decision making and it directly leads to exact small sample results. In addition, the Bayesian paradigm is particularly natural for prediction, taking into account all parameter or even model uncertainty. The predictive distribution is the sampling distribution where the parameters are integrated out with the posterior distribution and is exactly what we need for forecasting, often a key goal of time-series analysis.

Usually, the choice of a particular econometric model is not prespecified by theory and many competing models can be entertained. Comparing models can be done formally in a Bayesian framework through so-called posterior odds, which is the product of the prior odds and the Bayes factor. The Bayes factor between any two models is the ratio of the likelihoods integrated out with the corresponding prior and summarizes how the data favour one model over another. Given a set of possible models, this immediately leads to posterior model probabilities. Rather than choosing a single model, a natural way to deal with model uncertainty is to use the posterior model probabilities to average out the inference

*The New Palgrave Dictionary of Economics, Palgrave Macmillan, reproduced with permission of Palgrave Macmillan. This article is taken from the author's original manuscript and has not been edited. The definitive published version of this extract may be found in the complete New Palgrave Dictionary of Economics in print and online, available at http://www.dictionaryofeconomics.com.

(on observables or parameters) corresponding to each of the separate models. This is called Bayesian model averaging. The latter was already mentioned in Leamer (1978) and recently applied to economic problems in *e.g.* Fernández *et al.* (2001) (growth regressions) and in Garratt *et al.* (2003) and Jacobson and Karlsson (2004) for macroeconomic forecasting.

An inevitable prerequisite for using the Bayesian paradigm is the specification of prior distributions for all quantities in the model that are treated as unknown. This has been the source of some debate, a prime example of which is given by the controversy over the choice of prior on the coefficients of simple autoregressive models. The issue of testing for a unit root (deciding whether to difference the series before modelling it through a stationary model) is subject to many difficulties from a sampling-theoretical perspective. Comparing models in terms of posterior odds provides a very natural Bayesian approach to testing, which does not rely on asymptotics or approximations. It is, of course, sensitive to how the competing models are defined (*e.g.* do we contrast the stationary model with a pure unit root model or a model with a root larger than or equal to one?) and to the choice of prior. The latter issues have lead to some controversy in the literature, and prompted a special issue of *The Journal of Applied Econometrics* with animated discussion around the paper by Phillips (1991). The latter paper advocated the use of Jeffreys' principles to represent prior ignorance about the parameters (see also the discussion in Chapter 6 of Bauwens *et al.*, 1999).

Like the choice between competing models, forecasting can also be critically influenced by the prior. In fact, prediction is often much more sensitive than parameter inference to the choice of priors (especially on autoregressive coefficients) and Koop *et al.* (1995) show that imposing stationarity through the prior on the autoregressive coefficient in a simple AR(1) model need not lead to stabilization of the predictive variance as the forecast horizon increases.

2 Computational algorithms

Partly, the increased use of Bayesian methods in econometrics is a consequence of the availability of very efficient and flexible algorithms for conducting inference through simulation in combination with ever more powerful computing facilities, which have made the Bayesian analysis of non-standard problems an almost routine activity. Particularly, Markov chain Monte Carlo (MCMC) methods have opened up a very useful class of computational algorithms and have created a veritable revolution in the implementation of Bayesian methods. Whereas Bayesian inference before 1990 was at best a difficult undertaking in practice, reserved for a small number of specialized researchers and limited to a rather restricted set of models, it has now become a very accessible procedure which can fairly easily be applied to almost any model. The main idea of MCMC methods is that inference about an analytically intractable posterior (often in high dimensions) is conducted through generating a Markov chain which converges to a chain of drawings from the posterior distribution. Of course, predictive inference is also immediately available once one has such a chain of drawings. Various ways of constructing such a Markov chain exist, depending of the structure of the problem. The most commonly used are the Gibbs sampler and the Metropolis Hastings sampler. The use of data augmentation (*i.e.* adding auxiliary variables to the sampler) can facilitate implementation of the MCMC sampler, so that often the analysis is conducted on an augmented space including not only the model parameters but also things like latent variables and missing observations. An accessible reference to MCMC methods is *e.g.* Gamerman (1997).

As a consequence, we are now able to conduct Bayesian analysis of time series models that have been around for a long time (such as ARMA models) but also of more recent additions to our catalogue of models, such as Markov switching and nonparametric models, and the literature is vast. Therefore, I will have to be selective and will try to highlight a few areas which I think are of particular interest. I hope this can give an idea of the role that Bayesian methods can play in modern time series analysis.

3 ARIMA and ARFIMA models

Many models used in practice are of the simple ARIMA type, which have a long history and were formalised in Box and Jenkins (1970). ARIMA stands for Autoregressive Integrated Moving Average and an ARIMA(p, d, q) model for an observed series $\{y_t\}$, $t = 1, \ldots, T$ is a model where the dth difference $z_t = y_t - y_{t-d}$ is taken to induce stationarity of the series. The process $\{z_t\}$ is then modelled as $z_t = \mu + \varepsilon_t$ with

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \cdots + \phi_p \varepsilon_{t-p} + u_t - \theta_1 u_{t-1} - \cdots - \theta_q u_{t-q}$$

or in terms of polynomials in the lag operator L (defined through $L^s x_t = x_{t-s}$):

$$\phi(L)\varepsilon_t = \theta(L) u_t$$

where $\{u_t\}$ is white noise and usually Normally distributed as $u_t \sim N(0, \sigma^2)$. The stationarity and invertibility conditions are simply that the roots of $\phi(L)$ and $\theta(L)$, respectively, are outside the unit circle. An accessible and extensive treatment of the use of Bayesian methods for ARIMA models can be found in Bauwens *et al.* (1999). The latter book also has a useful discussion of multivariate modelling using Vector Autoregressive (VAR) models and cointegration.

The MCMC samplers used for inference in these models typically use data augmentation. Marriott *et al.* (1996) use a direct conditional likelihood evaluation and augment with unobserved data and errors to conduct inference on the parameters (and the augmented vectors $\varepsilon_a = (\varepsilon_0, \varepsilon_{-1}, \ldots, \varepsilon_{1-p})'$ and $u_a = (u_0, u_{-1}, \ldots, u_{1-q})$). A slightly different approach is followed by Chib and Greenberg (1994), who consider a state-space representation and use MCMC on the parameters augmented with the initial state vector.

ARIMA models will either display perfect memory (if there are any unit roots) or quite short memory with geometrically decaying autocorrelations (in the case of a stationary ARMA model). ARFIMA (Autoregressive Fractionally Integrated Moving Average) models (see Granger and Joyeux, 1980) have more flexible memory properties, due to fractional integration which allows for hyperbolic decay.

Consider $z_t = \Delta y_t - \mu$, which is modelled by an ARFIMA(p, δ, q) model as:

$$\phi(L)(1 - L)^{\delta} z_t = \theta(L) u_t,$$

where $\{u_t\}$ is white noise with $u_t \sim N(0, \sigma^2)$, and $\delta \in (-1, 0.5)$. The fractional differencing operator $(1-L)^\delta$ is defined as

$$(1-L)^\delta = \sum_{j=0}^{\infty} c_j(\delta) L^j,$$

where $c_0(\cdot) = 1$ and for $j > 0$:

$$c_j(a) = \prod_{k=1}^{j} \left(1 - \frac{1+a}{k}\right).$$

This model takes the entire past of z_t into account, and has as a special case the ARIMA$(p, 1, q)$ for y_t (for $\delta = 0$). If $\delta > -1$, z_t is invertible (Odaki, 1993) and for $\delta < 0.5$ we have stationarity of z_t. Thus, we have three regimes:

$\delta \in (-1, -0.5)$: y_t trend-stationary with long memory
$\delta \in (-0.5, 0)$: z_t stationary with intermediate memory
$\delta \in (0, 0.5)$: z_t stationary with long memory.

Of particular interest is the Impulse Response Function $I(n)$, which captures the effect of a shock of size one at time t on y_{t+n}, and is given by

$$I(n) = \sum_{i=0}^{n} c_i(-\delta - 1) J(n-i),$$

with $J(i)$ the standard ARMA(p, q) impulse responses (*i.e.* the coefficients of $\phi^{-1}(L)\theta(L)$). Thus, $I(\infty)$ is 0 for $\delta < 0$, $\theta(1)/\phi(1)$ for $\delta = 0$ and ∞ for $\delta > 0$. Koop *et al.* (1997) analyse the behaviour of the impulse response function for real U.S. GNP data using a set of 32 possible models containing both ARMA and ARFIMA models for z_t. They use Bayesian model averaging to conduct predictive inference and inference on the impulse responses, finding about one third of the posterior model probability concentrated on the ARFIMA models. Koop *et al.* (1997) use importance sampling to conduct inference on the parameters, while MCMC methods are used in Pai and Ravishanker (1996) and Hsu and Breidt (2003).

4 State-Space Models

The basic idea of such models is that an obervable y_t is generated by an observation or measurement equation

$$y_t = F_t' \theta_t + v_t,$$

where $v_t \sim N(0, V_t)$, and is expressed in terms of an unobservable state vector θ_t (capturing *e.g.* levels, trends or seasonal effects) which is itself dynamically modelled through a system or transition equation

$$\theta_t = G_t \theta_{t-1} + w_t,$$

with $w_t \sim N(0, W_t)$ and all error terms $\{v_t\}$ and $\{w_t\}$ are mutually independent. Normality is typically assumed, but is not necessary and a prior distribution is required to describe the initial state vector θ_0.

Models are defined by the (potentially time-varying) quadruplets $\{F_t, G_t, V_t, W_t\}$ and the time-varying states θ_t make them naturally adaptive to changing circumstances. This feature also fits very naturally with Bayesian methods, which easily allow for sequential updating. These models are quite general and include as special cases *e.g.* ARMA models, as well as stochastic volatility models, used in finance (see later in this entry).

There is a relatively long tradition of state-space models in econometrics and a textbook treatment can already be found in Harvey (1981). Bayesian methods for such models were discussed in *e.g.* Harrison and Stevens (1976) and a very extensive treatment is provided in West and Harrison (1997), using the terminology "Dynamic Linear Models". An accessible introduction to Bayesian analysis with these models can be found in Koop (2003, Ch. 8).

On-line sequential estimation and forecasting with the simple Normal state-space model above can be achieved with Kalman filter recursions, but more sophisticated models (or estimation of some aspects of the model besides the states) usually require numerical methods for inference. In that case, the main challenge is typically the simulation of the sequence of unknown state vectors. Single-state samplers (updating one state vector at a time) are generally less efficient than multi-state samplers, where all the states are updated jointly in one step. Efficient algorithms for multi-state MCMC sampling schemes have been proposed by Carter and Kohn (1994) and de Jong and Shephard (1995). For fundamentally non-Gaussian models, the methods in Shephard and Pitt (1997) can be used. A recent contribution of Harvey *et al.* (2005) uses Bayesian methods for state space models with trend and cyclical components, exploiting informative prior notions regarding the length of economic cycles.

5 Markov switching and mixture models

Markov switching models were introduced by Hamilton (1989) and essentially rely on an unobserved regime indicator s_t, which is assumed to behave as a discrete Markov chain with, say, K different levels. Given $s_t = i$ the observable y_t will be generated by a time series model which corresponds to regime i, where $i = 1, \ldots, K$. These models are often stationary ARMA models, and the switching between regimes will allow for some non-stationarity, given the regime allocations. Such models are generally known as hidden Markov models in the statistical literature.

Bayesian analysis of these models is very natural, as that methodology provides an immediate framework for dealing with the latent states, $\{s_t\}$, and a simple MCMC framework for inference on both the model parameters and the states was proposed in Albert and Chib (1993). A bivariate version of the Hamilton model was analysed in Paap and van Dijk (2003), who also examine the cointegration relations between the series modelled and find evidence for cointegration between U.S. per capita income and consumption. Using a similar model, Smith and Summers (2005) examine the synchronisation of business cycles across countries and find strong evidence in favour of the multivariate Markov switching model over a linear VAR model.

When panel data are available, another relevant question is whether one can find clusters of entities (such as countries or regions) which behave similarly, while allowing for differences between the clusters. This issue is addressed from a fully Bayesian perspective in Frühwirth-Schnatter and Kaufmann (2006), where model-based clustering (across countries) is integrated with a Markov switching framework (over time). This is achieved by a finite mixture of Markov switching autoregressive models, where the number of elements in the mixture corresponds to the number of clusters and is treated as an unknown parameter. Frühwirth-Schnatter and Kaufmann (2006) analyse a panel of growth rates of industrial production in 21 countries and distinguish two clusters with different business cycles. This also feeds into the important debate on the existence of so-called "convergence clubs" in terms of income per capita as discussed in Durlauf and Johnson (1995) and Canova (2004).

Another popular way of inducing nonlinearities in time series models is through, so-called, threshold autoregressive models, where the choice of regimes is not governed by an underlying Markov chain, but depends on previous values of the observables. Bayesian analyses of such models can be found in *e.g.* Geweke and Terui (1993) and are extensively reviewed in Ch. 8 of Bauwens *et al.* (1999). The use of Bayes factors to choose between various nonlinear models, such as threshold autoregressive and Markov switching models was discussed in Koop and Potter (1999).

Geweke and Keane (2005) present a general framework for Bayesian mixture models where the state probabilities can depend on observed covariates. They investigate increasing the number of components in the mixture, as well as the flexibility of the components and the specification of the mechanism for the state probabilities, and find their mixture model approach compares well with ARCH-type models (as described in the next section) in the context of stock return data.

6 Models for time-varying volatility

The use of conditional heteroskedasticity initially introduced in the ARCH (autoregressive conditional heteroskedasticity) model of Engle (1982) has been extremely successful in modelling financial time series, such as stock prices, interest rates and exchange rates. The ARCH model was generalised to GARCH (generalised ARCH) by Bollerslev (1986). A simple version of the GARCH model for an observable series $\{y_t\}$, given its past which is denoted by I_{t-1}, is the following:

$$y_t = u_t\sqrt{h_t} \tag{1}$$

where $\{u_t\}$ is white noise with mean zero and variance one. The conditional variance of y_t given I_{t-1} is then h_t, which is modelled as

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_i h_{t-i} \tag{2}$$

where all parameters are positive and usually $p = q = 1$ is sufficient in practical applications. Bayesian inference for such models was conducted through importance sampling in Kleibergen and van Dijk (1993) and using MCMC methods in Bauwens and Lubrano (1998).

An increasingly popular alternative model allows for the variance h_t to be determined by its own stochastic process. This is the so-called stochastic volatility model, which in its basic form replaces (2) by the assumption that the logarithm of the conditional volatility is driven by its own AR(1) process

$$\ln(h_t) = \alpha + \delta \ln(h_{t-1}) + v_t,$$

where $\{v_t\}$ is a white noise process independent of $\{u_t\}$ in (1). Inference in such models requires dealing with the latent volatilities, which are incidental parameters and have to be integrated out in order to evaluate the likelihood. MCMC sampling of the model parameters and the volatilities jointly is a natural way of handling this. An MCMC sampler where each volatility was treated in a separate step was introduced in Jacquier *et al.* (1994) and efficient algorithms for multi-state MCMC sampling schemes were suggested by Carter and Kohn (1994) and de Jong and Shephard (1995). Many extensions of the simple stochastic volatility model above have been proposed in the literature, such as correlations between the $\{u_t\}$ and $\{v_t\}$ processes, capturing leverage effects, or fat-tailed distributions for u_t. Inference with these more general models and ways of choosing between them are discussed in Jacquier *et al.* (2004).

Recently, the focus in finance has shifted more towards continuous-time models, and continuous-time versions of stochastic volatility models have been proposed. In particular, Barndorff-Nielsen and Shephard (2001) introduce a class of models where the volatility behaves according to an Ornstein-Uhlenbeck process, driven by a positive Lévy process without Gaussian component (a pure jump process). These models introduce discontinuities (jumps) into the volatility process. The latter paper also considers superpositions of such processes. Bayesian inference in such models through MCMC methods is complicated by the fact that the model parameters and the latent volatility process are often highly correlated in the posterior, leading to the problem of overconditioning. Griffin and Steel (2006b) propose MCMC methods based on a series representation of Lévy processes, and avoid overconditioning by dependent thinning methods. In addition, they extend the model by including a jump component in the returns, leverage effects and separate risk pricing for the various volatility components in the superposition. An application to stock price data shows substantial empirical support for a superposition of processes with different risk premiums and a leverage effect. A different approach to inference in such models is proposed in Roberts *et al.* (2004), who suggest a reparameterisation to reduce the correlation between the data and the process. The reparameterised process is then proposed only in accordance with the parameters.

7 Semi- and Nonparametric Models

The development and use of Bayesian nonparametric methods has been a rapidly growing topic in the statistics literature, some of which is reviewed in Müller and Quintana (2004). However, the latter review does not include applications to time series, which have been perhaps less prevalent than applications in other areas, such as regression, survival analysis and spatial statistics.

Bayesian nonparametrics is sometimes considered an oxymoron, since Bayesian methods are inherently likelihood-based, and, thus, require a complete probabilistic specification of the model. However, what is usually called Bayesian nonparametrics corresponds to models with priors defined over infinitely-dimensional parameter spaces (functional spaces) and this allows for very flexible procedures, where the data are allowed to influence virtually all features of the model.

Defining priors over collections of distribution functions requires the use of random probability measures. The most popular of these is the so-called Dirichlet process prior introduced by Ferguson (1973). This is defined for a space Θ and a σ-field $\mathcal{B}$ of subsets of Θ. The process is parameterised in terms of a probability measure H on $(\Theta, \mathcal{B})$ and a positive scalar M. A random probability measure, F, on $(\Theta, \mathcal{B})$ follows a Dirichlet process DP(MH) if, for any finite measurable partition, $B_1, \ldots, B_k$, the vector $(F(B_1), \ldots, F(B_k))$ follows a Dirichlet distribution with parameters $(MH(B_1), \ldots, MH(B_k))$. The distribution H centres the process and M can be interpreted as a precision parameter.

The Dirichlet process is (almost surely) discrete and, thus, not always suitable for modelling observables directly. It is, however, often incorporated into semiparametric models using the hierarchical framework

$$y_i \sim g(y_i|u_i) \quad \text{with} \quad u_i \sim F \quad \text{and} \quad F \sim \text{DP}(MH), \tag{3}$$

where $g(\cdot)$ is a probability density function. This model is usually referred to as a Mixture of Dirichlet Processes. The marginal distribution for y_i is a mixture of the distribution characterized by $g(\cdot)$. This basic model can be extended: the density $g(\cdot)$ or the centring distribution H can be (further) parameterised and inference can be made about these parameters. In addition, inference can be made about the mass parameter M. Inference in these models using MCMC algorithms has become quite feasible, using methods based on MacEachern (1994) and Escobar and West (1995).

However, the model in (3) assumes independent and identically distributed observations and is, thus, not directly of interest for time series modelling. A simple approach followed by Hirano (2002) is to use (3) for modelling the errors of an autoregressive model specification. However, this does not allow for the distribution to change over time. Making the random probability measure F itself depend on lagged values of the variable under consideration y_t (or, generally, any covariates) is not a straightforward extension. Müller *et al.* (1997) propose a solution by modelling y_t and y_{t-1} jointly, using a Mixture of Dirichlet Processes. The main problem with this approach is that the resulting model is not really a conditional model for y_t given y_{t-1}, but incorporates a contribution from the marginal model for y_{t-1}. Starting from the stick-breaking representation of a Dirichlet process, Griffin and Steel (2006a) introduce the class of Order-based Dependent Dirichlet Processes, where the weights in the stick-breaking representation induce dependence between distributions that correspond to similar values of the covariates (such as time). This class induces a Dirichlet process at each covariate value, but allows for dependence. Similar weights are associated with similar orderings of the elements in the representation and these orderings are derived from a point process in such a way that distributions that are close in covariate space will tend to be highly correlated. One proposed construction (the arrivals ordering) is particularly suitable for time series and is applied to stock index returns, where the volatility is modelled through an order-based dependent Dirichlet process. Results illustrate the flexibility and the feasibility of this approach. Jensen (2004) uses a Dirichlet process prior on the wavelet representation of the observables to conduct Bayesian inference in a stochastic volatility model with long memory.

8 Conclusion: where are we heading?

In conclusion, Bayesian analysis of time series models is alive and well. In fact, it is an ever growing field, and we are now starting to explore the advantages that can be gained from using Bayesian methods on time series data. Bayesian counterparts to the classical analysis of existing models, such as AR(F)IMA models are by now well-developed and a lot of work has already been done there to make Bayesian inference in these models a fairly routine activity. The main challenge ahead for methodological research in this field is perhaps to further develop really novel models that constitute not merely a change of inferential paradigm, but are inspired by the new and exciting modelling possibilities that are available through the combination of Bayesian methods and MCMC computational algorithms. In particular, nonparametric Bayesian time-series modelling falls in that category and I expect that more research in this area will be especially helpful in increasing our understanding of time series data.

References

Albert, J.H. and Chib, S. 1993. Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. Journal of Business and Economic Statistics 11, 1-15.

Barndorff-Nielsen, O.E. and Shephard, N. 2001. Non-Gaussian OU based models and some of their uses in financial economics. Journal of the Royal Statistical Society Series B 63, 167-241 (with discussion).

Bauwens, L. and Lubrano, M. 1998. Bayesian inference on GARCH models using the Gibbs sampler. The Econometrics Journal 1, C23-C46.

Bauwens, L., Lubrano, M. and Richard, J.F. 1999. Bayesian Inference in Dynamic Econometric Models. Oxford: Oxford University Press.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327.

Box, G.E.P. and Jenkins, G.M. 1970. Time Series Analysis: Forecasting and Control. San Francisco: Holden Day.

Canova, F. 2004. Testing for convergence clubs in income per capita: A predictive density approach. International Economic Review 45, 49-77.

Carter, C.K. and Kohn, R. 1994. On Gibbs sampling for state space models. Biometrika 81, 541-553.

Chib, S. and Greenberg, E. 1994. Bayes inference in regression models with ARMA(p, q) errors. Journal of Econometrics 64, 183-206.

de Jong, P. and Shephard, N. 1995. The simulation smoother for time series models. Biometrika 82, 339-350.

Durlauf, S. and Johnson, P. 1995. Multiple regimes and cross-country growth behaviour. Journal of Applied Econometrics 10, 365-384.

Engle, R.F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987-1008.

Escobar, M. D. and West, M. 1995. Bayesian density-estimation and inference using mixtures. Journal of the American Statistical Association 90, 577-588.

Ferguson, T.S. 1973. A Bayesian analysis of some nonparametric problems. Annals of Statistics 1, 209-230.

Fernández, C., Ley, E. and Steel, M. F. J. 2001. Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics 16, 563-576.

Frühwirth-Schnatter, S. and Kaufmann, S. 2006. Model-based clustering of multiple time series. Journal of Business and Economic Statistics, forthcoming.

Gamerman, D. 1997. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Boca Raton: Chapman and Hall/CRC.

Garratt, A., Lee, K., Pesaran, H.M. and Shin, Y. 2003. Forecast uncertainties in macroeconometric modelling: An application to the UK economy. Journal of the American Statistical Association 98, 829-838.

Geweke, J. and Terui, N. 1993. Bayesian threshold autoregressive models for nonlinear time series. Journal of Time Series Analysis 14, 441-454.

Geweke, J. and Keane, M. 2005. Smoothly mixing regressions. mimeo, University of Iowa.

Granger, C.W.J. and Joyeux, R. 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15-39.

Griffin, J.E. and Steel, M.F.J. 2006a. Order-based dependent Dirichlet processes. Journal of the American Statistical Association 101, 179-194.

Griffin, J.E. and Steel, M.F.J. 2006b. Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. Journal of Econometrics, forthcoming.

Hamilton, J.D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357-384.

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

Harvey, A.C. 1981. Time Series Models. Oxford: Philip Allen.

Harvey, A.C., Trimbur, T.M. and van Dijk, H.K. 2005. Trends and cycles in economic time series: A Bayesian approach. mimeo, University of Cambridge, UK.

Hirano, K., 2002. Semiparametric Bayesian inference in autoregressive panel data models. Econometrica 70, 781-799.

Hsu, N.J. and Breidt, F.J. 2003. Bayesian analysis of fractionally integrated ARMA with additive noise. Journal of Forecasting 22, 491-514.

Jacobson, T. and Karlsson, S. 2004. Finding good predictors for inflation: A Bayesian model averaging approach. Journal of Forecasting 23, 479-496.

Jacquier, E., Polson, N.G. and Rossi, P.E. 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371-417 (with discussion).

Jacquier, E., Polson, N.G. and Rossi, P.E. 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122, 185-212.

Kleibergen, F. and van Dijk, H.K. 1993. Non-stationarity in GARCH models: A Bayesian analysis. Journal of Applied Econometrics 8, S41-S61.

Koop, G. 2003. Bayesian Econometrics. Chichester: Wiley.

Koop, G, Ley, E., Osiewalski J. and Steel, M.F.J. 1997. Bayesian analysis of long memory and persistence using ARFIMA models. Journal of Econometrics 76, 149-169.

Koop, G, Osiewalski J. and Steel, M.F.J. 1995. Bayesian long-run prediction in time series models. Journal of Econometrics 69, 61-80.

Koop, G. and Potter, S.M. 1999. Bayes factors and nonlinearity: Evidence from economic time series. Journal of Econometrics 88, 251-281.

Leamer, E.E. 1978. Specification Searches: Ad Hoc Inference with Nonexperimental Data. New York: Wiley.

MacEachern, S.N. 1994. Estimating Normal means with a conjugate style Dirichlet process prior. Communications in Statistics, B 23, 727-741.

Marriott, J., Ravishanker, N., Gelfand A. and Pai, J. 1996. Bayesian analysis of ARMA processes: Complete sampling-based inference under exact likelihoods, in *Bayesian Analysis in Statistics and Econometrics*, eds. D. Berry, K. Chaloner and J. Geweke. New York: Wiley, pp. 243-256.

Müller, P. and Quintana, F.A. 2004. Nonparametric Bayesian data analysis. Statistical Science 19, 95-110.

Müller, P. , West. M. and MacEachern, S.N. 1997. Bayesian models for nonlinear autoregressions. Journal of Time Series Analysis 18, 593-614.

Odaki, M. 1993. On the invertibility of fractionally differenced ARIMA processes. Biometrika 80, 703-709.

Paap, R. and van Dijk, H.K. 2003. Bayes estimation of Markov trends in possibly cointegrated series: An application to U.S. consumption and income. Journal of Business and Economic Statistics 21, 547-563.

Pai, J. and Ravishanker, N. 1996. Bayesian modeling of ARFIMA processes by Markov chain Monte Carlo methods. Journal of Forecasting 16, 63-82.

Phillips, P.C.B. 1991. To criticize the critics: An objective Bayesian analysis of stochastic trends. Journal of Applied Econometrics 6, 333-473 (with discussion).

Roberts, G.O., Papaspiliopoulos, O. and Dellaportas, P. 2004. Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. Journal of the Royal Statistical Society Series B 66, 369-393.

Shephard, N. and Pitt, M.K. 1997. Likelihood analysis of non-Gaussian measurement time series. Biometrika 84, 653-667.

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

Suggested cross-references: ARCH models; Bayesian econometrics; Bayesian methods in macroeconometrics; Bayesian nonparametrics; Bayesian statistics; Econometrics; Financial econometrics; Markov chain Monte Carlo methods; Long memory models; State space models; Statistics and economics; Stochastic volatility models; Time series analysis.

Mark F.J. Steel