Bayesian dynamic modelling for nonstationary hydroclimatic time series forecasting along with uncertainty quantification

D. Nagesh Kumar* and Rajib Maity†

Department of Civil Engineering, Indian Institute of Science, Bangalore - 560012, India

Abstract:

Forecasting of hydrologic time series, with the quantification of uncertainty, is an important tool for adaptive water resources management. Nonstationarity, caused by climate forcing and other factors, such as change in physical properties of catchment (urbanization, vegetation change, etc.), makes the forecasting task too difficult to model by traditional Box–Jenkins approaches. In this paper, the potential of the Bayesian dynamic modelling approach is investigated through an application to forecast a nonstationary hydroclimatic time series using relevant climate index information. The target is the time series of the volume of Devil's Lake, located in North Dakota, USA, for which it was proved difficult to forecast and quantify the associated uncertainty by traditional methods. Two different Bayesian dynamic modelling approaches are discussed, namely, a constant model and a dynamic regression model (DRM). The constant model uses the information of past observed values of the same time series, whereas the DRM utilizes the information from a causal time series as an exogenous input. Noting that the North Atlantic Oscillation (NAO) index appears to co-vary with the time series of Devil's Lake annual volume, its use as an exogenous predictor is explored in the case study. The results of both the Bayesian dynamic models are compared with those from the traditional Box–Jenkins time series modelling approach. Although, in this particular case study, it is observed that the DRM performs marginally better than traditional models, the major strength of Bayesian dynamic models lies in the quantification of prediction uncertainty, which is of great value in hydrology, particularly under the recent climate change scenario. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS   Bayesian dynamic models; nonstationarity; forecasting; uncertainty; Devil's Lake

Received 5 May 2007; Accepted 26 October 2007

INTRODUCTION

In hydrologic time series analysis and forecasting, Box–Jenkins models (Box *et al*., 1994) are widely used. These models are static in nature and use the statistical properties of time series. A contiguous sequence of a few recent observations and errors in prediction are used as input to these models. However, the non-contiguous nature of autocorrelation and partial autocorrelation of somewhat more complex time series motivate the use of non-contiguous types of Box–Jenkins models (Mujumdar and Nagesh Kumar, 1990).

However, the basic assumption of all such models is the temporal persistence of statistical properties of the time series. Thus, once the parameters are determined, they are assumed to remain constant over time. This is often not a valid assumption, since climatic and other factors influencing the dynamics may cause changes in the statistical properties of time series over time. In general, uncertainty associated with observations is considered in traditional time series modelling approaches. However, uncertainty in model parameter values and model structure, which is another important source of uncertainty,

is ignored in traditional approaches. Another drawback of Box–Jenkins models is the exploitation of a significant amount of data for determining the parameters and for validation the model before it becomes ready for use. Large data sets may not always be available, particularly in developing countries.

Shortcomings of traditional modelling approaches, as briefly mentioned above, lead researchers towards advanced modelling approaches. In this paper, forecasting of nonstationary hydrologic time series using climatic information as inputs, along with quantification of uncertainty is investigated. This approach is very useful in the context of recent climate change scenarios as the effect of climate change imparts nonstationarity to hydrological time series. Bayesian dynamic models (BDMs) are able to deal with nonstationary time series. It is shown with a case study that BDMs have potential for hydroclimatic time series forecasting and quantification of uncertainity. The superiority of such models over traditional modelling approaches is explained here. The common assumption of stationarity can be relaxed in BDMs (Bernier, 1994). Dynamic properties enable such models to gradually update the model parameters in light of any changes in the time series arising from climatic or other causes. If the change is abrupt and can be predicted earlier, typically in the case of man-made changes, it may be possible to incorporate its effect in the model manually, based on decisions of the forecaster, which is

* Correspondence to: D. Nagesh Kumar, Department of Civil Engineering, Indian Institute of Science, Bangalore - 560012, India.
E-mail: nagesh@civil.iisc.ernet.in
† Current address: Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai - 400076, India.

known as intervention. However, this is not easy from a practical point of view in the field of hydrology. Thus, in this study, external intervention is not utilized.

The major strength of BDMs is its uncertainty quantification of predicted values. Forecasting is a statement of an uncertain future value and Bayesian philosophy believes in probabilistic representation of all sources of uncertainty. Thus, prediction of uncertain future values is available as a probability statement. Although such models are used in other fields of study (Pole *et al.*, 1994), only a few applications have been reported in the field of hydrology (Bernier, 1994; Berger and Insua, 1998; Krishnaswamy *et al.*, 2000, 2001) and its potential use in time series analysis and forecasting is yet to be established. It may be noted here that classical Kalman filtering, which is based on Bayesian philosophy, has also been used in hydrology and water resources engineering problems (Bergman and Delleur, 1985; Bierkens *et al.*, 2001; Schreider *et al.*, 2001; Crow and Wood, 2003). However, Kalman filtering is not Bayesian forecasting (West and Harrison, 1997). In this study, the utility of Bayesian dynamic models for hydrologic time series forecasting and quantification of the uncertainty involved is illustrated with the annual volume time series of Devil's Lake, located in North Dakota, USA.

The rest of the paper is organized as follows. The analytical structure of two Bayesian dynamic models—a constant model (CM) and a dynamic regression model (DRM) is presented in the following section. A brief description of Devil's Lake annual volume time series is provided in the third section, followed in the next section by an illustration of the use of CM and its performance in the prediction of annual volumes in Devil's Lake. The North Atlantic Oscillation (NAO) index, an exogenous predictor, is discussed. Implementation of the DRM is discussed in the fifth section using the NAO index as an exogenous predictor. Comparison of the two Bayesian dynamic models and Box–Jenkins models is then presented, with conclusions presented in the seventh and final section.

BAYESIAN DYNAMIC MODELS

The mathematical framework and related proof of Bayesian dynamic models can be found in West and Harrison (1997). Application of such models in other fields can be found elsewhere (Berliner *et al.*, 2000; Maity and Nagesh Kumar, 2006). The superiority of Bayesian dynamic models over traditional modelling approaches is mentioned briefly in the previous section. Relaxation of the stationarity property is the most useful aspect in hydrology as explained in this paper. The time-varying property of hydrologic time series can be captured by such models. This property is discussed in the context of a particular Bayesian model, known as a constant model, which uses the inherent property of the time series and updates the model parameters by utilizing the observed values, at each time step. On the other hand, the time-varying relationship between the causal time series and

the response time series is captured by a DRM, which uses information from the causal time series and also the inherent properties of the time series. As with the constant model, model parameters are updated by utilizing the observed values, at each time step. However, the assumption of normality is the basic assumption of Bayesian dynamic models.

Two forms of Bayesian dynamic model (BDM) are investigated for application to the Devil's Lake volume time series.

Constant Model

The observation equation is expressed as

$$Y_t = \mu_t + \nu_t \qquad \nu_t \sim N[0, V] \qquad (1)$$

and the system equation is expressed as

$$\mu_t = \mu_{t-1} + \omega_t \qquad \omega_t \sim N[0, W] \qquad (2)$$

Equation (1) considers the uncertainty in the observations. This includes measurement error, as well as sampling error, or error associated with the representativeness of the observation of the underlying mean process. Here, Y_t is the observation, μ_t is the mean level or state of the underlying process, and ν_t is the associated error process, which is assumed to have a Normal distribution with mean zero and variance V. Equation (2), while evolving over successive time steps, imparts due weighting to previous states of the observation and considers the dynamics of the underlying process, which is assumed to be a random walk—i.e. the mean of the process is allowed to drift, subject to some random perturbation ω_t, which is also assumed to have a Normal distribution with mean zero and variance W. If the variance of this perturbation is small enough then the system equation is that of a constant mean. Conversely, if it is large, the underlying dynamics is one of a random walk. The variances V and W are known as observational and evolution variance, respectively. The constant model assumes the variance of both the random terms to be constant over time. The information available at time t is denoted as D_t, and statements of system state are made conditional on that information.

The initial information on model parameter $(\mu_0/D_0) \sim N[m_0, C_0]$ is provided by the forecaster, i.e. mean m_0 and variance C_0 are the initial beliefs of the forecaster. The model parameter is updated at each time step following Bayes theorem to obtain the one-step-ahead forecast and posterior distribution of the model parameter. The procedure is explained below. The following procedure utilizes Bayes rule to obtain the forecast distribution and to update the prior distribution to posterior distribution, which is outlined in Appendix A for the general interest of the reader.

Without the loss of generality, let us assume that the posterior distribution for μ_{t-1} is $(\mu_{t-1}/D_{t-1}) \sim N[m_{t-1}, C_{t-1}]$ with some mean m_{t-1} and variance C_{t-1}.

The prior for μ_t is $(\mu_t/D_{t-1}) \sim N[m_{t-1}, R_t]$ where variance R_t is

$$R_t = C_{t-1} + W \qquad (3)$$

where W is the evolution variance as used in the system equation, i.e. while calculating the prior from the posterior of the previous time step, the variance is increased by the amount W to reflect the variance of the perturbation ω_t. The one-step-ahead forecast is $(Y_t/D_{t-1}) \sim N[f_t, Q_t]$, where

$$f_t = m_{t-1} \qquad (4)$$
$$Q_t = R_t + V \qquad (5)$$

where V is the observational variance associated with v_t, the error process in the observational equation. The expectation of the one-step forecast distribution can be used as point prediction, and the variance as the uncertainty information associated with it. The posterior for μ_t is $(\mu_t/D_t) \sim N[m_t, C_t]$, where

$$m_t = m_{t-1} + A_t e_t \qquad (6)$$
$$C_t = A_t V_t \qquad (7)$$
$$A_t = R_t/Q_t \qquad (8)$$
$$e_t = y_t - f_t \qquad (9)$$

y_t is the observed value of Y_t for the time step t and A_t is known as the adaptive coefficient. This can be thought of as a weight function, since it prescribes how much weight is given to past information and how much to the present observation.

Dynamic Regression Model (DRM)

The basic principle of a DRM is to capture the continuously changing relationship between two time series over time. The complete model is defined as follows:

The observation equation is

$$Y_t = F_t \theta_t + v_t \qquad v_t \sim N[0, V] \qquad (10)$$

and the system equation is

$$\theta_t = \theta_{t-1} + \omega_t \qquad \omega_t \sim T_{t-1}[0, W_t] \qquad (11)$$

where F_t is the regressor parameter at time t and θ_t is the regression parameter at time t. v_t is the associated error process, which is assumed to have Normal distribution with mean zero and variance V. V is the observational variance and considered to be unknown in this case. In the system equation, T_{t-1} implies Students' t distribution with $(t-1)$ degrees of freedom. W_t is time-dependent evolution variance. Dealing with unknown observational variance and time-dependent evolution variance is discussed later.

One-step forecast and posterior distribution are calculated as follows—the initial information is $(\theta_0/D_0) \sim T_1[m_0, C_0]$, where mean, m_0 and variance C_0 are the initial beliefs of the forecaster. The posterior for θ_{t-1} is

$(\theta_{t-1}/D_{t-1}) \sim T_{t-1}[m_{t-1}, C_{t-1}]$, with some mean m_{t-1} and variance C_{t-1}. The prior for θ_t is $(\theta_t/D_{t-1}) \sim T_{t-1}[m_{t-1}, R_t]$, where

$$R_t = C_{t-1} + W_t \qquad (12)$$

It is difficult, practically, to assign the sequence of evolution variance $\{W_t\}$ without relating it to some previous known variance. So a discount factor δ ($0 < \delta < 1$) is introduced such that

$$R_t = C_{t-1}/\delta \qquad (13)$$

This reflects the reality that $R_t > C_{t-1}$. Using Equation (13) in Equation (12) it can be shown after some algebraic manipulation that

$$W_t = C_{t-1}(\delta^{-1} - 1) \qquad (14)$$

As mentioned earlier, the observational variance is considered to be unknown, so a new parameter ϕ is defined as $\phi = V^{-1}$, which reflects the precision of prediction. The distributional form of the precision parameter ϕ, for the time step $(t-1)$ is $(\phi/D_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2]$, where G denotes the Gamma distribution. One-step forecast distribution is $(Y_t/D_{t-1}) \sim T_{t-1}[f_t, Q_t]$, where

$$f_t = F_t m_{t-1} \qquad (15)$$
$$Q_t = F_t^2 R_t + S_{t-1} \qquad (16)$$
$$S_{t-1} = d_{t-1}/n_{t-1} \qquad (17)$$

As in the constant model, the expectation of the one-step forecast distribution can be used as a point prediction, and the variance as the uncertainty information associated with it. The posterior distribution for θ_t is $(\theta_t/D_t) \sim T_t[m_t, C_t]$, with

$$m_t = m_{t-1} + A_t e_t \qquad (18)$$
$$C_t = R_t S_t/Q_t \qquad (19)$$
$$S_t = d_t/n_t \qquad (20)$$
$$e_t = Y_t - f_t \qquad (21)$$
$$A_t = F_t R_t/Q_t \qquad (22)$$

The updated distributional form of precision parameter ϕ, for the time step t, is $(\phi/D_t) \sim G[n_t/2, d_t/2]$, where

$$n_t = n_{t-1} + 1 \qquad (23)$$
$$d_t = d_{t-1} + S_{t-1} e_t^2/Q_t \qquad (24)$$

Two things are to be noted here. The first is that the discount factor δ plays an important role regarding the amount of information loss over successive observations. The second point is to select a suitable regressor time series, which is also known as the influential or causal time series.

It is worthwhile to once again note here that the Bayesian models presented in this section, are able to capture changes in the time series arising from its
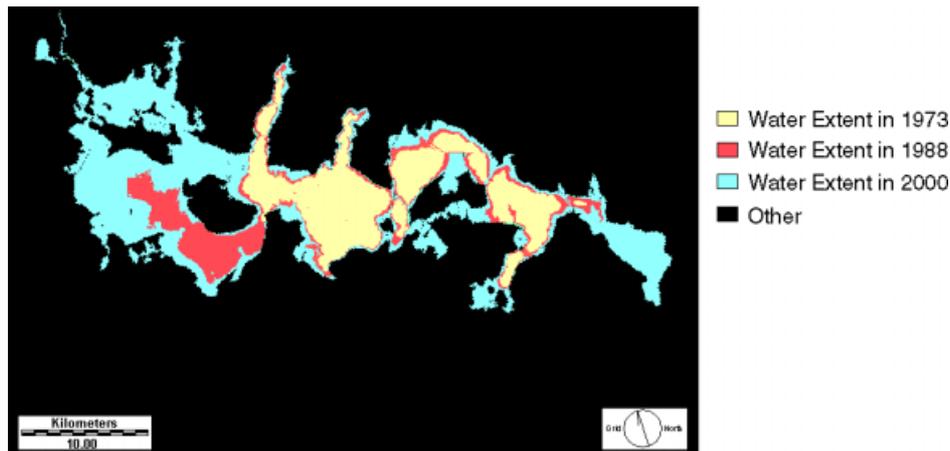
Figure 1. Volume change of Devil's Lake (Source: http://www.geocities.com/)
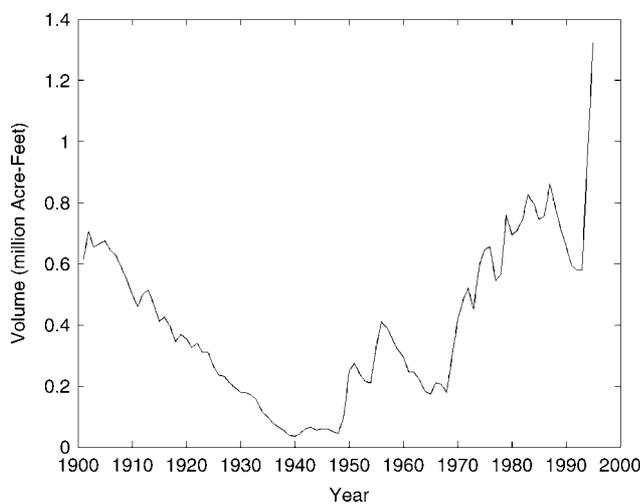

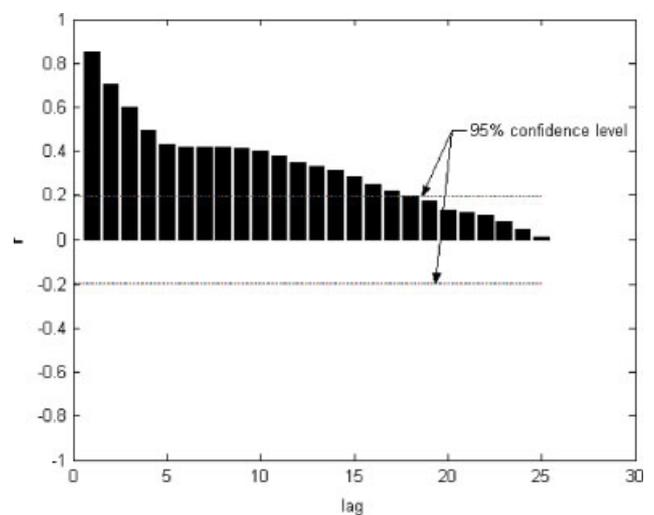Figure 2. Annual time series of Devil's Lake volume

Figure 3. Autocorrelogram of annual time series of Devil's Lake volume

dynamic nature. The factor A_t in Equations (6), (8), (18) and (22) is known as the adaptive coefficient, which can lie between 0 and 1. Equations (6) or (18) can also be rewritten as $m_t = A_t Y_t + (1 - A_t) m_{t-1}$, i.e. the present level depends on both the previous estimated level and the present observation (West and Harrison, 1997). Thus, any change in the time series is captured by the model. So, these models are suitable for capturing the time-varying relationship induced by climate forcing, and hence used in this paper. The superiority of these models over traditional Box–Jenkins models is discussed later.

DEVIL'S LAKE ANNUAL VOLUME TIME SERIES AND NONSTATIONARITY

Devil's Lake (98°52′30″W to 98°45′00″W and 47°59′53″N to 48°07′23″N) is situated in the state of North Dakota, USA. After 1970 the lake started increasing in extent gradually. A comparison of the lake area at three stages in the last three decades is shown in Figure 1. The time series of annual volume of Devil's Lake was obtained from the US Army Corps of Engineers for the period 1901 to 1995 (Figure 2).

A visual plot of a time series is often used to make a first judgment as to whether the data are stationary or nonstationary. The autocorrelogram has sometimes been used to test whether a series is stationary in mean or not—if it drops to zero relatively quickly, it is known to be stationary and on the other hand if it dies very slowly, i.e. autocorrelations remain significant at higher lags too, it is considered to be nonstationary (Makridakis *et al.*, 1998). In the present case the time series plot (Figure 2) and autocorrelogram (Figure 3) indicate that this time series is nonstationary. The assumptions of Normal distribution, as mentioned earlier, cannot be verified meticulously due to the long-term nonstationary nature of the data for this closed basin. However, the volume of a lake within a closed basin can be assumed to be the collective response of many hydrologic time series, such as, inflow through various streams, overland flow, contribution of subsurface flow, rainfall, losses due to various factors, and so on. With an assumption of their continuous distributions, it can be stated that the collective response, i.e. their summation (i.e. time series of volume) approximately follows a Normal density (Papoulis and Pillai, 2002). Thus, the

normality assumption can be statistically supported by the Central Limit Theorem for the time series of annual volume of a closed basin lake.

PERFORMANCE OF THE CONSTANT MODEL

As discussed earlier, to start prediction with the Bayesian dynamic model, the distributional form of the model parameters has to be defined, which comes from the initial belief of the forecaster. As the first value in the time series is 615,000 acre-feet, a mean level m_0 should be chosen close to it. However, as the precision of this value is not known, a high value of initial variance C_0 should be chosen. Considering these points, at time step 0 (1900) m_0 is assumed to be $600\,000$ and variance C_0 as 400×10^6. It may be noted that the effect of these initial assumed values dies down quickly after three or four time steps (West and Harrison, 1997). The constant values of observational and evolution variances are 'judicially' assumed to be $V = 250 \times 10^6$ and $W = 500 \times 10^6$, respectively. However, these 'judicial' assumptions are very difficult to be verified from a practical point of view. It may be noted that the assumed values of observational and evolution variances will have an effect on the uncertainty measurement or, in other words, confidence interval of the prediction. Thus, it is preferred to observe the performance of the model with some trial values of observational and evolution variances. Suitable values are then selected. However, this can be overcome in the DRM by introducing a precision parameter (inverse of variance) as discussed later.

With these initial beliefs the one-step-ahead forecast performance is compared with the actual observation. As already discussed, predictions are available in a distributional form and thus the confidence interval of predicted values can be obtained at any desired statistical confidence level. In Figure 4, the observation and one-step-ahead values forecast by the constant model along with the 95% confidence interval, are shown. Forecasting performances are evaluated based on R^2 values. As the time series is highly nonstationary, R^2 values are obtained for different stretches of the time series, as well as for the full length time series (Table I).

It may be observed from Figure 4 that, in the periods 1950–1952, 1956–1957, 1970–1973 and 1995–1998 the forecast performance is not as good as in other periods and is outside the 95% confidence limits, suggesting that improvements to the model are needed. This may be due to some sudden changes occurring at the beginning of those periods. If in advance it is possible to predict the expected effect, this can be incorporated in the model by external intervention, which is a major advantage of forecasting with Bayesian dynamic models. But such interventions are extremely difficult to implement operationally. It may also be noted that, using the recursive approach of the constant model, forecasts can be made available for longer lead times. However, the expectation of the distribution remains the same and the variance (uncertainty) goes up with forecast lead time. Thus, eventually the forecast information becomes 'diluted' as the forecast lead time is increased.

The possibility of external influence on the variability of the time series indicates that the inherent property of the time series is not sufficient for prediction with

Table I. Values of R^2 (Constant model)

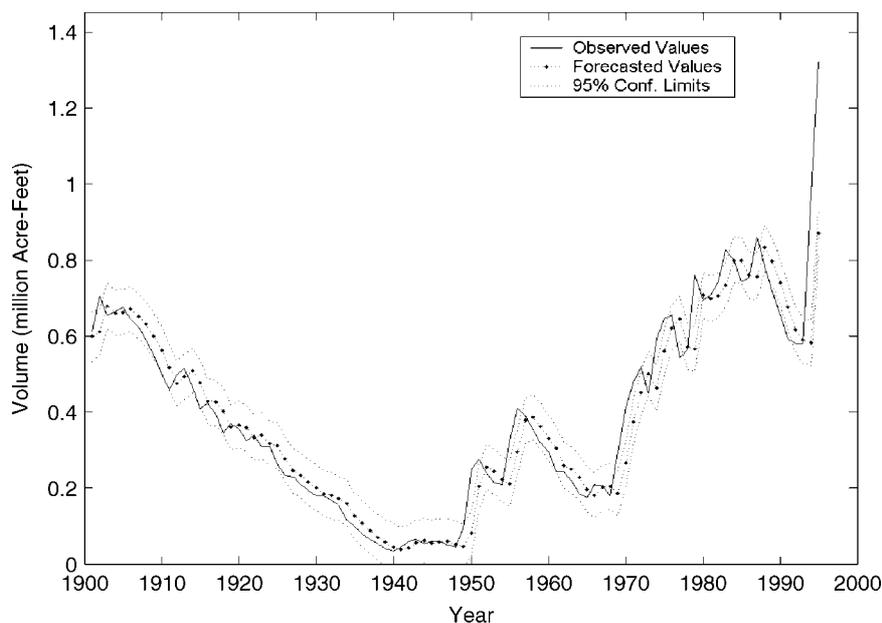| Period	R^2
1901 to 1995	0·90
1901 to 1949	0·98
1950 to 1995	0·82

Figure 4. Comparison of actual and one-step forecast by constant model

reasonable accuracy. This possibility is explored by using an exogenous predictor, which is elaborated in the next section.

PERFORMANCE OF THE DRM

The DRM needs a causal time series as an exogenous input. Based on recent studies, the North Atlantic Oscillation (NAO) is used as an exogenous input in the present study. A brief discussion of the North Atlantic Oscillation (NAO) and motivation for using it as an exogenous input is presented.

North Atlantic Oscillation (NAO)

The NAO refers to a north–south oscillation in atmospheric mass with centres of action near Iceland and over the subtropical Atlantic from the Azores across the Iberian Peninsula (Hurrell, 2000). The winter NAO index is defined as the anomalous difference between the polar low and the subtropical high during the winter season (December through March).

Effect of NAO Index

The NAO exerts a dominant influence on wintertime temperatures across much of the Northern Hemisphere. Surface air temperature and sea surface temperature (SST) across wide regions of the North Atlantic Ocean, North America, the Arctic, Eurasia and the Mediterranean are significantly correlated with NAO variability. Recent changes in the NAO trend are also reflected in pronounced changes in the transport and convergence of atmospheric moisture and, thus, the distribution of evaporation and precipitation (Hurrell, 2000). The NAO index's trend through the last three decades may be related to global climatological changes. The positive phase persistence appears to be unprecedented in the observational period (Hurrell, 1995a). This increasing trend started around 1970. From 1989 winter unprecedented positive values of the index have been recorded (Hurrell, 1995a; Watanabe and Nitta, 1999). This trend in the NAO index accounts for several remarkable climatological changes in the Northern Hemisphere, i.e. drier milder winter in Europe and a wetter, more severe winter over eastern Canada and north-west Atlantic (Hurrell, 1995a; Wallace et al., 1995; Hurrell, 1996; Shabbar et al., 1997), change in precipitation (Hurrell, 1995a; Dai et al., 1997; Hurrell and Loon, 1997), glacier movement (Sigurdsson and Jonsson, 1995), changes in sea-ice cover (Maslanik et al., 1996; Cavalieri et al., 1997), decrease in mean sea level pressure (Walsh et al., 1996), changes in strength and character of the Atlantic meridional overturning circulation due to change in convection in the Labrador and Greenland–Iceland Seas (Dickson et al., 1996; Houghton, 1996) and storm activities (Hurrell, 1995b).

NAO Index Data

The winter NAO index is based on the difference of normalized sea level pressure (SLP) between Lisbon, Portugal and Stykkisholmur/Reykjavik, Iceland since 1864. The data set is collected from the home page of Jim Hurrell, a scientist in the Climate Analysis Section in the Climate and Global Dynamics Division at National Centre for Atmospheric Research (NCAR). The SLP anomalies at each station were normalized by division of each seasonal mean pressure by the long-term mean (1864–1983) standard deviation. Positive values of the index indicate stronger-than-average westerlies over the middle latitudes. The time series of NAO index (Dec–Mar) for the period 1892–1995 are shown in Figure 5. Station index value for year N refers to an average of December year $N - 1$ and January, February, and March year N. For example, the 1990 value contains the average of December 1989 and January, February, and March 1990. [Source: http://www.cgd.ucar.edu/~jhurrell/nao.stat.winter.html].

Exogenous Predictor—the NAO Index

While a dynamical connection cannot be directly made between the Devil's Lake volume and the NAO index, from Figures 2 and 5, a correspondence can be seen between the trends in the NAO index time series and changes in the time series of Devil's Lake annual volume. A simple regression model is adopted for the preliminary study of the relationship between them. However, the relationship is not well captured by the simple regression model. The reason may be the static nature of the regression model whereas a time varying relationship is expected as indicated by the above analysis. Moreover, quantification of uncertainty is very important for the modelling of time varying relationships. Keeping these points in mind, a dynamic model is required for application, which is able to capture the dynamic relationship and provides the information uncertainty associated with the prediction.

Since the Devil's Lake is a closed basin lake, its volume reflects an integration of basin precipitation over some prior time period. Thus, a closed basin lake may reflect an aggregation or integration of the
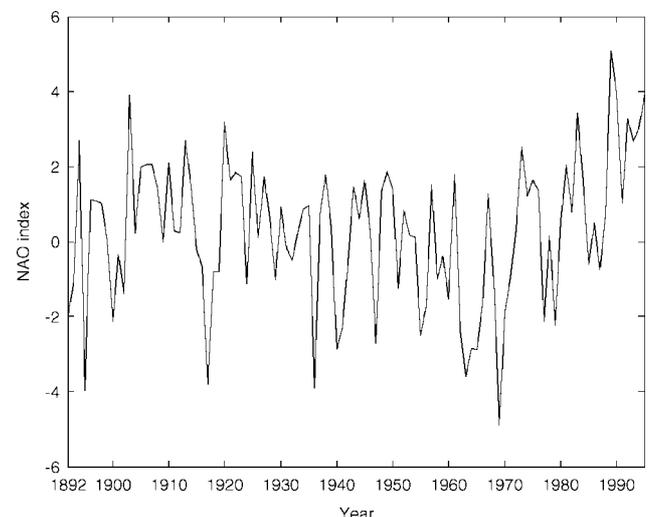

Figure 5. Time series of NAO index (Dec–Mar)

meteorologic processes represented by the climate index, over some preceding time period (Baldwin and Lall, 2000). Hence, the volumetric response of a closed basin like Devil's Lake is expected to vary with the aggregation or integration of climatic forces and varies slowly in time. Thus, the average value of the climate index over 'N' preceding years is used as precursor.

The value of 'N' is investigated using a measure of similarity between the two time series. Among the different similarity measures between two time series, Euclidean distance is widely used. Euclidean distance is a distance metric between two time series, considering each to be a point in n-dimensional space, where n is the length of the time series (Agrawal *et al.*, 1993; Ma and Manjunath, 1996; Park *et al.*, 1999; Maity and Nagesh Kumar, 2007). The Euclidean distance between two time series X and Y is defined as

$$D_E(X, Y) = \left(\sum_{i=1}^{n} (X_i - Y_i)^2 \right)^{1/2} \quad (25)$$

The smaller the Euclidean distance, the closer are the two time series considered. The Euclidean distances were calculated between the time series of standardized Devil's Lake volume and the time series of moving average NAO index for different window sizes (2 to 20 years). The results are shown in Figure 6.

It is observed that the Euclidean distance is lowest for the moving average of the preceding 8 years. The numerical values for the Euclidean distances for the moving average of 9 and 10 years are almost the same as those for 8 years. Here, a moving average of the preceding 10 years is used, this being somewhat more of a standard than 8 or 9 years.

Thus, NAO moving average of the preceding 10 years is selected as the exogenous predictor for Devil's Lake annual volume for the current year. Baldwin and Lall (2000) had explored a similar connection between the Devil's Lake, the NAO and other climate indices. In their study, a running sum of the departures of the NAO from its mean value is considered as a predictor. However, the considered NAO index is based on the anomaly values of sea level pressure, as mentioned earlier. Thus, the departures from the mean are not calculated, as it is already considered in the NAO index.

Performance of DRM model using the NAO index as an exogenous input

Initial values of the parameters m_0, C_0, d_0 and n_0, at time step 0 (1900), are assumed to be the forecaster's initial belief. The explanation for assumed m_0 and C_0 is the same as for the constant model and these are selected to be $1 \cdot 15 \times 10^5$ and $1 \cdot 5 \times 10^3$, respectively. As mentioned earlier, in the case of DRM, a precision parameter, denoted ϕ ($= V^{-1}$) is introduced, which is Gamma distributed with shape parameter $n_t/2$, scale parameter $d_t/2$. Expectation of this distribution gives an estimate of the observational variance. Initial information on these parameters, i.e. n_0 and d_0, comes from the initial belief of the forecaster. n_0 is selected to be 1 as it is increased by 1 during updating at each time step. d_0 is selected such that it will reflect the initial estimate of unknown observational variance, V. It is assumed to be 1×10^5, so that initial estimate of V ($= d_0/n_0 = 1 \times 10^5$) is a little higher, as the accuracy of this value is not known. However, as in the previous case, the effect of these initial assumed values dies down quickly after three or four time steps (West and Harrison, 1997). However, it is noticed that the effect of d_0 on the observational variance takes a comparatively longer time to die down.

The observation and one-step-ahead forecasted values along with the 95% confidence intervals, are shown in Figure 7. Forecasting performances are evaluated based on R^2 values as earlier (Table II). It can be observed that the model performance is improved, particularly in those years when the performance of the constant model is not satisfactory. This is because of the use of NAO index information as an exogenous input to this DRM, which is not present in the constant model. A comparison of performance of this model with those of the constant model and Box–Jenkins models is presented in the next section.

It may be noted that, in the DRM the coefficient of determination (R^2) may be worse than that in the constant model, if selection of the exogenous input is unsuitable. In the present problem, the worth of the NAO index is established by the improved performance of the DRM.

Another point is that during the time when the drastic change in Devil's Lake volume starts occurring, particularly in the period 1971–1995, the difference in the performance of the constant model and the DRM is conspicuous. In fact, 43% ($r = 0.648$) of the total variability
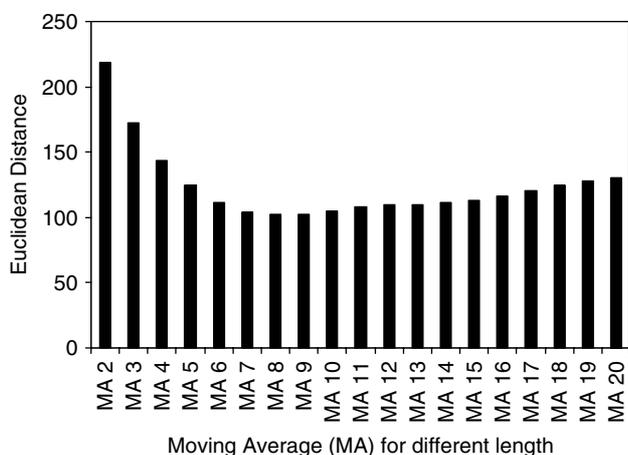

Figure 6. Euclidean distances between the time series of standardize Devil's Lake volume and the time series of moving average NAO index for different window sizes (2 to 20 years)

Table II. Values of R^2 (DRM)

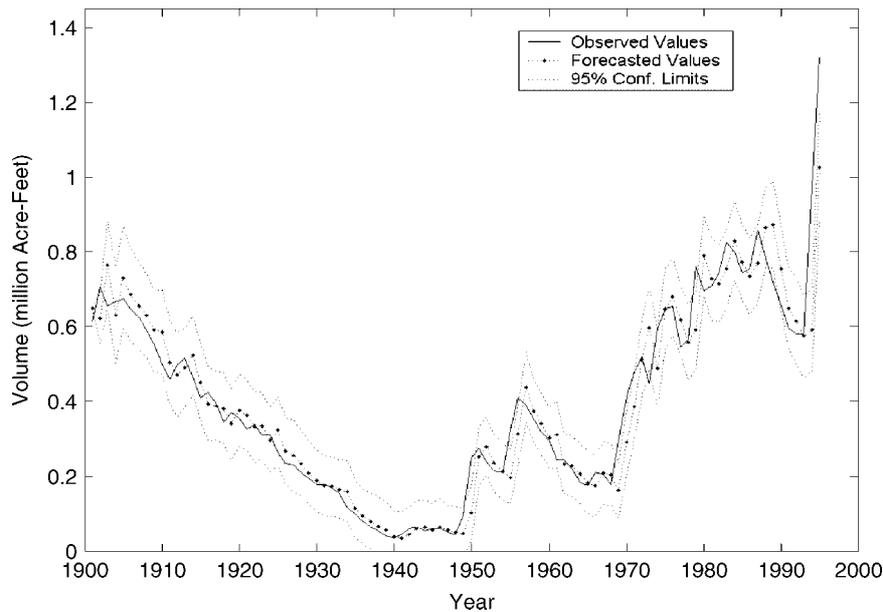| Period	R^2
1901 to 1995	0·91
1901 to 1949	0·98
1950 to 1995	0·85

Figure 7. One-step forecast with DRM with NAO index as an exogenous input

is explained in a constant model that does not use the NAO index information, whereas 52% ($r = 0{\cdot}721$) of the variability is explained in the DRM that does use the NAO index information.

Another point is that the possibility of longer lead time forecasting is more reasonable than with the constant model. However, the k-step-ahead forecast needs the exogenous input information (here NAO Index) for the kth step ahead from the present time step. As this information is not available, longer lead time forecasting was not attempted.

COMPARISON OF BOX–JENKINS MODELS WITH BAYESIAN MODELS

Before proceeding further, it is once again worth mentioning here that major strength of Bayesian dynamic models lies in the quantification of the uncertainty associated with predicted values by considering the uncertainty of the model parameters, which is not considered in Box–Jenkins models. Another point is that as Box–Jenkins models use a significant quantity of data for model development, comparison can be done only with a smaller part of the data. These obvious shortcomings of Box–Jenkins models should be kept in mind before evaluating the comparison results.

Tentative Box–Jenkins models, which are widely known as autoregressive integrated moving average (ARIMA) model, were developed based on data for the period 1901–1970, following the usual procedure (Box *et al*., 1994). Three tentatively selected models—AR(2), ARIMA(1,1,1) and ARIMA(2,2,1), were tested for the remaining period 1971–1995. It may be noted here that a naïve forecasting model such as '1 month ahead is equal to last month' is recommended sometimes as a baseline model, which is a special case of the AR(1) model ($Y_t = a_1 Y_{t-1} + e_t$) with parameter $a_1 = 1$, so that

the expectation at time step t, i.e. forecast at time step t will be equal to the observed value at time step $t - 1$. However, it is found that the partial correlation coefficient is significant up to the second lag and the possibility of AR(1) model is ruled out along with the naïve forecasting model. However, being the baseline model, the performance of such a model is also tested and reported.

It is also important to investigate the performance of autoregressive model with exogenous input (ARX) in this regard. The general form of ARX models [$ARX(n, m, k)$] is given by $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_n Y_{t-n} + b_1 X_{t-k} + b_2 X_{t-k-1} + \cdots + b_m X_{t-k-m+1} + e_t$, where, Y is the response variable, i.e. volume of Devil's Lake, and X is the exogenous input; i.e. NAO Index. $a_i (i = 1, \ldots, n)$ and $b_j (j = 1, \ldots, m)$ are model parameters, which are estimated based on the 1901–1970 data. n, m and k are the order of the autoregressive parameter, order of the exogenous input parameter and delay factor. n is decided by significant partial autocorrelation of Devil's Lake volume data and was found to be 2. m is decided by the significant cross-correlation between Devil's Lake volume data and NAO Index data and was found to be 1. k is equal to zero because the nth year NAO index refers to an average of December from the $(n - 1)$th year and January, February and March from the nth year. Thus the model is reduced to $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + b_1 X_t + e_t$, i.e. ARX(2,1,0). The model is tested for the period 1971–1995, as for the other Box–Jenkins models.

It is necessary to test statistically the residual for (a) zero mean, (b) no significant periodicity, and (c) absence of correlation to validate the ARIMA models (Mujumdar and Nagesh Kumar, 1990). Brief details of these statistical tests are presented in Appendix B. All these tests were carried out and test results are given in Table III. Cells marked with the asterisk indicate failure to pass the test.

To compare the performance of Bayesian dynamic models with that of ARIMA and ARX models, mean square error (MSE), mean absolute deviations (MAD) and R-square (R^2) values were calculated for all cases for the period 1971–1995, as this was the model testing period for the ARIMA models. Model performance statistics are shown in Table IV. It may be mentioned here that the naive forecasting model, i.e. forecast at time step t equal to the observed value at time step $t − 1$, is tested and the statistics are obtained as $MSE = 1·62 \times 10^{10}$, $MAD = 8·46 \times 10^4$ and $R^2 = 0·52$ for the period 1971–1995. These values were compared with the statistics for the AR(2) model (Table IV) and found to be inferior, which is expected as explained earlier. It is further noticed that the AR(2) and the DRM with NAO as exogenous predictor are quite close in their performance as far as their MSE and MAD are concerned. However, residuals of the AR(2) model fail the test of zero mean and the test of significant periodicity during this period.

Moving to ARIMA models of progressively higher order (e.g. 2,2,1) leads to substantially higher MSE and MAD, i.e. degradation in predictive performance, but the mean of the residuals is close to zero. For the ARIMA (2,2,1) model, the residuals are autocorrelated, suggesting its rejection. The performance of ARX(2,1,0) is found to be better than other Box–Jenkins models. However residuals of this model fail the test of significant periodicity (Table III). Although the performance, in terms of statistics (shown in Table IV), is comparable with the DRM, quantification of uncertainty associated with the predicted values is the main strength of the latter. This is achieved by considering the uncertainty of the DRM model parameters as mentioned earlier.

Essentially, the Bayesian dynamic models by considering the mean change over time are able to capture the nonstationarity in the mean. The DRM with NAO does bring in some information beyond that contained in the constant model. Since the lake evolves slowly, the lag-1 autocorrelation is typically very high (0·95, with lag-2 equal to −0·25 for the fitted AR(2) model). In the absence of external, nonstationary forcing, this sort of system will exhibit long memory and significant, persistent excursions about the long-term mean. So, one question is whether the recent rise of Devil's Lake is simply such an excursion and does not reflect a change in the forcing function. The performance of the ARIMA (p,d,q) models considered here suggests otherwise. There is a persistent bias in the mean for the lower order models, and once the model with d = 2, consistent with the test of nonstationarity, is considered, it is found that the residual structure is still not consistent with the underlying assumptions. The DRM allows the flexibility of changing dynamics with the inclusion of additional information in the linear autoregressive model. Further, it provides information regarding the uncertainty associated with the forecasted values, which is a valuable contribution towards better management of water resources.

CONCLUSIONS

In this paper, the Bayesian dynamic modelling approach is explored for nonstationary hydrologic time series forecasting. Two linear models were considered with relatively consistent assumptions as to the model form and underlying density functions. Performance of the models was compared with that of the traditional Box–Jenkins models with a case study of Devil's Lake annual volume

Table III. Residual tests for ARIMA models during testing period (1971–1995)

Model	Test for zero mean		Test for significant periodicities					Test for absence of correlation	
			Test statistic (P-value)						
	Test Statistic (P-value)	$t_{0.975}(24)$	1	2	3	$f_{0.975}(2,23)$	Test Statistic (P-value)	$\chi^2_{0.975}(5)$
AR(2)	2·09* (0·0474)	2·06	14·28* (0·0001)	0·01 (0·993)	0·39 (0·681)	4·35	2·47 (0·781)	12·83
ARIMA(1,1,1)	2·17* (0·0401)	2·06	15·25* (0·0001)	0·60 (0·557)	1·24 (0·308)	4·35	6·16 (0·291)	12·83
ARIMA(2,2,1)	−0·01 (0·9921)	2·06	0·00 (1·000)	0·04 (0·961)	0·38 (0·688)	4·35	18·14* (0·0028)	12·83
ARX(2,1,0)	1·789 (0·0781)	2·06	10·94* (0·0005)	0·002 (0·998)	0·36 (0·702)	4·35	1·96 (0·855)	12·83

Table IV. Statistical comparisons between different ARIMA models and Bayesian models for the period 1971–1995

	AR(2)	ARIMA (1,1,1)	ARIMA (2,2,1)	ARX (2,1,0)	Constant model	DRM taking NAO index as exogenous input
MSE	$1·58 \times 10^{10}$	$3·15 \times 10^{10}$	$4·84 \times 10^{10}$	$1·60 \times 10^{10}$	$1·98 \times 10^{10}$	$1·51 \times 10^{10}$
MAD	$8·5 \times 10^4$	$12·2 \times 10^4$	$18·6 \times 10^4$	$8·22 \times 10^4$	$9·10 \times 10^4$	$8·44 \times 10^4$
R^2	0·57	0·21	0·45	0·54	0·43	0·52

time series. Based purely on split sample forecasting performance, an AR(2) model and the DRM with NAO as predictor appear to be best. However, the AR(2) model is contra-indicated by the stationarity test and also by the non-zero mean residual structure. In this case, the nonstationary ARIMA models improve on the residual attributes but at the expense of a significant degradation of the predictive variance. The Bayesian models, on the other hand, seem to offer a better compromise with respect to these attributes. Conceptually, the 'constant model' is a close approximation to the AR(2) behaviour in that if the lag 1 correlation is 0·95, then the random walk dynamics embedded in the observational noise formulation of the 'constant model' is likely to give similar results. However, the mean is allowed to change, and the changes in the mean are constrained by the inferred posterior variance of the mean state in the Bayesian procedure. Thus, the model reflects a trade-off between the nonstationary and low order linear dynamics that corresponds to an AR(1) sort of model. On the other hand, the DRM extends the observation equation and dynamics equations to allow the changing mean to depend on an exogenous predictor. In this case, the predictor was selected based on some prior investigations and literature that suggest that the NAO may influence winter/spring storm tracks into the region. It may also have some influence on the summer storms through its influence on the Bermuda High region in the North Atlantic. The idea is that if the predictor is useful, the associated regression coefficients may have a 'tight' posterior distribution that does not intersect with zero over a significant area. This was the case in this application. Moreover, the ability to quantify the uncertainty associated with the prediction is another valuable contribution. Thus, the observed overall performance, together with the improvement in the residual attributes, suggests that Bayesian dynamic modelling approach is superior to that of ARIMA models.

APPENDIX A: ONE-STEP AHEAD FORECAST DISTRIBUTION AND UPDATING OF PRIOR DISTRIBUTION TO POSTERIOR DISTRIBUTION USING BAYES RULE

This discussion illustrates the use of Bayes theorem to update the prior distribution to posterior distribution and one-step-ahead forecast distribution.

$$\text{Observation equation}: \ Y_t = \mu_t + v_t$$
$$v_t \sim N[0, V] \quad \text{(A-1)}$$

$$\text{and the system equation}: \ \mu_t = \mu_{t-1} + \omega_t$$
$$\omega_t \sim N[0, W] \text{(A-2)}$$

Let the posterior distribution for μ_{t-1} at time step $(t-1)$ be $P(\mu_{t-1}/D_{t-1}) \sim N[m_{t-1}, C_{t-1}]$ with some mean m_{t-1} and variance C_{t-1}. Thus from Equation (A-2), μ_t is the summation of two normally distributed random variables: $N[m_{t-1}, C_{t-1}]$ and $N[0, W]$. Assuming that the error term is independent of the level of the process, the summation will be another normally distributed random variable with mean m_{t-1} and variance $(C_{t-1} + W)$. Thus, the prior distribution of μ_t for the time step t is $P(\mu_t/D_{t-1}) \sim N[m_{t-1}, R_t]$, where, $R_t = C_{t-1} + W$.

In a similar way, from equation (A-1), Y_t is the summation of $N[m_{t-1}, R_t]$ and $N[0, V]$, which is also normally distributed with mean m_{t-1} and variance $(R_t + V)$. Thus, one-step-ahead forecast distribution is $P(Y_t/D_{t-1}) \sim N[f_{t-1}, Q_t]$, where $f_t = m_{t-1}$ and $Q_t = R_t + V$.

It can be noted that, until now, information up to time step $(t-1)$ is available, which is denoted D_{t-1}. At the end of the time step t, the observed value of Y_t (denoted as y_t), for this time step, is available. Thus, the available information is improved, which is denoted as D_t (consisting of D_{t-1} and y_t). The posterior distribution for μ_t is obtained by Bayes rule. According to the Bayes rule:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior} \quad \text{(A-3)}$$

The likelihood for μ_t is proportional to the observed density of Y_t viewed as a function of μ_t, i.e. the probability density of Y_t given that μ_t has already occurred, i.e. $P(Y_t/D_{t-1}, \mu_t)$, which is, from observation equation (A-1), normally distributed with mean μ_t and variance V_t.

Thus, from Bayes rule (Equation (A-3)):

$$P(\mu_t/D_t) \propto P(Y_t/D_{t-1}, \mu_t)$$
$$\times P(\mu_t/D_{t-1})$$
$$\Rightarrow P(\mu_t/D_t) \propto \frac{1}{\sqrt{2\pi V}} \exp\left[-\frac{(Y_t - \mu_t)^2}{2V}\right]$$
$$\times \frac{1}{\sqrt{2\pi R_t}} \exp\left[-\frac{(\mu_t - m_{t-1})^2}{2R_t}\right]$$

After some algebraic operations, it can be shown that

$$P(\mu_t/D_t) \propto \frac{1}{\sqrt{2\pi\left(\frac{R_t V}{R_t + V}\right)}}$$
$$\times \exp\left[-\frac{\left(\mu_t - \frac{Y_t R_t + m_{t-1} V}{R_t + V}\right)^2}{2\left(\frac{R_t V}{R_t + V}\right)}\right]$$

i.e. $P(\mu_t/D_t)$ is normally distributed with mean

$$\frac{Y_t R_t + m_{t-1} V}{R_t + V} \ \text{and variance} \ \frac{R_t V}{R_t + V}$$

As denoted earlier, $R_t + V = Q_t$ and let $\frac{R_t}{Q_t} = A_t$. So variance, $\frac{R_t V}{R_t + V} = \frac{R_t V}{Q_t} = A_t V$ and mean

$$\frac{Y_t R_t + m_{t-1} V}{R_t + V} = \frac{Y_t R_t + m_{t-1} V}{Q_t}$$

$$= A_t Y_t + \frac{m_{t-1}(Q_t - R_t)}{Q_t}$$

$$= A_t Y_t + m_{t-1} - A_t m_{t-1} = m_{t-1}$$

$$+ A_t(Y_t - m_{t-1}) = m_{t-1} + A_t e_t$$

Thus finally, $P(\mu_t/D_t) \sim N[m_t, C_t]$, where $m_t = m_{t-1} + A_t e_t$ and $C_t = A_t V$.

For DRM, the observation variance is considered to be unknown. Thus a precision parameter $\phi \ (= V^{-1})$ is introduced, which is Gamma distributed. The inverse of the expectation of this distribution is considered to be the estimate of the variance, which is $S_t = d_t/n_t$ (Equation (20)). This is multiplied by the factor $\frac{R_t}{V_t}$ to obtain C_t, i.e. $C_t = R_t S_t/Q_t$ (Equation (19)). Equations to update the precision parameter are also obtained by Bayes principle following a similar procedure.

APPENDIX B: STATISTICAL TEST FOR ANALYSIS OF RESIDUALS FROM ARIMA MODELS

To validate a model, three necessary conditions should be satisfied by the series of residuals obtained from the model. (a) The residual series $\{e_t\}$ should have zero mean. (b) No significant periodicity or other temporal structure (e.g. autocorrelation) should be present. (c) The residual series should be uncorrelated. Statistical tests to ensure the above requirements are explained here. Details of these tests can be found in standard texts and Mujumdar and Nagesh Kumar (1990).

Test for zero mean

To test whether the residual series has zero mean or not, the statistic $T(e_t)$ is calculated using

$$T(e_t) = \frac{N^{1/2}\overline{e}}{\hat{\rho}^{1/2}} \tag{B1-1}$$

where $\overline{e}$ is the estimated residual mean, $\hat{\rho}$ is the estimated residual variance and N is the number of data points. The statistic $T(e_t)$ approximately follows Student distribution $[\sim T_\alpha(N - 1)]$, where α is the significance level. If $T(e_t) \leq t_{\alpha/2}(N - 1)$ or $T(e_t) \geq t_{1-\alpha/2}(N - 1)$, then the mean is considered not to significantly differ from zero.

Test for periodicities

To test whether the residual series has any significant periodicities or not, the statistic $F(e_t)$ is calculated using

$$F(e_t) = \frac{\gamma^2(N - 2)}{4\hat{\rho}_1} \tag{B2-1}$$

where,

$$\gamma^2 = \hat{\alpha}^2 + \hat{\beta}^2 \tag{B2-2}$$

$$\hat{\rho}_1 = \frac{\sum_{t=1}^{N}[e_t - \hat{\alpha}\cos(W_1 t) - \hat{\beta}\sin(W_1 t)]^2}{N} \tag{B2-3}$$

$$\hat{\alpha} = \frac{2\sum_{t=1}^{N} e_t \cos(W_1 t)}{N} \tag{B2-4}$$

$$\hat{\beta} = \frac{2\sum_{t=1}^{N} e_t \sin(W_1 t)}{N} \tag{B2-5}$$

where $\frac{2\pi}{W_1}$ is the periodicity for which the test is carried out and N is the number of data points. The statistic $F(e_t)$ approximately follows F-distribution $[\sim F_\alpha(2, N - 2)]$, where α is the significance level. If $F(e_t) \leq f_\alpha(2, N - 2)$, then the periodicity corresponding to W_1 is considered not significant.

Test for absence of correlation (Portmanteau test)

To test whether the residual series is a white noise or not the Portmanteau test is carried out. According to this test the statistic $\chi(e_t)$ is calculated using

$$\chi(e_t) = (N - n_1)\sum_{k=1}^{n_1}\left(\frac{R_k}{R_0}\right)^2 \tag{B3-1}$$

where,

$$R_k = \frac{\sum_{j=k+1}^{N}[e_j e_{j-k}]}{N - k} \tag{B3-2}$$

where N is the number of data points and n_1 is normally assumed to be $0 \cdot 15N$. The statistic $\chi(e_t)$ approximately follows chi-square distribution $[\sim X_\alpha^2(n_1)]$, where α is the significance level. If $\chi(e_t) \leq \chi_\alpha^2(n_1)$, then the residual series is considered to be a white noise.

However, Whittle's test (Whittle, 1952) can also be used to test the absence of correlation in residual series.

REFERENCES

Agrawal R, Faloutsos C, Swami A. 1993. Efficient similarity search in sequence databases. In *Foundations of Data Organization and Algorithms: 4th International Conference, FODO '93 Chicago, Illinois, USA, October 13-15, 1993 Proceedings*, Lomet DB (ed.). *Lecture Notes in Computer Science*, No. 730. Springer Verlag: Berlin; 69–84.

Baldwin CK, Lall U. 2000. *Dramatic fluctuations of Devil's Lake, North Dakota: climate connections and forecasts*. Utah Water Research Laboratory, Utah State University, Logan, UT 84322-8200, US Army Corps of Engineers, Institute of Water Resources Project Completion Report; 41.

Berliner LM, Wikle CK, Cressie N. 2000. Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *Journal of Climate* **13**(22): 3953–3968.

Berger J, Insua DR. 1998. Recent developments in Bayesian inference with applications in hydrology. In *Statistical and Bayesian Methods in*

Hydrological Sciences, E. Parent, *et al* (eds). UNESCO Press: Paris; 43–62.

Bergman MJ, Delleur JW. 1985. Kalman filter estimation and prediction of daily stream flows: I. Review, algorithm, and simulation experiments. *Water Resources Bulletin WARBAQ* **21**(5): 815–826.

Bernier J. 1994. Statistical detection of changes in geophysical series. In *Engineering Risk in Natural Resources Management with Special References to Hydrosystems Under Changes of Physical or Climatic Environment*, Duckstein L, Parent E (eds). Kluwer Academic Publishers: Netherlands, NATO ASI Series E: Applied Sciences 275: 343–357.

Bierkens MFP, Knotters M, Hoogland T. 2001. Space-time modeling of water table depth using a regionalized time series model and the Kalman filter. *Water Resources Research* **37**(5): 1277–1290.

Box GEP, Jenkins GM, Reinsel GC. 1994. *Time Series Analysis, Forecasting and Control*. Pearson Education: Indian reprint, India.

Cavalieri DJ, Gloersen P, Parkinson CL, Comiso JC, Zwally HJ. 1997. Observed hemispheric asymmetry in global sea ice changes. *Science* **278**: 1104–1106.

Crow WT, Wood EF. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. *Advances in Water Resources* **26**(2): 137–149.

Dai A, Fung IY, Del Genio AD. 1997. Surface observed global land precipitation variations during 1900–88. *Journal of Climate* **10**: 2943–2962.

Dickson RR, Lazier J, Meincke J, Rhines P, Swift J. 1996. Long-term co-ordinated changes in the convective activity of the North Atlantic. *Progress in Oceanograpphy* **38**: 241–295.

Houghton RW. 1996. Subsurface quasi-decadal fluctuations in the North Atlantic. *Journal of Climate* **9**: 1363–1373.

Hurrell JW. 1995a. Transient eddy forcing of the rotational flow during northern winter. *Journal of Atmospheric Science* **52**: 2286–2301.

Hurrell JW. 1995b. Decadal trends in the North Atlantic Oscillation regional temperatures and precipitation. *Science* **269**: 676–679.

Hurrell JW. 1996. Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperatures. *Geophysical Research Letters* **23**: 665–668.

Hurrell JW. 2000. Climate variability: North Atlantic and Arctic Oscillation (NAO/AO). In *Encyclopedia of Atmospheric Sciences*, Holton J, Pyle J, Curry J (eds). Academic Press: 439–445.

Hurrell JW, Loon HV. 1997. Decadal variations in climate associated with the North Atlantic oscillation. *Climatic Change* **36**: 301–326.

Krishnaswamy J, Halpin PN, Richter DD, Korfmacher K. 2001. Dynamics of sediment discharge in relation to land-use and hydro-climatology in a humid tropical watershed in Costa Rica. *Journal of Hydrology* **253**: 91–109.

Krishnaswamy J, Lavine M, Richter DD, Korfmacher K. 2000. Dynamic modeling of long-term sedimentation in the Yadkin River basin. *Advances in Water Resources* **23**: 881–892.

Ljung L. 1999. *System Identification—Theory for the User* 2nd edn. Prentice Hall: Upper Saddle River, NJ.

Ma WY, Manjunath BS. 1996. Texture features and learning similarity. In *Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition (CVPR '96)*. IEEE Computer Society: Washington, DC; 425–430.

Maity R, Nagesh Kumar D. 2006. Bayesian dynamic modeling for monthly Indian summer monsoon rainfall using ENSO and EQUINOO. *Journal of Geophysical Research* **111**: D07104. DOI: 10·1029/2005JD006539.

Maity R, Nagesh Kumar D. 2007. Hydroclimatic teleconnection between global sea surface temperature and rainfall over India at subdivisional monthly scale. *Hydrological Processes*. **21**(14): 1802–1813. DOI: 10·1002/hyp.6300.

Makridakis S, Wheelwright SC, Hyndman RJ. 1998. *Forecasting Methods and Applications*, 3rd edn. John Wiley & Sons: New York.

Maslanik JA, Serreze MC, Barry R G. 1996. Recent decreases in Arctic summer ice cover and linkages to atmospheric circulation anomalies. *Geophysical Research Letters* **23**: 1677–1680.

Mujumdar PP, Nagesh Kumar D. 1990. Stochastic models of streamflow: some case studies. *Hydrological Sciences–Journal–des Sciences Hydrologiques* **35**: 395–410.

Papoulis A, Pillai SU. 2002. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill: New York.

Park S, Lee D, Chu WW. 1999. Fast retrieval of similar subsequences in long sequence databases. In *3rd IEEE Knowledge and Data Engineering Exchange Workshop*, Chicago, USA.

Pole A, West M, Harrison J. 1994. *Applied Bayesian Forecasting and Time Series Analysis*. Chapman & Hall: New York.

Schreider SY, Young PC, Jakeman AJ. 2001. An application of the Kalman filtering technique for streamflow forecasting in the upper Murray basin. *Mathematical and Computer Modelling* **33**: 733–743.

Shabbar A, Higuchi K, Skinner W, Knox JL. 1997. The association between the BWA index and winter surface temperature variability over eastern Canada and west Greenland. *International Journal of Climatology* **17**: 1195–1210.

Sigurdsson O, Jonsson T. 1995. Relation of glacier variations to climate changes in Iceland. *Annals of Glaciology* **21**: 263–270.

Wallace JM, Zhang Y, Renwick JA. 1995. Dynamic contribution to hemispheric mean temperature trends. *Science* **270**: 780–783.

Walsh JE, Chapman WL, Shy TL. 1996. Recent decrease of sea level pressure in the central Arctic. *Journal of Climate* **9**: 480–486.

Watanabe M, Nitta T. 1999. Decadal changes in the atmospheric circulation and associated surface climate variations in the Northern Hemisphere winter. *Journal of Climate* **12**: 494–510.

West M, Harrison J. 1997. *Bayesian Forecasting and Dynamic Models*. Spinger-Varlag New York.

Whittle P. 1952. Tests for fit in time series, *Biometrika* **39**: 309–318.