Local polynomial method for ensemble forecast of time series

S. K. Regonda[1,2], **B. Rajagopalan**[1,2], **U. Lall**[3], **M. Clark**[2], and **Y.-I. Moon**[4]

[1]Dept. of Civil, Env. and Arch. Eng., University of Colorado, Boulder, CO, USA
[2]CIRES, University of Colorado, Boulder, CO, USA
[3]Dept. of Earth and Env. Eng., Columbia University, New York, NY, USA
[4]Dept. of Civil Engg., University of Seoul, Seoul, Korea

Abstract. We present a nonparametric approach based on local polynomial regression for ensemble forecast of time series. The state space is first reconstructed by embedding the univariate time series of the response variable in a space of dimension (D) with a delay time (τ). To obtain a forecast from a given time point t, three steps are involved: (i) the current state of the system is mapped on to the state space, known as the feature vector, (ii) a small number ($K = \alpha * n$, α=fraction (0,1] of the data, n=data length) of neighbors (and their future evolution) to the feature vector are identified in the state space, and (iii) a polynomial of order p is fitted to the identified neighbors, which is then used for prediction. A suite of parameter combinations (D, τ, α, p) is selected based on an objective criterion, called the Generalized Cross Validation (GCV). All of the selected parameter combinations are then used to issue a T-step iterated forecast starting from the current time t, thus generating an ensemble forecast which can be used to obtain the forecast probability density function (PDF). The ensemble approach improves upon the traditional method of providing a single mean forecast by providing the forecast uncertainty. Further, for short noisy data it can provide better forecasts. We demonstrate the utility of this approach on two synthetic (Henon and Lorenz attractors) and two real data sets (Great Salt Lake bi-weekly volume and NINO3 index). This framework can also be used to forecast a vector of response variables based on a vector of predictors.

1 Introduction

It has been always intriguing to forecast various natural and physical processes (e.g. rainfall, runoff, lake volumes, etc.), which appear as a result of coupling of different components of the Earth system. Theoretically, it is possible to forecast many of these natural processes using first principles given

their basic mechanisms, but, unfortunately, it is not practically possible, since it is associated with the following problems: a) measurements taken at discrete locations of space and time, b) initial conditions are associated with significant uncertainty, c) many of the natural processes are dynamic, nonlinear, and extremely complex, and are affected by several interconnected physical variables, d) heterogeneity of the model variables on space and time scales, and e) noisy and finite amount of real data. Additionally, another major concern is the unavailability of an appropriate mathematical model. These problems, and the fact that the dependent and independent variables are coupled, limit the forecast success for many variables via the development of physical models. Consequently, time series methods that model the response variable statistically have become increasingly popular.

The theory of traditional statistical forecasting views the time series of a response variable as a realization of a random process (e.g. Pandit and Yu, 1983), and it is appropriate only if effective randomness arises from complicated motion involving many independent and irreducible degrees of freedom (Farmer and Sidorowich, 1987). As a result, their predictability is limited. An alternative cause of randomness is deterministic chaos, which can occur even in very simple deterministic systems. It suggests that complex and unpredictable processes are not necessarily of high degrees of freedom but might be result of low dimensional dynamical systems. In a simple way, deterministic chaos or chaotic systems appear as random processes but internally have a definite relationship among the variables. Though chaotic systems result from deterministic phenomenon, have fundamental limit on long-term forecasting because its future evolutions are unstable. Nonlinear time series methods exploit this recognition and reconstruct the dynamics of the system, thereby developing the potential for short and long-lead forecast.

Much research using geophysical time series (i.e. precipitation, streamflow, etc.) demonstrated the existence of lower order chaotic behavior on different scales (e.g. Rodriguez-Iturbe et al., 1989; Sharifi et al., 1990; Jayawardena and

Lai, 1994; Porporato and Ridolfi, 1996, 1997; Sangoyomi et al., 1996; Sivakumar et al., 1998, 1999a, 2002; Wang and Gan, 1998; Krasovskaia et al., 1999; Stehlik, 1999, 2000; Elshorbagy et al., 2002; Regonda et al., 2004), but there have been limited efforts in using this information for short term forecasts. Notable among them are the attempts at forecasting the Great Salt Lake (GSL) bi-weekly volumes (Abarbanel and Lall, 1996; Abarbanel et al., 1996; Lall et al., 1996), daily rainfall (e.g. Jayawardena and Lai, 1994), flood (Laio et al., 2003), and streamflow on daily and monthly scales (e.g. Jayawardena and Lai, 1994; Porporato and Ridolfi, 1996, 1997, 2001; Liu et al., 1998; Sivakumar et al., 2001, 2002). However, in other fields, especially, medical, there are widespread forecast applications (e.g. Kantz and Schreiber, 1997, 1998; Richter and Schreiber, 1998).

A key assumption behind nonlinear time series methods is that even if the exact mathematical description of the dynamical system is not known, the state space (or phase space) can be reconstructed from a single observed time series (Packard et al., 1980). The state space is defined as a multidimensional space in which axes correspond to variables of a dynamical system. The state space is reconstructed by embedding the univariate time series of the response variable with a delay time τ in a D-dimensional space (this will be described in the following section). The dimension D can be thought of as the minimum number of state variables required to describe the system and delay time τ is the average length of memory in the system. Two popular methods for estimating the embedding dimension (D) in practice are: the Grassberger-Procaccia (Grassberger and Procaccia, 1983a, b) approach (GPA), which estimates the dimension mostly as "fractal" or "non-integer", and the False Nearest Neighbor (FNN) method (Kennel et al., 1992), which computes the integer dimension. The delay time (τ) is estimated using the Mutual Information approach (Fraser and Swinney, 1986; Moon et al., 1995). Once the state space is reconstructed, the forecasting proceeds as follows: (i) the current state of the system, say, X_t, at the current time step t is identified in the state space, (ii) a small number (say K) of nearest neighbors of current state, say $\mathbf{X_K}$, and their corresponding successors, $\mathbf{X_{K+1}}$, are identified in the state space, and (iii) a local function is fitted to the identified neighbors, $X_{K+1} = f(\mathbf{X_K})$, and (iv) the fitted function is then used to estimate the forecast. The function $f(.)$ can be a simple weighted average (Farmer and Sidorowich, 1987) or it can be a higher order polynomial ($p \geq 1$). The parameters (D, τ, α, p) and the form of the function $f(.)$ are typically fixed. This works well if the estimated parameters D and τ capture the dynamics accurately in the embedded space. In such cases, the forecasts from this approach will handily outperform linear time series methods, such as Auto Regressive (AR) models (Casdagli et al., 1990; Grassberger et al., 1991; Tsonis, 1992; Jayawardena and Lai, 1994; Jayawardena and Gurung, 2000; Lisi and Villi, 2001).

Real data sets, especially geophysical time series, are short and noisy (due to instrumental and dynamical errors). Consequently, the estimates of the parameters are not very reliable, thereby significantly impacting the forecast skills

(Schreiber and Kantz, 1996). Smoothing can reduce the noise (Schreiber and Grassberger, 1991; Porporato and Ridolfi, 1996, 1997), but then if not done properly it can alter the underlying dynamics (Sivakumar et al., 1999b). To address this, it is intuitive that an ensemble of forecasts has to be generated from a suite of plausible parameter combinations that vary within the state space. This ensemble approach has several advantages over issuing a single mean forecast as described above. They are that the ensembles provide: (i) a natural estimate of the forecast uncertainty, and (ii) the probability density function (PDF) of the response variable and, consequently, threshold exceedance probabilities which can be very useful in decision making.

The goal of this paper is to present a new method that provides a suite of model parameters and, consequently, enables ensemble forecast of time series. The methodology is first described. We then demonstrate its utility by applying it to two synthetic data sets from Henon and Lorenz attractors and two real data sets, the Great Salt Lake bi-weekly volumes and the NINO3 (an index of El Niño Southern Oscillation, ENSO).

2   Methodology

We provide only a limited description of reconstruction of state space of a dynamical system from a scalar time series of one of the state variables, and refer the reader to Abarbanel et al. (1993), Abarbanel and Lall (1996), and Lall et al. (1996) for background information and details.

A method for state space reconstruction was originally proposed by Packard et al. (1980), and put on a firm mathematical basis by Takens (1981). The dynamics of a time series $\{x_1, x_2, \ldots, x_n\}$ are fully captured or embedded in the D-dimensional phase space defined by $X_t = \{x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\}$. According to the embedding theorem of Takens (1981), to characterize a dynamic system with an attractor dimension d, $D=2d+1$-dimensional phase space is sufficient, whereas Abarbanel et al. (1993) suggested that, in practice, $D>d$ may be adequate.

The False Nearest Neighbor (FNN) method proposed by Kennel et al. (1992) is used in this study to estimate the embedding dimension of the phase space. Other methods such as the correlation dimension method (Grassberger and Procaccia, 1983a, 1983b) can also be used but the FNN method is much more robust (Abarbanel and Lall, 1996). FNN method estimates the minimum embedding dimension of the phase space by computing the number of false neighbors in each embedded phase space. The delay time τ needs to be appropriately chosen. If the value of τ is less than the appropriate value then the data values will not be independent, resulting in a loss of information and characteristics on the attractor structure. If τ is too large, i.e. much larger than the information decay time, then there is no dynamical correlation between the state vectors and it causes a loss of information on the original system (Jayawardena and Lai, 1994). The choice of τ is usually made with the help of autocorrelation

function or mutual information content (Fraser and Swinney, 1986; Jayawardena and Lai, 1994; Moon et al., 1995). Limitations of these methods are discussed in Abarbanel and Lall (1996) and Porporato and Ridolfi (1996, 1997).

Using a selected value of the embedding dimension and delay time, the state space is reconstructed (as described above) and the forecast for T time step into the future is given as

$$X_{t+T} = f(X_t) + \varepsilon_t. \tag{1}$$

X_t is the feature vector at time t and $f(.)$ is a function (linear or nonlinear) that maps current state to a future point in the state space. Typically, the function $f(.)$ is estimated locally within a neighborhood of the feature vector in the state space (Farmer and Sidorowich, 1987). We use locally weighted polynomials, also defined as LOCFIT (Loader, 1997) to estimate the function. In this, a polynomial of order p is fitted to K ($K=\alpha*n$, $\alpha=(0, 1]$) nearest neighbors of the feature vector in the state space. Clearly, if $K=n$ (or $\alpha=1$) and $p=1$, then it collapses onto a traditional multiple linear regression. Since the function is estimated locally, it has the flexibility to approximate any differentiable functional form appropriate for the data. For theoretical details of this approach, see Loader (1999). This approach has been widely applied with good results in ensemble streamflow forecast (Grantz, 2003), flood frequency analysis (Apipattanavis et al., 2004[1]), ensemble streamflow simulation (Prairie, 2002), and in spatial interpolation of hydroclimatic fields (Rajagopalan and Lall, 1998; Owosina, 1992).

General Cross Validation (GCV) is used to guide the selection of optimal parameters, α and p. The optimal combination of parameters is, typically, the one that produces a minimum GCV. The GCV (α, p) score function is defined as:

$$\text{GCV}(\alpha, p) = \frac{\sum_{i=1}^{n} \frac{e_i^2}{n}}{\left(1 - \frac{m}{n}\right)^2}, \tag{2}$$

where e_i is the error, n is the number of data points, m is the number of parameters.

Because of limited sample size and noise (measurement noise and internal noise) in the data, there is always uncertainty in the computed embedding dimension of real systems and in estimated parameters. Hence, a single estimate of embedding dimension and delay time might not adequately represent the dynamics of the system, especially for forecasting. Therefore, it is intuitive to obtain a suite of plausible parameters of the state space. To this end, we propose the selection of all the parameters (D, τ, α, p) using the GCV score function. Thus, the function in Eq. (2) would be GCV(D, τ, α, p).

---

[1] Apipattanavis, S., Rajagopalan, B., and Lall, U.: Local polynomial technique for flood frequency analysis, J. Hydro. Eng., in review, 2004.

2.1 Forecast algorithm

The algorithm for implementation is as follows:

1. Compute the embedding dimension and delay time using the standard methods (e.g. FNN and MI). Using this as reference, choose a suitable range of values of D and τ.

2. Reconstruct the phase space for a selected D and τ combination.

3. Calculate the GCV for the reconstructed phase space by varying the smoothing parameter of local polynomial function (i.e. neighborhood size) and the order of local polynomial.

4. Repeat steps 2 and 3 for all combinations of D, τ, α and p.

5. Select a suite of "best" parameter combinations that is within 5% of the lowest GCV.

6. Each selected "best" combination is then used to generate a forecast.

7. The suite of "best" combination forecast then provides an ensemble of conditional mean forecast that reflect parameter/model selection uncertainty.

3 Description of test cases

As mentioned earlier, we applied the ensemble forecast methodology on synthetic data from Henon and Lorenz attractors and on two geophysical time series (Great Salt Lake bi-weekly volume and NINO3 index). The data sets are described below. In all the cases, the model is fitted (i.e. suite of parameters obtained) on a subset of the data and blind predictions (no data outside the fitting subset is used for predictions) are made for T time steps ahead. The forecasts are made from several starting points to show the change in the predictability – one of the aspects of nonlinear systems.

3.1 The Henon system

The Henon system is described by the following coupled equations (Henon, 1976):

$$\begin{aligned} x_{t+1} &= 1 - a*x_t^2 + y_t \\ y_{t+1} &= b*x_t \end{aligned}.$$

Values of $a=1.4$, $b=0.3$, and initial observation (x_0, y_0) as (0, 0) are chosen to generate a time series of 4000 observations using the above equations. Figure 1a shows the times series of the x variable of the Henon system.
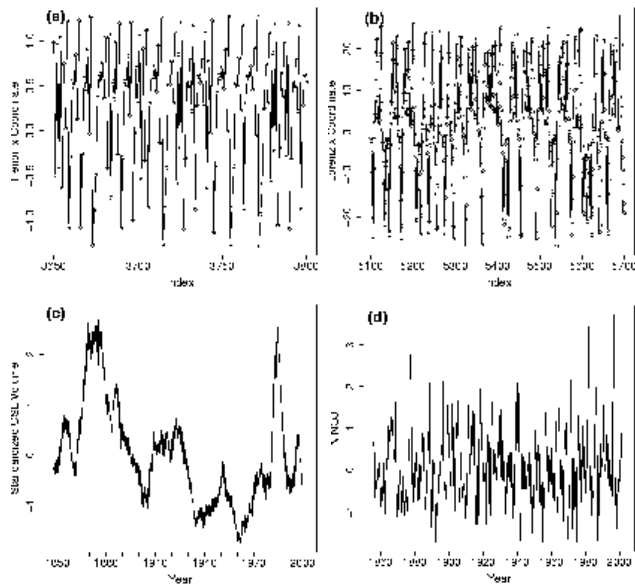
Fig. 1. Time series of the synthetic and real data sets: **(a)** Henon *x* coordinate, **(b)** Lorenz *x* coordinate, **(c)** Standardized bi-weekly volumes of Great Salt Lake, and **(d)** monthly NINO3 index.

3.2   The Lorenz system

The Lorenz System is described by the following equations (Lorenz, 1963):

$$\dot{x} = -\sigma(x + y)$$
$$\dot{y} = -x * z + r * x - y \,.$$
$$\dot{z} = x * y - b * z$$

Here we take $\sigma=16$, $r=45.92$, $b=4$, $\Delta t=0.05$ and initial observation (x_0, y_0, z_0) as $(1, 0, 0)$ to generate a time series of 6000 observations. Figure 1b shows the times series of the *x* variable of the Lorenz system.

3.3   The Great Salt Lake time series

The Great Salt Lake of Utah is located at approximately 40° to 42° N and 110° to 112° W and is the fourth largest, perennial, closed basin, saline lake in the world. The GSL drains an area of $90\,000\,\text{km}^2$ and its water level has been recorded since 1847 at 15 days (bi-weekly) interval. Sangoyomi (1993) compiled biweekly data sets of GSL time series and reported variability on annual, interannual, and, interdecadal scales. The time series is shown in Fig. 1c. Of particular interest is the fall and rise of the lake in the mid 1920s and 1980s, which had tremendous implications to the regional economy (James et al., 1979). Large scale ocean-atmospheric features over the Pacific Ocean have been shown to modulate the interannual and interdecadal variability of the Lake volumes (Mann et al., 1995; Lall and Mann, 1995; Moon, 1996; Moon and Lall, 1996) and fluctuations in the GSL are related with patterns in regional precipitation, temperature, and streamflow over quasi-periodic interannual and interdecadal scales (Sangoyomi, 1993; Mann et al., 1995;

Abarbanel et al., 1996; Lall et al., 1996; Moon and Lall, 1996; Sangoyomi et al., 1996).

Dynamical characteristics of the GSL, stochastic or deterministic chaos, dimensionality and predictability are described in Sangoyomi et al. (1996), which indicate the GSL to be a low dimensional chaotic system (i.e. dominated by a few degrees of freedom). Being a closed lake of arid region and having a large surface area, long-term average evaporation rate exceeds the average precipitation; and it integrates the basin hydrologic response by filtering out the noisy processes into a few dominant processes, thus supporting the low dimensional characteristics. Here, we forecast the lake volumes from several starting points of fall (1925–1930) and rise (1983–1987) of the GSL.

3.4   NINO3

NINO3 is a widely used index of the tropical Pacific ocean-atmospheric phenomenon, ENSO. It is a time series of averaged monthly Sea Surface Temperature (SST) anomalies in the tropical Pacific covering the domain of 4° N–4° S and 90°–150° W. This time series starts from 1856 and is shown in Fig. 1d. Details of data prior to the modern observational period can be found in Kaplan et al. (1998). ENSO has significant implications to global climate (e.g. Ropelewski and Halpert, 1986) and, consequently, to the global socio-economy. Hence, understanding its dynamics and predictability is crucial to improve seasonal climate forecast. There are several approaches to ENSO forecasting – statistical (e.g. Barnett et al., 1988; Barnston and Ropelewski, 1992; Balmaseda et al., 1994; Latif et al., 1994; Xue et al., 1994; Mason and Mimmack, 2002) and dynamical (e.g. Cane et al., 1986; Zebiak and Cane, 1987; Barnett et al., 1993; Balmaseda et al., 1994; Latif et al., 1994; Chen et al., 1995, 1997, 2004; Tziperman et al., 1995; Xue et al., 1997). Zebiak and Cane (1987) were the first to develop a simple model for ENSO and successfully predicted the El Niño of the 1986–1987.

Several researchers explored the ENSO dynamics and explained some of its attributes, e.g. onset, termination, and cyclic nature to seasonal cycle (see references in Tziperman et al., 1994, 1995), whereas attributes irregularity and partial locking to seasonal cycle are unexplained because of its complexity (Rasmusson and Carpenter, 1982). Lower order chaotic dynamics (e.g. Vallis, 1986, 1988; Munnich et al., 1991; Chang et al., 1994; Tziperman et al., 1994, 1995) was shown to be a plausible reason for irregularity. In fact, Tziperman et al. (1994, 1995) and Jin et al. (1994) proposed a model that presumed ENSO dynamics as low order chaotic and explained all the attributes of ENSO dynamics. Tziperman et al. (1994) suggested ENSO as a chaotic system with few (i.e. ≤ 9) dominated degrees of freedom.
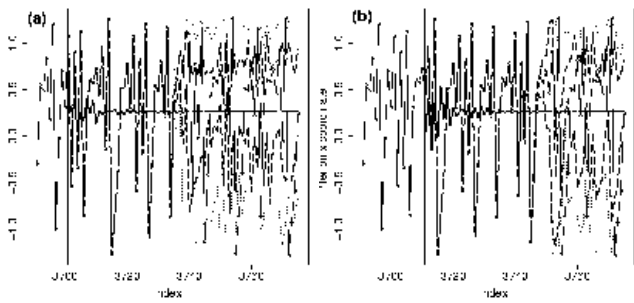
Fig. 2. Ensemble forecasts of Henon x coordinate **(a)** Blind prediction starts from index 3701, and **(b)** Blind prediction starts from 3711. The inter quartile range of the ensembles are shown as dashed lines, the 5th and 95th percentiles are as dotted lines, the true values as points connected by solid lines, and the best AR forecast shown as solid line.

4 Results

As mentioned in the previous section, the ensemble forecasts are generated using the methodology developed earlier in the paper. For all the cases, the inter quartile range of the ensembles are shown as dashed lines, the 5th and 95th percentiles are as dotted lines, and the true values shown as points connected by solid lines. Forecasts from the best AR models that use Akaike Information Criterion (AIC) (Brockwell and Davis, 2002; Chatfield, 2003), are also shown as solid lines for comparison with traditional linear time series methods.

4.1 The Henon system

This system has two variables (i.e. x and y), and so the true embedding dimension is 2; the estimates from FNN and GPA methods also confirm this. Plotting x [i] vs. x [$i+1$] unfolds the attractor in a two-dimensional phase space (not shown), suggesting the phase space dimension and delay time as 2 and 1, respectively (Henon, 1976; Kennel et al., 1992). Considering these values as reference, we search over $D=1$ to 5 and $\tau=1$ to 10 for the best combination of state space parameters. We compute GCV for each of these combinations by varying neighbor size (α) and degree of local polynomial fit (p). The "best" combinations with GCV values within the 5% range of the lowest value are selected, resulting in 15 combinations. These combinations are used to generate forecasts for 100 time steps into the future. Interestingly, all the selected combinations exhibit the same parameter values of $D=2$, $\tau=1$ and $p=2$ but with various neighborhood sizes (i.e. α). Note that the D and τ values are equal to the true values of the system.

Blind predictions start from index 3701. Each of the selected combinations is used to obtain a forecast for 100 steps into the future. Forecasts from the best AR model fitted to the data are also generated. The forecasts are shown in Fig. 2. Notice that the ensemble forecasts predict extremely well the first 40 to 50 time steps, that they are indistinguishable from the true values (Fig. 2a). Subsequently, the forecast trajectories start to diverge, i.e. the inter quartile range starts to
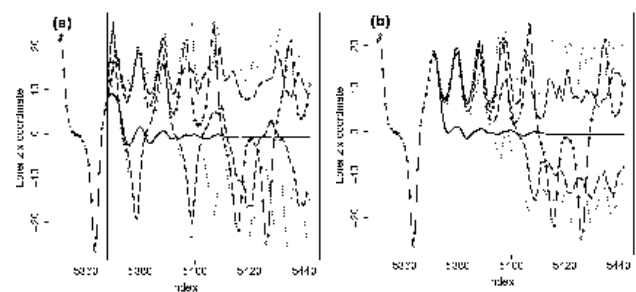

Fig. 3. Same as Fig. 2 but for Lorenz x coordinate, blind prediction starts from **(a)** index 5368, and **(b)** index 5371.

expand. Unlike this, the best AR forecast is barely able to predict correctly the first $\sim$5 to 10 time steps and, subsequently, stays flat at the mean value. Similar observations can be seen with forecasts starting from index 3711 shown in Fig. 2b.

4.2 The Lorenz system

For the Lorenz system, there are three variables and, hence, the embedding dimension is 3. GPA and FNN methods indicate a dimension of 2.06 and 3, respectively. GCV scores are computed for the same ranges of D and τ as with the Henon system. Here too, 15 "best" combinations have their GCV values within 5% of the least GCV combination. These combinations have $D=2$ and 3, $\tau=1$ and 2, and $p=2$ with various neighbor sizes. Blind forecasts are issued from indices 5368 and 5371. Forecasts starting from index 5368 (Fig. 3a) show a wide inter quartile range indicative of large uncertainty. This is consistent with the fact that the region ($x=0$) is the unstable part of the attractor and, hence, low predictability. On the other hand, forecasts issued from index 5371 (Fig. 3b) which is quite away from $x=0$ has a tighter inter quartile range for the first $\sim$35 points and then starts to expand significantly just when the true value crosses the ($x=0$) region. The best AR predictions barely capture the first 3 points indicating their inadequacy and incapability to provide good forecasts even in the part of the system that has good predictability. Notice that the Lorenz system exhibits less predictability compared to Henon. This is consistent with their Lyapunov exponents (Wolf et al., 1985), i.e. largest Lyapunov exponent of Lorenz is 2.2 bits/s, whereas for Henon it is 0.6 bits/s (larger the Lyapunov exponent lesser the predictability). However, it should be remembered that prediction is highly dependent on the position of the initial point in the phase space, as noted in the above two examples.

4.3 The Great Salt Lake time series

We present blind predictions (i.e. no data outside the fitting subset is used for prediction) for two cases: (i) the fall of the lake volume (during 1925–1930), and (ii) the dramatic rise and fall (during 1983–1987). The GPA and FNN methods suggest an embedding dimension of 4 and the mutual information (Moon et al., 1995) indicates a delay time of 14. Var-
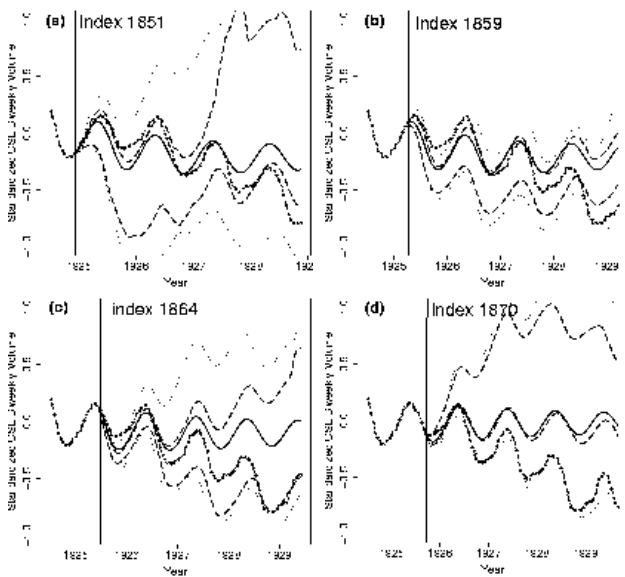
Fig. 4. Same as Fig. 2 but for the standardized bi-weekly volume of GSL. Blind forecasts for four years starting from **(a)** 15 December 1924 (index 1851), **(b)** 15 April 1925 (index 1859), **(c)** 1 July 1925 (index 1864), and **(d)** 15 October 1925 (index 1870).

Fig. 5. Same as Fig. 4, blind forecasts starting from **(a)** 15 October 1983 (index 3264), **(b)** 15 January 1984 (index 3270), **(c)** 15 September 1984 (index 3286), and **(d)** 1 January 1987 (index 3341).

ious sections of the GSL time series yield similar embedding dimension and delay time values. These are consistent with those obtained by Sangoyomi et al. (1996). GCV values are computed over the range of D=2 to 6, τ=10 to 20, and p=1 to 2 with various neighbor sizes. Fifteen combinations with GCV values within 1.2% of the least GCV value are selected.

For the low volume region, we start the blind forecast from December 1, 1924 (index 1851) and also from three different points thereafter, just when the lake was in the process of undergoing a regime transition (see Fig. 1c) from an "average" volume to a "low" volume state. In this case, the fifteen selected combinations have the range of parameters as D=4 and 5, τ=10, 14, and 15, p=1 and 2, and α=0.1–0.5. The inter quartile range of the ensembles is tight and is able to capture the transition to the "low" volume regime (for forecasts starting from 1 March and 1 June 1925, Figs. 4b and 4c) – the best AR model performs well at the beginning but then fails to make the transition to the "low" volume regime. For the other cases, the ensemble forecasts do quite well for the first 1∼2 years and then the inter quartile range tends to expand. Note the sensitivity of the predictability to different starting points (see Figs. 4a, 4b, 4c, and 4d). These results are qualitatively similar to those reported in Lall et al. (1996). The difference is that we are now able to offer some uncertainty bounds on the forecast.

The period 1983–1987 covers the dramatic rise of the GSL and its subsequent decline. Various auto regressive methods fail to blind forecast this event (Lall et al., 1996). The blind forecasts are started from index 3264, which corresponds to October 15, 1983, and from several points along the rising limb. It can be seen that transition from low-volume to average-volume regime has already taken place. The best
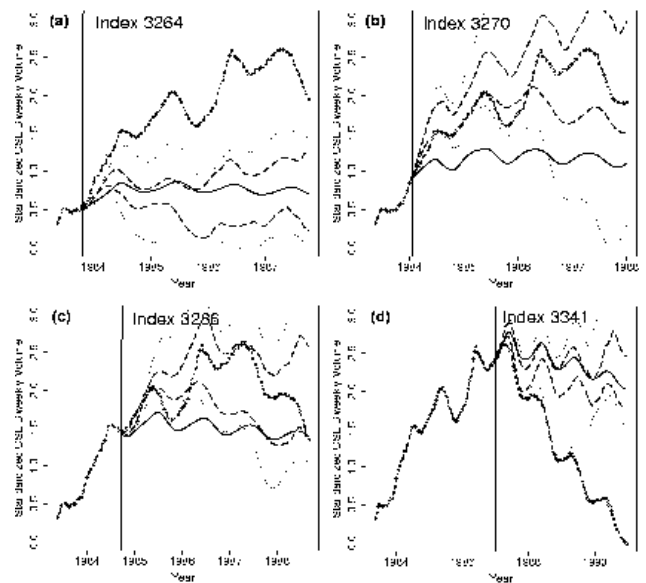
fifteen combinations based on GCV result in the parameter range of D=4 and 5, τ=10 and 14, p=2, and α=0.1–0.4. Polynomial order two (i.e. quadratic) is selected in all the combinations, given the rise of the time series a higher order is intuitive. Ensembles from these combinations for different starting points are shown in Fig. 5. The ensemble forecasts generally do a good job performing particularly well in capturing the rise and fall of the Lake volume when the starting point is slightly along the rising limb (Fig. 5c). The best AR model performs poorly in almost all the situations, regardless of the starting point, and it always tends to the mean (which is 0 in the case of the standardized volume data). Given that the forecasts in both the cases are blind and almost four years into the future, the ensemble forecasts' skill in capturing the PDFs is impressive.

It could be argued that the high GSL volume in the early part (1868–1877) of the record is responsible for better forecasting the high volume period of the mid 1980s (Fig. 5). To test this, data for the 1909 to 1984 period is used to select the model parameters and blind predictions are made for the rise and fall of the lake volume in the mid 1980s. Interestingly, the predictions are quite good (figure not shown), albeit with a decreased lead time in comparison to Fig. 5. This shows that the embedding of the time series is able to capture the underlying dynamics and, consequently, can predict features not observed in the past.

4.4   NINO3

We apply the method to two El Niño events of 1982 and 1997 and two La Niña events of 1984 and 1999. Unlike the GSL time series, the embedding dimension and the delay time computed from the FNN and MI methods yield slightly different values for different lengths of the data. This is some-
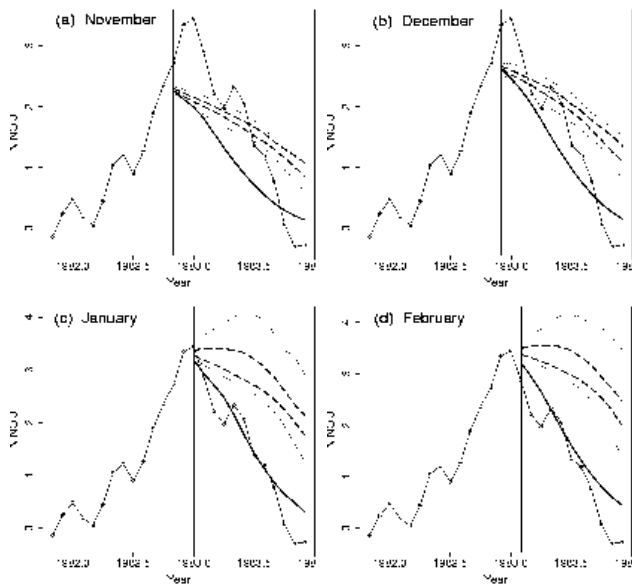
Fig. 6. Same as Fig. 2 but for the NINO3 index. Blind forecast of 1982–1983 El Niño event issued from different starting points **(a)** November 1982, **(b)** December 1982, **(c)** January 1983, and **(d)** February 1983.
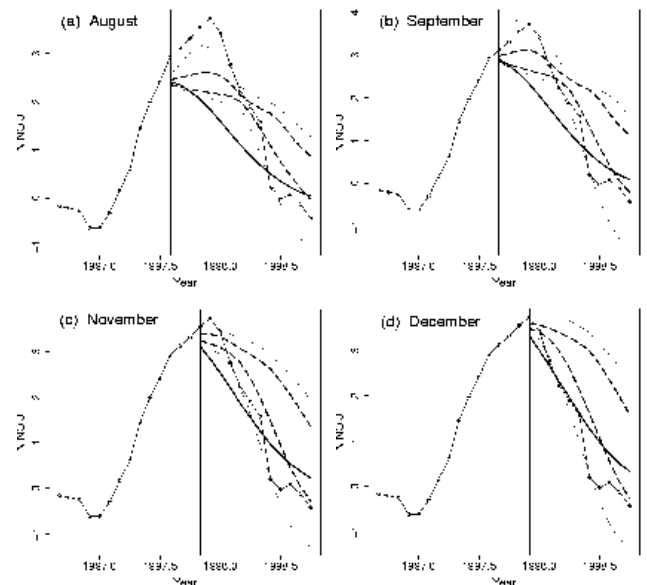

Fig. 7. Same as Fig. 6, but blind forecast of 1997–1998 El Niño event issued from **(a)** August 1997, **(b)** September 1997, **(c)** November 1997, and **(d)** December 1997.

what to be expected, as the NINO3 data is noisier and shorter than the GSL time series.

Two El Niño events (1982–1983 and 1997–1998) are selected for this study. The data spanning 1856–1981 indicate an embedding dimension 4 with a delay of 16. For the period 1856–1996, they are 5 and 13, respectively. The 1982–1983 event selected 600 combinations that have GCV values within the 1.05% of the least GCV with parameters range of D=2 to 5, τ=11 to 21, p=1 and 2, and α=0.1–1.0. Forecasts are made at the start of each month from November 1982 (Fig. 6a). Both methods perform similarly at the start but the ensemble predictions are better for forecasts from December (Fig. 6b). Note that the AR forecast issued from January (Fig. 6c) quickly tends towards the mean (which is 0, in this case), while the ensembles indicate a rise and then a gradual fall. For the forecast issued in February (Fig. 6d), the AR seems to do better, due to the fact that the AR methods tend towards the mean value. The ensembles do a slightly better job but the overall this event is difficult to predict.

For the 1997–1998 event, the combinations have similar parameter ranges as with the 1982–1883 event, except τ ranges from 8 to 16. Figure 7 shows the forecasts issued at the start of different months i.e. August, September (Figs. 7a and 7b), and November, December (Figs. 7c and 7d). Note that in all the cases the ensemble predictions perform much better than the AR-model. In particular, it is able to reproduce the rise and fall starting from September (Fig. 7b).

The data prior to the 1984 and 1999 La Niña events yield a dimension and delay time of 5 and 17, respectively. The "best" combinations have parameters in the range of D=2 to 5, τ=12 to 22, p=1 and 2, and α=0.1–1.0. For the 1984 event, predictions are issued in September and Decem-
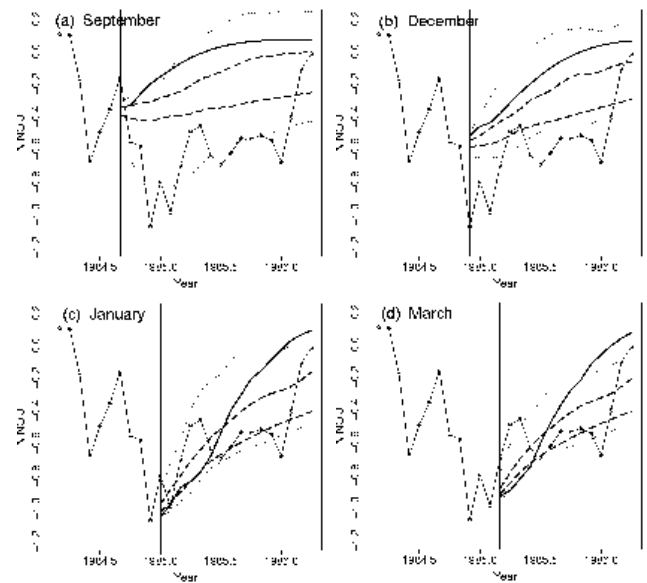

Fig. 8. Same as Fig. 6 but for blind forecasts of 1984–1985 La Niña event issued from **(a)** September 1984, **(b)** December 1984, **(c)** January 1985, and **(d)** March 1985.

ber (Figs. 8a and 8b, before La Niña reaches it's negative peak), in January (Fig. 8c), and in March (Fig. 8d). For all the cases, predictions from both the methods perform similarly and yield good results during its peak (Fig. 8c) and after its peak (Fig. 8d). For the 1999 event, predictions are issued in June (Fig. 9a), November (Fig. 9b), December (Fig. 9c), and in February (Fig. 9d). Here too the methods perform similarly, with increasing skill of the predictions when issued closer to the negative peak of the events. Lastly, we issue forecasts for the recent past starting 1 May 2002
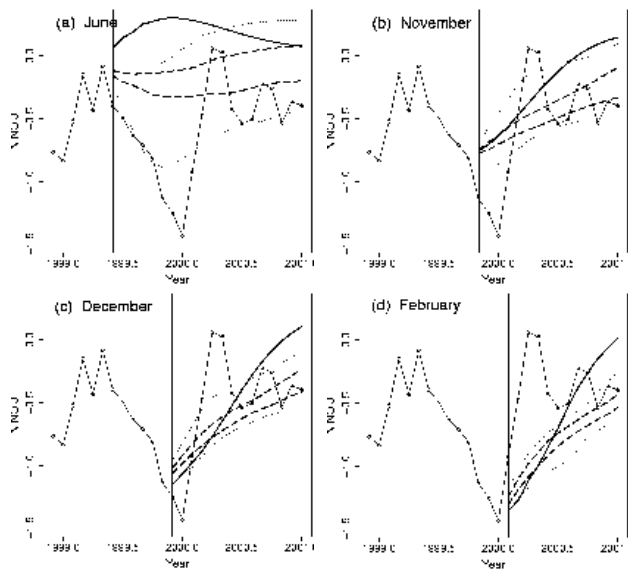
Fig. 9. Same as Fig. 6, but for blind forecast of 1999–2000 La Niña event issued from **(a)** June 1999, **(b)** November 1999, **(c)** December 1999, and **(d)** February 2000.
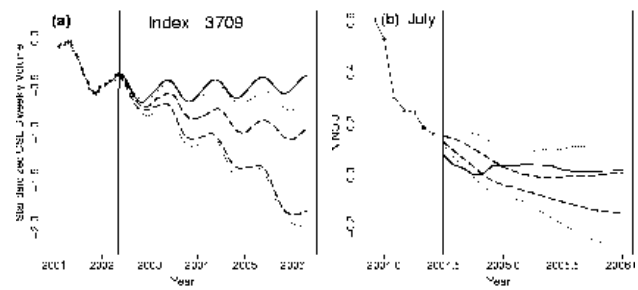

Fig. 10. Near term forecast of the real data sets. Blind forecast of **(a)** GSL starting at 1 May 2002 (index 3709), and **(b)** NINO3 index, issued on 1 July 2004.

5 Summary and discussions

We developed a new framework to generate ensemble forecasts of univariate time series. In this, the state space is first reconstructed by embedding the univariate time series of the response variable in a space of dimension D with a delay time τ. To obtain a forecast from a given time point t, three steps are involved: (i) the current state of the system is mapped on to the state space, known as the feature vector, (ii) a small number (K) of neighbors and their future evolution to the feature vector are identified in the state space, and (iii) a polynomial of order p is fitted to the identified neighbors, which is then used for prediction. Whereas in traditional nonlinear dynamical based forecasting approach, the parameters D and τ are obtained using standard algorithms and are fixed throughout the forecast period, and so are the other parameters p and α.

With short noisy data, which is often the case in real life, the parameters have significant uncertainty; hence, keeping them fixed does not yield good forecasts. Furthermore, it is only appropriate to generate ensemble forecasts that can naturally provide the forecast uncertainty. To facilitate this, our proposed framework entails the use of an objective criterion, the Generalized Cross Validation. The GCV score is computed for several parameter combinations of (D, τ, α, p). From this, a suite of parameter combinations with low GCV

and 1 July 2004 for the GSL volumes and NINO3, respectively (Fig. 10). For the ENSO, neutral to weak La Niña conditions during the coming months are indicated by the forecasts. This seems to be consistent with the forecasts issued by the climate centers (http://iri.ldeo.columbia.edu; http://www.cpc.ncep.noaa.gov) using dynamical and statistical models.

scores is selected. Each of the selected combinations is then used to issue a T-step iterated forecast starting from the current time t, thus generating an ensemble forecast. We demonstrated this method on two synthetic and two real data sets and the performance was quite good. The unique aspect of this approach is that the varying parameter combinations at different points in the state space provide the ability to better capture the underlying dynamics from noisy data. Another useful aspect of the proposed methodology is that the confidence intervals provided by the ensembles are wider and narrower for different forecast starting points, reflecting the uncertainty at different parts of the state space. However, traditional time series models (e.g. AR, ARMA, etc.) and even nonparametric methods tend to provide a fixed confidence interval that is often narrow and also unrealistic.

The method presented here may be improved in several ways. First, explicitly accounting for the error in the forecast model (Eq. 1) should allow the production of more realistic probabilistic forecasts in situations where there is low predictability. For example, when the Great Salt Lake is in a regime transition (Figs. 5a and 5d), the observed time series lies well outside the forecast ensemble spread – explicitly accounting for forecast model error for each parameter combination should inflate the ensemble spread in these situations. Second, assigning weights to different parameter combinations, e.g. based on the Generalized Cross Validation metric, may also provide more reliable probabilistic forecasts. Third, smoothing of geophysical time series using wavelet or singular spectrum analysis may result in further improvements in probabilistic forecast skill for noisy time series. Rigorous evaluation of these methods using ensemble diagnostics (e.g. the ranked histogram and spread-skill relationships, Whittaker and Loughe, 1998; Hamill, 2001) and comparison with physical models (e.g. the Cane-Zebiak ENSO forecasting model, Zebiak and Cane, 1987) is necessary before these methods can be used in real-time applications.

Edited by: B. Sivakumar
Reviewed by: two referees

References

Abarbanel, H. D. I., Rabinovich, M. I., and Sushchik, M. M.: Introduction to nonlinear dynamics for physicists, World Scientific, New Jersey, 1993.

Abarbanel, H. D. I. and Lall, U.: Nonlinear dynamics of Great Salt Lake: system identification and prediction, Clim. Dynam., 12, 287–297, 1996.

Abarbanel, H. D. I., Lall, U., Moon, Y., Mann, M. E., and Sangoyomi, T.: Nonlinear dynamics of the Great Salt Lake: A predictable indicator of regional climate, Energy, 21, 655–665, 1996.

Balmaseda, M. A., Anderson, D. L. T., and Davey, M. K.: ENSO prediction using a dynamical ocean model coupled to statistical atmospheres, Tellus A, 46, 497–511, 1994.

Barnett, T. P., Graham, N. E., Cane, N. A., Zeibak, S. E., Dolan, S. C., O'Brien, J. J., and Legeler, D. M.: On the prediction of the El Niño of 1986–1987, Science, 241, 192–196, 1988.

Barnett, T. P., Latif, M., Graham, N., Flugel, M., Pazan, S., and White, W.: ENSO and ENSO-related predictability: Part 1 – Prediction of equatorial Pacific sea surface temperatures with a hybrid coupled ocean-atmosphere model, J. Climate, 6, 1545–1566, 1993.

Barnston, A. G. and Ropelewski, C. F.: Prediction of ENSO episodes using canonical correlation analysis, J. Climate, 5, 1316–1345, 1992.

Brockwell, P. J. and Davis, R. A.: Introduction to time series and forecasting, 2nd ed., Springer-Verlag, New York, 2002.

Cane, M. A., Zebiak, S. E., and Dolan, S. C.: Experimental Forecasts of El Niño, Nature, 321, 827–832, 1986.

Casdagli, M., Eubank, S., Farmer, J. D., Gibson, J., Des Dardins, D., Hunter, N., and Theiler, J.: Nonlinear modeling of chaotic time series: theory and applications, Los Alamos National Laboratories, LA-UR-91-1637, 1990.

Chang, P., Wang, B., Li, T., and Ji, L.: Interactions between the seasonal cycle and the southern oscillation-Frequency entrainment and chaos in a coupled ocean-atmosphere model, Geophys. Res. Lett., 21, 2817–2820, 1994.

Chatfield, C.: The analysis of time series: an introduction, 6th ed., Chapman & Hall/CRC, 2003.

Chen, D., Zebiak, S. E., Busalacchi, A. J., and Cane, M. A.: An improved procedure for El Niño forecasting – implications for predictability, Science, 269, 1699–1702, 1995.

Chen, D. K., Zebiak, S. E., Cane, M. A., and Busalacchi, A. J.: Initialization and predictability of a coupled ENSO forecast model, Mon. Weath. Rev., 125, 773–788, 1997.

Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E., and Huang, D. J.: Predictability of El Niño over the past 148 years, Nature, 428, 733–736, 2004.

Elshorbagy, A., Simonovic, S. P., and Panu, U. S.: Estimation of missing streamflow data using principles of chaos theory, J. Hydrol., 255, 123–1333, 2002.

Farmer, J. D. and Sidorowich, J. J.: Prediction chaotic time series, Phys. Rev. Lett., 59, 845–848, 1987.

Fraser, A. M. and Swinney, L. H.: Independent coordinates for strange attractor from mutual information, Phys. Rev. Lett., 33, 1134–1140, 1986.

Grantz, K.: Using large-scale climate information to forecast seasonal streamflow in the Truckee and Carson rivers, M. S. Thesis, University of Colorado, Boulder, Colorado, 2003.

Grassberger, P. and Procaccia, I.: Characterization of strange attractors, Phys. Rev. Lett., 50, 346–349, 1983a.

Grassberger, P. and Procaccia, I.: Measuring the strangeness of strange attractors, Physica D, 9, 189–208, 1983b.

Grassberger, P., Schreiber, T., and Schaffrath, C.: Nonlinear time sequence analysis, Int. J. Bifur. Chaos, 1, 521–547, 1991.

Hamill, T. M.: Interpretation of rank histograms for verifying ensemble forecasts, Mon. Weath. Rev., 129, 550–560, 2001.

Henon, M.: A two-dimensional mapping with a strange attractor, Comm. Math. Phys., 50 69–77, 1976.

James, L, D., Bowles, D. S., James, W. R., and Canfiled, R. V.: Estimation of water surface elevations probabilities and associated damages for the Great Salt Lake, Utah Water Res. Lab., Utah Sate Univ., Logan, Utah, 1979.

Jayawardena, A. W. and Lai, F.: Analysis & prediction of chaos in rainfall & stream flow time series, J. Hydrol., 153, 23–52, 1994.

Jayawardena, A. W. and Gurung, A. B.: Noise reduction & prediction of hydrometeorological time series: dynamical system approach vs. stochastic approach, J. Hydrol., 228, 242–264, 2000.

Jin, F.-F., Neelin, D., and Ghil, M.: ENSO on the devil's stair-case, Science, 264, 70–72, 1994.

Kantz, H. and Schreiber, T.: Nonlinear time series analysis, Cambridge University Press, Cambridge, UK, 1997.

Kantz, H. and Schreiber, T.: Human ECG: nonlinear deterministic versus stochastic aspects, IEE P-Sci, M., 145, 279–284, 1998.

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., and Rajagopalan, B.: Analyses of global sea surface temperature 1856–1991, J. Geophys. Res. (Oceans), 103, 18 567–18 589, 1998.

Kennel, M. B., Brown, R., and Abarbanel, H. D. I.: Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A., 45, 3403–3411, 1992.

Krasovskaia, I., Gottschalk, L., and Kundzewicz, Z. W.: Dimensionality of Scandinavian river flow regimes, Hydro. Sci. J., 44, 705–723, 1999.

Laio, F., Porporato, A., Revelli, R., and Ridolfi, L.: A comparison of nonlinear flood forecasting methods, Water Resour. Res., 39, 1129, doi:10.1029/2002WR001551, 2003.

Lall, U. and Mann, M.: The Great Salt Lake: A barometer of low frequency climatic variability, Water Resour.Res., 31, 2503–2515, 1995.

Lall, U., Sangoyomi, T., and Abarbanel, H. D. I.: Nonlinear dynamics of the Great Salt Lake: Nonparametric short-term forecasting, Water. Resour. Res., 32, 975–985, 1996.

Latif., M., Barnett, T. P., Cane, M. A., Flugel, M., Graham, N. E., Storch, H. V., Xu, J-S., and Zebiak, S. E.: a review of ENSO prediction studies, Clim. Dynam., 9, 167–179, 1994.

Lisi, F. and Villi, V.: Chaotic forecasting of discharge time series: a case study, J. Am. Wat. Resour. Assoc., 37, 271–279, 2001.

Liu, Q., Islam, S., Rodriguez-Iturbe, I., and Le., Y.: Phase-space analysis of daily streamflow: characterization and prediction, Adv. Water Resour., 21, 463–475, 1998.

Loader, C. R.: Locfit: An introduction, Statistical computing and Statistical Graphics Newsletter, 8, 11–17, 1997.

Loader, C.: Local regression and likelihood, Springer-Verlag, New York, 1999.

Lorenz, E. N.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141, 1963.

Mann, M., Lall, U., and Saltzman, B.: Decadal to century scale climatic variability: Understanding the rise and fall of the Great Salt Lake, Geophys. Res. Lett., 22, 937–940, 1995.

Mason, S. J. and Mimmack, G. M.: Comparison of some statistical methods of probabilistic forecasting of ENSO, J. Climate, 15, 8–29, 2002.

Moon, Y.-I.: PhD Thesis, Utah State University, Logan, Utah, 1996.

Moon, Y.-I., Rajagopalan, B., and Lall, U.: Estimation of mutual information using kernel density estimators, Phys. Rev. E., 52, 2318–2321, 1995.

Moon, Y.-I. and Lall, U.: Atmospheric flow indices and internannual Great Salt Lake variability, J. Hydro. Eng., 2, 55–62, 1996.

Munnich, M., Cane, M. A., and Zebiak, S. E.: A study of self excited oscillations of the tropical ocean-atmosphere system. Part II: Nonlinear cases, J. Atmos. Sci., 48, 1238–1248, 1991.

Owosina, A.: Methods for assessing the space and time variability of groundwater data, M.S. Thesis, Utah State University, Logan, Utah, 1992.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. D.: Geometry from a time series, Phys. Rev. Lett., 45, 712–716, 1980.

Pandit, S. M. and Yu, S.-M.: Time series and system analysis with applications, Wiley, New York, 1983.

Porporato, A. and Ridolfi, L.: Clues to the existence of deterministic chaos in river flow, Int. J. Mod. Phys. B., 10, 1821–1862, 1996.

Porporato, A. and Ridolfi, L.: Nonlinear analysis of river flow time sequences, Water Resour. Res., 33, 1353–1367, 1997.

Porporato, A. and Ridolfi, L.: Multivariate nonlinear prediction of river flows, J. Hydrol, 248, 109–122, 2001.

Prairie, J.: Long-term salinity prediction with uncertainty analysis: application for Colorado river above glenwood springs, CO, M. S. Thesis, University of Colorado, Boulder, Colorado, 2002.

Rajagopalan, B. and Lall, U.: Locally weighted polynomial estimation of spatial precipitation, J.Geographic Information and Decision Analysis, 2, 44–51, 1998.

Rasmusson, E. and Carpenter, T.: Variations in the tropical sea surface temperature and surface wind fields associated the Southern Oscillation/El Niño, Mon. Weath. Rev., 110, 354–384, 1982.

Richter, M. and Schreiber, T.: Phase space embedding of electrocardiograms, Phys. Rev. E., 58, 6392–6398,1998.

Regonda, S. K., Sivakumar, B., and Jain, A.: Temporal scaling in river flow: can it be chaotic?, Hydro. Sci. J., 49, 373–384, 2004.

Rodriguez-Iurbe, I., dePower, B. F., Sharifi, M. B., and Georgakakos, K. P.: Chaos in rainfall, Water Resour. Res., 25, 1667–1675, 1989.

Ropelewski, C. F. and Halpert, M. S.: North American precipitation and temperature patterns associated with the El Niño/Southern Osciallation (ENSO), Mon. Weath. Rev., 115, 2352–2362, 1986.

Sangoyomi, T.: Climatic variability and dynamics of Great Salt Lake hydrology, PhD Dissertation, Department of Civil and Environmental Engineering, Utah State Univ., Logan, Utah, 1993.

Sangoyomi, T., Lall, U., and Abarbanel, H. D. I.: Nonlinear dynamics of the Great Salt Lake: dimension estimation, Water. Resour. Res., 32, 149–160, 1996.

Schreiber, T. and Grassberger, P.: A simple noise-reduction method for real data, Phys. Lett. A., 160, 411–418, 1991.

Schreiber, T. and Kantz, H.: Observing and predicting chaotic signals: Is 2% noise too much?, in Predictability of Complex Dynamical Systems, Springer Ser. In Synergetics, edited by Kravtsov, Y. A. and Kadtke, J. B., Springer-Verlag, New York,

43–65, 1996.

Sharifi, M. B., Georgakakos, K. P., and Rodriguez-Iturbe, I.: Evidence of deterministic chaos in the pulse of storm rainfall, J. Atmos. Sci., 47, 888–893, 1990.

Sivakumar, B., Liong, S.-Y., and Liaw, C.-Y.: Evidence of chaotic behavior in Singapore rainfall, J. Am. Wat. Resour. Assoc., 34, 301–310, 1998.

Sivakumar, B., Phoon, K.-K., Liong, S.-Y., and Liaw, C.-Y.: Singapore rainfall behavior: Chaotic?, J. Hydro. Eng., 4, 38–47, 1999a.

Sivakumar, B., Phoon, K-K., Liong, S-Y., and Liaw, C-Y.: Comment on "Nonlinear analysis of river flow time sequences" by A. Porporato and L. Ridolfi, Water. Resour. Res., 35, 895–897, 1999b.

Sivakumar, B., Berndtsson, R., and Persson, M.: Monthly runoff prediction using phase space reconstruction, Hydro. Sci. J., 46, 377–387, 2001.

Sivakumar, B., Persson, M., Berndtsson, R., and Uvo, C. B.: Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series?, Water. Resour. Res., 38, doi:10.1029/2001WR000333, 2002.

Stehlik, J.: Deterministic chaos in runoff series, J. Hydrol. Hydrom., 47, 271–287, 1999.

Stehlik, J.: Searching for Chaos in rainfall and temperature trends – a nonlinear analysis of time series from an experimental basin, J. Hydrol., 153, 28–52, 2000.

Takens, F.: Detecting strange attractors in turbulence. In: Dynamcial Systems and Turbulence, Lecture Notes in Mathematics 898, edited by Rand, D. A. and Young, L. S., Springer-Verlag, Berlin, Germany, 366–381, 1981.

Tsonis, A. A.: Chaos – From theory to applications, Plenum Press, New York, 1992.

Tziperman, E., Stone, S., Cane M. A., and Jarosh, H.: El-Niño chaos – overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator, Science, 264, 72–74, 1994.

Tziperman, E., Cane, M. A., and Zebiak, S.: Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos, J. Atmos., Sci., 52, 293–306, 1995.

Vallis, G. K.: El Niño: A chaotic dynamical system?, Science, 232, 243–245, 1986.

Vallis, G. K.: Conceptual models of El Niño and the Southern Oscillations, J. Geophys. Res., 93C, 13 979–13 971, 1988.

Wang, Q. and Gan, T. Y.: Biases of correlation dimension estimates of streamflow data in the Canadian prairies, Water Resour. Res., 34, 2329–2339, 1998.

Whittaker, J. S. and Loughe, A. S.: The relationship between ensemble spread and ensemble mean skill, Mon. Weath. Rev., 126, 3292–3302, 1998.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A.: Deterministic lyapunov exponents from a time series, Physica D, 16, 285–317, 1985.

Xue, Y., Cane, M. A., Zebiak, S. E., and Blumenthal, M. B.: On the prediction of ENSO – a study with a low-order markov model, Tellus A, 46, 512–528, 1994.

Xue, Y., Cane, M. A., and Zebiak, S. E.: Predictability of a coupled model of ENSO using singular vector analysis, 2. Optimal growth & forecast skill, Mon. Weath. Rev., 125, 2057–2073, 1997.

Zebiak, S. E. and Cane, M. A.: A Model El Niño – Southern Oscillation, Mon. Weath. Rev., 115, 2262–2278, 1987.