Monograph:

Guo, L.Z., Billings, S.A., Coca, D. et al. (1 more author) (2009) Multiscale time series modelling with an application to the relativistic electron intensity at the geosynchronous orbit. Research Report. ACSE Research Report no. 999 . Automatic Control and Systems Engineering, University of Sheffield

Multiscale time series modelling with an application to the relativistic electron intensity at the geosynchronous orbit

Guo, L. Z., Billings, S. A., Coca, D., and Balikhin, M.

Department of Automatic Control and Systems Engineering
University of Sheffield
Sheffield, S1 3JD
UK

Research Report No. 999
September 2009

Multiscale time series modelling with an application to the relativistic electron intensity at the geosynchronous orbit

L. Z. Guo, S. A. Billings, D. Coca, and M. Balikhin

Department of Automatic Control and Systems Engineering
University of Sheffield
Sheffield S1 3JD, UK

Abstract

In this paper, a Bayesian system identification approach to multiscale time series modelling is proposed, where multiscale means that the output of the system is observed at one (coarse) resolution while the input of the system is observed at another (fine) resolution. The proposed method identifies linear models at different levels of resolution where the link between the two resolutions is realised via non-overlapping averaging process. This averaged time series at the coarse level of resolution is assumed to be a set of observations from an implied process so that the implied process and the output of the system result in an errors-in-variables ARMAX model at the coarse level of resolution. By using a Bayesian inference and Markov Chain Monte Carlo (MCMC) method, such a modelling framework results in different dynamical models at different levels of resolution at the same time. The new method is also shown to have the ability to combine information across different levels of resolution. An application to the analysis of the relativistic electron intensity at the geosynchronous orbit is used to illustrate the new method.

1 Introduction

In time series modelling and analysis, it is sometimes worthwhile to investigate the underlying processes from different levels of resolution. This need may arise from the fact that in some cases the measurements can only be made at different levels of resolution (e.g. hourly and daily) due to a variety of technical limitations. The data obtained as such are of a multiple nature. Another reason may come from the following consideration. Whilst investigating the data at a single level of resolution may often be satisfactory, such an approach may not be able to capture all the features of the underlying processes and some important structures of the signals may exist at different levels of resolution. It follows that the levels at which the observations and inferences are made may lead to different conclusions with respect to the dynamics of the underlying processes. These considerations resulted in a series of studies on multiscale time series modelling methods including temporal aggregation (Drost and Nijman 1993), reconstruction of time series based on higher order multiscale statistics (Nawrotha and Peinke 2006), multiscale and hidden resolution time series models (Ferreira and Lee 2007), and Bayesian methods (Oigard, Rue, and Godtliebsen 2006). In this paper, a new Bayesian system identification approach is proposed to tackle the multiscale data modelling problem, which can be considered as an extension of the multiscale and hidden resolution time series model given by Ferreira and Lee (2007).

The multiscale problem studied here is the case where the output and input of a system are measured at different levels of resolution; where the output is at a coarse temporal scale while the input is at a fine temporal scale. It is assumed that the input at the fine scale is subjected to a dynamical process while at the coarse level the output follows an ARMAX model. The actual input at the coarse level is subjected to an implied dynamical model. The link across the two resolutions is realised via a non-overlapping averaging process. The averaged values are considered as observations from the implied input process at the coarse level. In this way, the implied input process and the output of the system result in an errors-in-variables ARMAX model at the coarse level of resolution. The objective of the paper is to estimate the parameters of the models at different levels of resolution and to produce predictions for the whole system. By using a Bayesian inference and Markov Chain Monte Carlo (MCMC) method, such a modelling framework results in different dynamical models at different levels of resolution at the same time. Jeffrey's rule of conditioning is used to ensure consistent modelling of the time series across the different levels of resolution. The method has the ability to combine information across different levels of resolution and to produce integrated modelling and prediction.

One of the motivations of the study, arises from an application related to the analysis of the relativistic electron intensity at a geosynchronous orbit. The real data were obtained from National Geophysical Data Centre, USA, where the relativistic electron flux was measured daily and the solar wind velocity and the geomagnetic index $SymH$ were measured minutely/hourly. Relativistic electrons in the Earth's inner magnetosphere are of great importance from both scientific and practical standpoints and the flux in the outer radiation belt exhibits a highly dynamic behaviour, which has been identified to be most relevant to the solar wind velocity and $SymH$ index, and is therefore an important influence on space weather. Therefore, in this paper an ARMAX model of the relativistic electron intensity, with the flux as output, and the solar wind velocity and the $SymH$ index as input, is analysed by using the new algorithm.

The paper begins in section 2 with a formulation of the problem, followed by a detailed discussion of the likelihood function and posterior probability densities. Section 3 presents the inference and prediction problem. Section 4 illustrates the proposed approach using the example mentioned earlier. Finally conclusions are drawn in section 5.

2 Problem description

2.1 Basic framework

In the section, the problem will be formulated within a general framework. The problem begins with two univariate time series (it is straightforward to extend the analysis to multivariate time series) observed at different temporal scales: $y_s, s = 1, 2, \cdots$ and $x_t, t = 1, 2, \cdots$, where y_s is assumed to be the output of a underlying dynamical system while x_t is assumed to be the input. To specify the difference between the temporal scales, it is also assumed that the time-indexes s and t are related together within a window of length $l > 0$, a positive integer. This means the output series y_s is l times coarser than the input series x_t, that is, in this time-indexing the sequence of inputs and outputs ordered in the way how they are observed is

$$x_1, x_2, \cdots, x_l, y_1, x_{l+1}, x_{l+2}, \cdots, x_{2l}, y_2, \cdots. \tag{1}$$

For any $n_y > 0$ and $n_x = l \times n_y$, denote $x_{1:n_x} = (x_1, x_2, \cdots, x_{n_x})^T$ and $y_{1:n_y} = (y_1, y_2, \cdots, y_{n_y})^T$ the observations up to time instant n_x. Although the problem can be investigated within a mixed temporal scale framework such as given by Nawrotha and Peinke (2006), in this paper it is studied under a general multiple scale modelling framework motivated by Ferreira, et al (2006). This multiple scale framework assumes that the dynamics at different scales are subjected to different models, between which the link is built via a link equation of the following form

$$u_s = u_s^o + \mu_s \tag{2}$$

in which μ_s is noise and the known variable $u_s^o = f(x_{(s-1)l+1}, x_{(s-1)l+2}, \cdots, x_{sl})$, where f can take many different forms such as maximising and averaging. Note that in this way the influence of the input x_t at the fine level on the output y_s at the coarse level is realised via the intermediate quantity u_s at the coarse level. It follows that the modelling problem of the pair (u_s, y_s), forming an input-output pair of the underlying dynamical process at the coarse level, is of interest to us. Apart from this coarse level model, the time series $x_t, t = 1, 2, \cdots$, is assumed to be subjected to a dynamical model which is of interest to us also. At the same time, it is interesting to investigate what impact it will have on both levels if $u_s, s = 1, 2, \cdots$, follows a dynamical model. The objective of the paper is to identify these three models from the multiscale observations y_s and x_t by using a Bayesian method.

Assume that the pair $y_s, u_s, s = 1, 2, \cdots$ follows an Errors-In-Variables (EIV) ARMAX model as follows

$$y_s = G(q^{-1})u_s + H(q^{-1})w_s = \frac{B(q^{-1})}{A(q^{-1})}u_s + \frac{C(q^{-1})}{A(q^{-1})}w_s \tag{3}$$

where $A(q^{-1}), B(q^{-1})$, and $C(q^{-1})$ are polynomials in the backshift operator q^{-1}

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_0 q^{-1} + b_1 q^{-2} + \cdots + b_{n_b} q^{-n_b} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c} \end{aligned} \tag{4}$$

and the error $w_s \overset{i.i.d}{\sim} N(0, \sigma_y^2)$, N denotes the normal distribution. y and u are the output and input of the process respectively. The fact that the input u_s comes from a process whose observations are the averaged values of the measurement from a finer level justifies this is an EIV problem. Assume also that time series $x_t, t = 1, 2, \cdots$ follows an AR(1) model as follows

$$x_t = \phi_x x_{t-1} + \varepsilon_t \tag{5}$$

where $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma_x^2)$. Thus for all $n_x > 0$, $p(x_{1:n_x})$ is an n_x-dimensional stationary distribution

$$p(x_{1:n_x}) = N(x_{1:n_x}|\mathbf{0}, V_x) \tag{6}$$

where $x_{1:n_x} = (x_1, \cdots, x_{n_x})^T$, and $\mathbf{0}$ is the zero-vector and $V_x = (\sigma_x^2 \phi_x^{|i-j|}/(1-\phi_x^2))_{ij}$ is an $n_x \times n_x$ covariance matrix.

The link (2) in this paper is assumed to be averaging. Let $u_s^o = f(x_{(s-1)l+1}, x_{(s-1)l+2}, \cdots, x_{sl}) = 1/l \sum_{i=1}^{l} x_{(s-1)*l+i}$ be the non-overlapping averaged value of l consecutive values of x over the window l and it is assumed that $\mu_s \overset{i.i.d}{\sim} N(0, \tau)$. Let $u_{1:n_u}^o = (u_1^o, \cdots, u_{n_u}^o)^T$. If u_s are conditionally independent with τ, the link between these two different scales is then given by the following relationship

$$p(u_{1:n_u}|u_{1:n_u}^o, x_{1:n_x}) = p(u_{1:n_u}|x_{1:n_x}) = N(u_{1:n_u}|Ax_{1:n_x}, \Gamma) \tag{7}$$

where $n_u \times l = n_x$, $u_{1:n_u} = (u_1, \cdots, u_{n_u})^T$, and $\Gamma = \tau \mathbf{I}$. Note that A is a sparse matrix whose non-zero elements are all $1/l$ of which the row i are those in columns $(i-1)l+1$, and $p(u_{1:n_u}|u^o_{1:n_u}) = p(u_{1:n_u}|x_{1:n_x})$.

Now it is assumed that the unobserved input variable u_s at a coarse level follows an implied AR(1) process

$$u_s = \phi_u u_{s-1} + \eta_s \tag{8}$$

where $\eta_s \overset{i.i.d}{\sim} N(0, \sigma_u^2)$. Thus for all $n_u > 0$, $q(u_{1:n_u})$ is the implied n_u-dimensional stationary distribution

$$q(u_{1:n_u}) = N(u_{1:n_u}|0, Q_u) \tag{9}$$

where $u_{1:n_u} = (u_1, \cdots, u_{n_u})^T$, and 0 is the zero-vector and $Q_u = (\sigma_u^2 \phi_u^{|i-j|}/(1-\phi_u^2))_{ij}$ is an $n_u \times n_u$ covariance matrix.

The following assumptions are set to make the problem well defined.

Assumption 1. The n_a, n_b, n_c values and the link matrix A are known (Model).

Assumption 2. $A(q^{-1})$, and $C(q^{-1})$ have no common roots.

Assumption 3. $A(q^{-1})$, and $C(q^{-1})$ are asymptotically stable polynomials (Stationarity and invertibility).

Assumption 4. Zero initial conditions for all the involved processes are assumed, which can be justified by the discussion given by Peterka (1981). A similar discussion about non-zero initial conditions can be made following Chib and Greenberg (1994) if necessary.

2.2 Implied processes induced by u_s

From (7), the marginal densities of $u_{1:n_u}$ can be calculated as

$$p(u_{1:n_u}) = \int p(u_{1:n_u}|x_{1:n_x}) \times p(x_{1:n_x}) dx_{1:n_x} = N(\mathbf{0}, A^T V_x A + \Gamma) \tag{10}$$

Obviously, due to the assumption of the implied process (8), the involved probability densities may not be compatible. According to Jeffrey's rule of conditioning, the revised probability distribution must not change the conditional probability degrees of any random variable given uncertain variable $u_{1:n_u}$, which means

$$
\begin{aligned}
q(x_{1:n_x}|u_{1:n_u}) = p(x_{1:n_x}|u_{1:n_u}) \\
q(y_{1:n_y}|u_{1:n_u}) = p(y_{1:n_y}|u_{1:n_u})
\end{aligned} \tag{11}
$$

The induced implied process at the fine level has been discussed in detail in Ferreira, et al (2006), which are shown in here as

$$
q(x_{1:n_x}) = \int p(x_{1:n_x}|u_{1:n_u}) \times q(u_{1:n_u})du_{1:n_u} = N(x_{1:n_x}|0, Q_x) \tag{12}
$$

with $Q_x = V_x - B(W - Q_u)B^T$ where $W = AV_xA^T + \Gamma$ and $B = V_xA^TW^{-1}$.

The induced implied process at the coarse level is given by the following formula

$$
q(y_{1:n_y}) = \int p(y_{1:n_y}|u_{1:n_u}) \times q(u_{1:n_u})du_{1:n_u} \tag{13}
$$

Under the assumptions that $H(q^{-1})$ is monic, w_s is independent, and $C(q^{-1})$ is asymptotically stable, the mean squares optimal one-step ahead predictor of y_s, in the sense that $\hat{y}_{s|s-1} = E[y_s|\mathcal{F}_{s-1}]$ ($\mathcal{F}_{s-1}$ is the sub-σ algebra generated by data up to time $s-1$), is

$$
\hat{y}_{s|s-1} = H^{-1}(q^{-1})G(q^{-1})u_s + (1 - H^{-1}(q^{-1}))y_s \tag{14}
$$

and we have

$$
y_s = \hat{y}_{s|s-1} + w_s \tag{15}
$$

which is generally used to calculate the likelihood. Equation (14) can be written alternatively as (Lemma 7.4.1, Goodwin and Sin 1984)

$$
\begin{aligned}
y_s &= \hat{y}_{s|s-1} + w_s \\
C(q^{-1})\hat{y}_{s|s-1} &= B'(q^{-1})u_{s-1} + C'(q^{-1})y_{s-1}
\end{aligned} \tag{16}
$$

with $B'(q^{-1}) = qB(q^{-1})$ and $C'(q^{-1})$ is the unique $n-1$th order polynomial satisfying

$$C(q^{-1}) = A(q^{-1}) + q^{-1}C'(q^{-1}) \tag{17}$$

It follows that

$$p(y_{1:n_y}|u_{1:n_u}) = N(y_{1:n_y}|\hat{y}_{1:n_y}, \sigma_y^2\mathbf{I}) \tag{18}$$

where $\hat{y}_{1:n_y} = (\hat{y}_{1|0}, \cdots, \hat{y}_{n_y|n_y-1})^T$ so that

$$q(y_{1:n_y}) = \int p(y_{1:n_y}|u_{1:n_u}) \times q(u_{1:n_u})du_{1:n_u} \propto \int N(y_{1:n_y}|\hat{y}_{1:n_y}, \sigma_y^2\mathbf{I}) \times N(u_{1:n_u}|0, Q_u)du_{1:n_u} \tag{19}$$

Therefore, the resulting probability distribution $q(y_{1:n_y})$ is also a normal distribution.

3 The Bayesian approach

3.1 Inference

When using the Bayesian approach in our problem, the measured data available for inference are $(y_{1:n_y}, x_{1:n_x}), n_x = l \times n_y$ while the unknowns are those parameters $a_1, \cdots, a_{n_a}$, $b_0, \cdots, b_{n_b}$, $c_1, \cdots, c_{n_c}$, $\phi_x, \phi_u, \sigma_y^2, \sigma_x^2, \sigma_u^2, \tau$ and the implied process $u_{1:n_u}$. The inference about the unknowns is made via posterior probability densities conditional on the measured data. It should be pointed out that conditional on the implied process, the inference about the unknowns can be done separately at two different temporal levels. At the fine level, the parameters are $\alpha = (\phi_x, \sigma_x^2)$ while at the coarse level the unknowns are $\beta = (a_1, \cdots, a_{n_a}, b_0, \cdots, b_{n_b}, c_1, \cdots, c_{n_c}, \phi_u, \sigma_y^2, \sigma_u^2, \tau)$ and $u_{1:n_u}$.

By considering $u_{1:n_u}^o$ as the observations of the implied process u_s at the coarse level, the likelihood function for the coarse level is then $p(y_{1:n_y}, u_{1:n_u}^o|\beta, u_{1:n_u})$. Instead of calculating the exact likelihood function, a spectral factorisation method is adopted in this paper to obtain an approximation of the likelihood function for this EIV ARMAX model (e.g. Soderstrom and Stoica 1989). Substituting (8) into (3) and (2) yields

$$\begin{aligned} (1-\phi_u q^{-1})A(q^{-1})y_s &= B(q^{-1})\eta_s + (1-\phi_u q^{-1})C(q^{-1})w_s \\ (1-\phi_u q^{-1})u_s^o &= \eta_s + (-1+\phi_u q^{-1})\mu_s \end{aligned} \tag{20}$$

Letting

$$z_s = \begin{pmatrix} (1-\phi_u q^{-1})A(q^{-1})y_s \\ (1-\phi_u q^{-1})u_s^o \end{pmatrix} \tag{21}$$

then z_s is an MA process so that by the spectral factorisation theorem there exists a unique matrix $S(q^{-1})$ such that

$$z_s = S(q^{-1})\epsilon_s \tag{22}$$

where $E\epsilon_{s_1}\epsilon_{s_2}^T = \Lambda\delta_{s_1,s_2}, \Lambda > 0$ and $\delta_{i,j}$ Kronecker delta. The approximation of the likelihood function is then given by

$$p(y_{1:n_y}, u_{1:n_u}^o|\beta, u_{1:n_u}) \approx (2\pi)^{-n_y}|\Lambda)|^{-n_y/2}\exp(-\frac{1}{2}\sum_{s=1}^{n_y}\epsilon_s^T\Lambda^{-1}\epsilon_s) \tag{23}$$

It follows that the full conditional distributions at the coarse level are then

$$\begin{aligned} q(u_{1:n_u}|\beta, y_{1:n_x}, u_{1:n_x}^o) &\propto p(y_{1:n_y}, u_{1:n_u}^o|\beta, u_{1:n_u}) \times q(u_{1:n_u}|\beta) \\ p(\beta|u_{1:n_u}, y_{1:n_x}, u_{1:n_x}^o) &\propto p(y_{1:n_y}, u_{1:n_u}^o|\beta, u_{1:n_u}) \times p(\beta|u_{1:n_u}) \end{aligned} \tag{24}$$

At the fine level, the joint posterior density for the unknown parameters α, $p(\alpha|x_{1:n_x}, u_{1:n_u})$, is given by

$$p(\alpha|x_{1:n_x}, u_{1:n_u}) \propto q(u_{1:n_u}|\alpha, x_{1:n_x}) \times p(\alpha|x_{1:n_u}) \tag{25}$$

Due to the difficulty in computing these full conditional distributions directly, the simulation of the unknowns will be conducted via the Gibbs sampler and the Metropolis-Hastings proposal after specifying some a priori chosen prior probability densities.

3.2 Prediction

The next objective of the paper is about the prediction problem. It is interesting to see that the prediction problem can be divided into two different cases: synchronous prediction and asynchronous prediction, due to the characteristics of the multiple scale.

1) Synchronous prediction scheme. This scheme uses the data up to n_x to produce a prediction for the quantities concerned. To perform the prediction, it is sufficient to generate a sample from the predictive distributions. The following predictive distributions are of interest

$$
\begin{array}{lcl}
p(y_{n_y+1}|u_{1:n_u}, y_{1:n_y}) & & \\
q(u_{n_u+1}|u_{1:n_u}, x_{1:n_x}) & = & q(u_{n_u+1}|u_{1:n_u}, x_{n_x}) \\
p(x_{n_x+1}|u_{1:n_u}, x_{1:n_x}) & \text{or} & p(x_{n_x+1:n_x+l}|u_{1:n_u}, x_{1:n_x})
\end{array}
\tag{26}
$$

where the first one is to forecast the system output of the coarse level at n_y+1 given observations (input and output) up to n_y, the second one is to simulate the implied process at n_u+1 given observations from the fine level up to n_x, and finally the third one is to generate the forecast of the dynamics at a fine level, given observations up to the last time instant.

2) Asynchronous prediction scheme. This scheme is to use the data up to n_y at the coarse level to produce a prediction for the quantity in the fine level. This is the advantage of the multiscale method, which enables the forecast at the fine level to be guided by the forecast at the coarse level so as to reduce the forecasting error. The following relationship

$$
p(x_{n_x-l+1:n_x}|u_{1:n_u}, x_{1:n_x-l}) = p(x_{n_x-l+1:n_x}|u_{n_u}, x_{n_x-l}) \tag{27}
$$

indicates that the forecast at some point at fine level can be given by considering the effect from the coarse level prediction ahead of this point.

The derivation of some of the relations can be found in (Ferreira, West, Lee, and Higdon 2006).

4 Applications to the relativistic electron intensity at a geosynchronous orbit

Relativistic electrons in the Earth's inner magnetosphere are very important, where the flux in the outer radiation belt ($L > 3.5$) exhibits highly dynamical behaviour. High relativistic electrons on orbiting satellites can cause electric discharges across internal satellite components, which in turn leads to possible satellite damage or failures. Therefore it is crucial to be able to understand the mechanisms of electron transport and acceleration, and the way how they depend on other processes extending from the solar wind into the inner magnetosphere. Obviously, developing a physical model is highly desirable and has been a challenge. To obtain a model from observations at a geosynchronous orbit(GO), there are many methods available now such as moving average (MA) linear filters driven with the K_p index (Nagai 1988), MA driven with the K_p index, AE index, and solar wind velocity v (Baker, McPherron, Cayton, and Klebesadel 1990), a nonlinear method (Rodgersa, Clucasa, Dyera, and Smitha 2003), and a delay embedding method (Ukhorskiy, Sitnov, Sharma, Anderson, Ohtani, and Lui 2004). While a general correlation between the growth of relativistic electron flux at a geosynchronous orbit and geomagnetic activity is well established, the existing investigations into relativistic electron flux have revealed that the effectiveness of the influence of magnetic storms on electron fluxes strongly depends on the geomagnetic index $SymH$, solar wind velocity, and other geomagnetic indices, and good results in predicting electron fluxes in GOs is obtained when they are used as input data for forecast models.

In this application, the proposed multiple scale Bayesian modelling approach is applied and the parameters of the model are derived from the correlated data. The output of the system used was the relativistic electron intensity, the daily time series of the omnidirectional flux (Je) of > 2 MeV electrons from GOES 7 and 8 satellites for years 1995 and 1996 to 2000, respectively. The input of the system was chosen to be solar wind velocity v and geomagnetic indices $SymH$, obtained hourly from Kyoto World Data Centre from 1995 to 2000. The different temporal scales in the data here motivate our method. Obviously, this is a two input and one output problem. The model at the coarse level in this simulation was chosen as ARMAX(1,1). The prior distributions of the involved quantities were chosen as: $a_1 \sim N(m_a, S_a)$, $b_0 \sim N(m_{b_0}, S_{b_0})$, $b_1 \sim N(m_{b_1}, S_{b_1})$, $\phi_{x_1}, \phi_{x_2} \sim N(m_{\phi_x}, S_{\phi_x})$, $\sigma^2_{x_1}, \sigma^2_{x_2} \sim IG(\nu_{\sigma_x}/2, \nu_{\sigma_x} s_{\sigma_x}/2)$, $\phi_{u_1}, \phi_{u_2} \sim N(m_{\phi_u}, S_{\phi_u})$, $\sigma^2_{u_1}, \sigma^2_{u_2} \sim IG(\nu_{\sigma_u}/2, \nu_{\sigma_u} s_{\sigma_u}/2)$, $\sigma^2_y \sim IG(\nu_{\sigma_y}/2, \nu_{\sigma_y} s_{\sigma}/2)$, $\tau_1, \tau_2 \sim IG(\nu_\tau/2, \nu_\tau s_\tau/2)$, and $u_{1:n_u} \sim U(\max(-1, u^o_{1:n_u} - \delta), \min(1, u^o_{1:n_u} + \delta))$, where the hyperparameters were $m_{a_1} = 0, S_{a_1} = 1000$, $m_{b_0} = 0, S_{b_0} = 1000$, $m_{b_1} = 0, S_{b_1} = 1000$, $m_{\phi_x} = 0, S_{\phi_x} = 1000$, $\nu_{\sigma_x} = 0.001$, $\nu_{\sigma_x} s_{\sigma_x} = 0.001$, $m_{\phi_u} = 0$, $S_{\phi_u} = 1000$, $\nu_{\sigma_u} = 0.001$, $\nu_{\sigma_u} s_{\sigma_u} = 0.001$, $\nu_{\sigma_y} = 0$, $\nu_{\sigma_y} s_{\sigma_y} = 0.001$, $\nu_\tau = 0.001$, $\nu_\tau s_\tau = 0.001$, and $\delta = 10\% of u^o_{1:n_u}$. The data were normalised into an interval of $(-1, 1)$. The Gibbs sampler was used to generate a total of 1000 iterations and the last 500 iterations were used to perform the statistical analysis, which is shown in Table (1). Figure (1) shows the prediction results from Nov 19, 1999 to Nov 08, 2000, with a mean square error 0.0415. For the purpose of comparison, the forecast was also calculated from a single scale method, that is using the averaged values $u^o_{1:n_u}$ directly, whose mean square error is 0.0522, which shows the efficiency of the proposed multiscale method.

5 Conclusions

A Bayesian system identification approach to the modelling and prediction of multiscale time series has been presented in this paper. The proposed approach provided a multiple scale modelling framework for time series analysis and prediction. Due to the characteristics of multiple scale, the prediction can be performed synchronously or asynchronously. It has been shown by numerical studies that the proposed method works well in applications to real data.

The proposed method can actually be considered as a basic modelling framework for multiscale

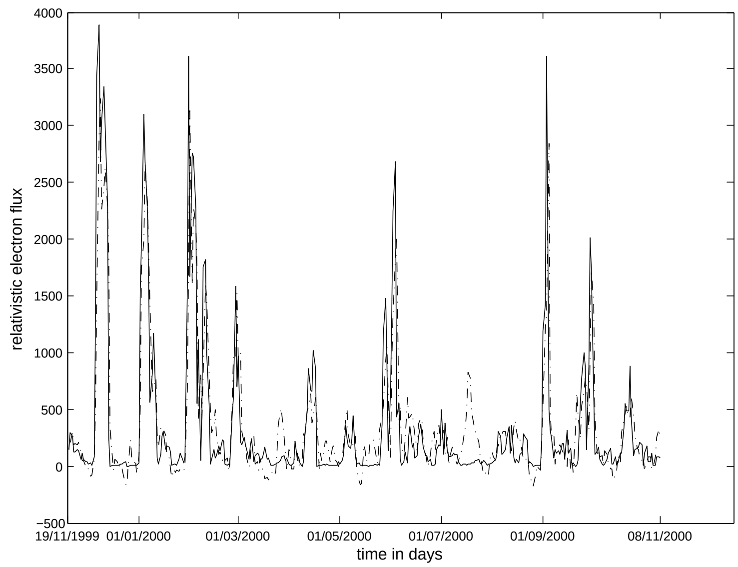

Figure 1: One-day prediction of relativistic electron intensity at geosynchronous orbit for a period from Nov 19, 1999 to Nov 08, 2000 (solid – data; dashed – one-day prediction)

Parameter	Mean	Standard Deviation
a_1	0.7329	0.0072
b_0	0.4254	0.0343
b_1	0.0351	0.0102
ϕ_{u_1}	0.9670	0.0410
ϕ_{u_2}	0.9787	0.0181
ϕ_{x_1}	0.9985	0.0267
ϕ_{x_2}	0.9983	0.0392
τ_1	0.0014	0.0007267
τ_2	0.0015	0.0007027
σ_y^2	0.2002	0.0192
$\sigma_{u_1}^2$	0.0686	0.0203
$\sigma_{u_2}^2$	0.0831	0.0202
$\sigma_{x_1}^2$	0.0183	0.0096
$\sigma_{x_2}^2$	0.1295	0.0223

Table 1: Statistical summaries for the parameters

time series. The models within this framwork can be chosen as different types such as MA and ARMA depending on the characteristics of the time series.

6 Acknowledgement

The authors gratefully acknowledge financial support from EPSRC (UK) and the European Research Council.

References

[1] Baker, D. N., McPherron, R. L., Cayton, T.E., and Klebesadel, R.W., (1990), Linear prediction filter analysis of relativistic electron properties at 6.6 RE, *J. Geophys. Res.*, Vol. 95, pp. 15113-15140.

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

[3] Daoudi, K., Frakt, A., and Willsky, A., (1999) Multiscale autoregressive models and wavelets, *IEEE Trans. on Information Theory*, Vol. 45, No. 3, pp. 828-845.

[4] Drost, F. C. and Nijman, T. E., (1993), Temporal aggregation of GARCH processes, *Econometrica*, Vol. 61, No. 4, pp. 909-927 .

[5] Ferreira, M. and Lee, H., (2007) *Multiscale Modeling - A Bayesian Perspective*, Springer, New York.

[6] Ferreira, M. A. R., West, M., Lee, H. K. H., and Higdon, D. M., (2006), Multi-scale and hidden resolution time series models, *Bayesian Analysis*, Vol.1, No. 4, pp. 947-968.

[7] Lai, T. L. and Wei, C. Z., (1982) Least-squares estimates in stochastic regression models with applications to identification and control of dynamical systems, *Annals Statistics*, Vol. 10, No. 1, pp. 154-166.

[8] Nagai, T., (1988), Space weather forecast: Prediction of relativistic electron intensity at synchronous orbit, *Geophys. Res. Lett.*, Vol. 15, pp. 425428.

[9] Nawrotha, A. P. and Peinke, J. (2006) Multiscale reconstruction of time series, *Physics Letters A*, Vol. 360, No. 2, pp. 234-237.

[10] Oigard, T. A., Rue, H., and Godtliebsen, F., (2006), Bayesian multiscale analysis for time series data, *Computational Statistics & Data Analysis*, Vol. 51, No. 3, pp. 1719-1730.

[11] Peterka, V. (1981), Bayesian approach to system identification, in *Trends And Progress in System Identification*, P. Eykhoff eds. Pergamon Press: Oxford, pp. 239-304.

[12] Rodgersa,D. J., Clucasa, S. N., Dyera, C. S., and Smitha, R. J. K. (2003), Non-linear prediction of relativistic electron flux in the outer belt, *Advances in Space Research*, Vol. 31, No. 4, pp. 1015-1020.

[13] Soderstrom, T, and Stoica, P. (1989) *System Identification*, Prentice Hall, New York.

[14] Ukhorskiy, A. Y., Sitnov, M. I., Sharma, A. S., Anderson, B. J., Ohtani, S., and Lui, A. T. Y., (2004), Data-derived forcasting model for relativistic electron intensity at geosynchronous orbit, *Geophysical Research Letters*, Vol. 31, L09806, doi:10.1029/2004GL019616.