

Predictive uncertainty of chaotic daily streamflow using ensemble wavelet networks approach

C. T. Dhanya¹ and D. Nagesh Kumar¹

Received 27 October 2010; revised 24 February 2011; accepted 17 March 2011; published 11 June 2011.

[1] Perfect or even mediocre weather predictions over a long period are almost impossible because of the ultimate growth of a small initial error into a significant one. Even though the sensitivity of initial conditions limits the predictability in chaotic systems, an ensemble of prediction from different possible initial conditions and also a prediction algorithm capable of resolving the fine structure of the chaotic attractor can reduce the prediction uncertainty to some extent. All of the traditional chaotic prediction methods in hydrology are based on single optimum initial condition local models which can model the sudden divergence of the trajectories with different local functions. Conceptually, global models are ineffective in modeling the highly unstable structure of the chaotic attractor. This paper focuses on an ensemble prediction approach by reconstructing the phase space using different combinations of chaotic parameters, i.e., embedding dimension and delay time to quantify the uncertainty in initial conditions. The ensemble approach is implemented through a local learning wavelet network model with a global feed-forward neural network structure for the phase space prediction of chaotic streamflow series. Quantification of uncertainties in future predictions are done by creating an ensemble of predictions with wavelet network using a range of plausible embedding dimensions and delay times. The ensemble approach is proved to be 50% more efficient than the single prediction for both local approximation and wavelet network approaches. The wavelet network approach has proved to be 30%–50% more superior to the local approximation approach. Compared to the traditional local approximation approach with single initial condition, the total predictive uncertainty in the streamflow is reduced when modeled with ensemble wavelet networks for different lead times. Localization property of wavelets, utilizing different dilation and translation parameters, helps in capturing most of the statistical properties of the observed data. The need for taking into account all plausible initial conditions and also bringing together the characteristics of both local and global approaches to model the unstable yet ordered chaotic attractor of a hydrologic series is clearly demonstrated.

Citation: Dhanya, C. T., and D. Nagesh Kumar (2011), Predictive uncertainty of chaotic daily streamflow using ensemble wavelet networks approach, *Water Resour. Res.*, 47, W06507, doi:10.1029/2010WR010173.

1. Introduction

[2] A major breakthrough in routine weather prediction occurred in the mid-20th century with the development of various climate models that numerically integrate an adequate set of mathematical equations which explain the physical laws governing the climatic processes occurring. However, these mathematical equations form a nonlinear dynamic system in which an infinitesimally small uncertainty in the initial conditions will eventually grow exponentially even with a perfect model, leading to a chaotic behavior [Smith *et al.*, 1998]. Lorenz [1972] identified the sensitivity of any deterministic system, even to a slight change in the initial con-

ditions, leading to a vast change in the final solution and termed it as “butterfly effect” in the field of weather forecasting. Hence, Earth’s weather can be treated as a chaotic system with a finite limit in predictability, arising mainly because of the incompleteness in initial conditions. The complexity and irregularity in chaotic systems may be due to the nonlinear interdependent variables involved, with sensitive dependence on initial conditions. An infinitesimal initial uncertainty ∂_0 grows exponentially with time at a rate of separation given by the highest Lyapunov exponent λ [Wolf *et al.*, 1985; Rosenstein *et al.*, 1993]. Thus, the separation or uncertainty after Δt time steps ahead is $\partial_{\Delta t} \cong e^{\lambda \Delta t} \partial_0$. Hence, the predictability of a chaotic system is limited primarily because of (1) the indefiniteness in the initial conditions (given a perfect model) and also (2) the imperfection of the model itself [Smith, 2000].

[3] Variations in the reliability of any individual forecast due to the first factor, i.e., indefiniteness in the initial conditions, can be quantified by generating an ensemble of fore-

¹Department of Civil Engineering, Indian Institute of Science, Bangalore, India.

casts with slightly varying initial conditions. Any uncertainty in the initial conditions is reflected in the evolution of the ensemble and hence the nonlinearity. *Smith* [2000] stated that since the ensembles can accurately reflect the likelihood of occurring of various future conditions; given a perfect model, any chaos places no a priori limits on predictability. Also, an estimate of the stability of the forecasts can be obtained by observing how quickly the ensemble spreads out (or shrinks). Many operational centers now adopt the ensemble approach, replacing the traditional best guess initial condition approach. *Shukla* [1998] made different runs of the model with different initial conditions to predict sea surface temperatures. The National Centers for Environmental Prediction Climate Forecasting System also produces retrospective predictions by initiating runs from successive days to generate a set of ensembles with different initial conditions [*Saha et al.*, 2006; *Lee Drbohlav and Krishnamurthy*, 2010].

[4] However, research into the applications of chaos theory in the field of hydrology have been focused primarily on investigating the existence of the chaotic nature of rainfall [*Jayawardena and Lai*, 1994; *Puente and Obregón*, 1996; *Rodriguez-Iturbe et al.*, 1989] and runoff [*Liu et al.*, 1998; *Porporato and Ridolfi*, 1997; *Wang and Gan*, 1998]. The concept of chaos has also been applied in other areas like rainfall disaggregation [*Sivakumar et al.*, 2001a], missing data estimation [*Elshorbagy et al.*, 2002], rainfall-runoff process [*Sivakumar et al.*, 2001b], lake volume [*Sangoyomi et al.*, 1996], and sediment transport [*Shang et al.*, 2009]. The outcomes of these studies affirm the existence of low-dimensional chaos, thus indicating the possibility of only short-term predictions. Even so, the existence of low-dimensional deterministic chaos in hydrologic processes is a debatable issue. *Schertzer et al.* [2002] questioned the applicability of correlation dimension in hydrological processes since it may lead to the depiction of complex systems as low-dimensional chaotic systems. *Khan et al.* [2005] observed that real hydrologic data may or may not have a detectable chaotic component. However, they further clarified that the absence of chaotic component in their streamflow data may be due to the stochastic mode of dam operation and the subsequent randomness occurring in the streamflow process.

[5] The nonlinear predictions of the rainfall and runoff series considering them as a univariate series [*Islam and Sivakumar*, 2002; *Sivakumar et al.*, 1999a, *Jayawardena and Lai*, 1994; *Porporato and Ridolfi*, 1997] and also as a multivariate series [*Porporato and Ridolfi*, 2001], utilizing the information from other time series, are also attempted in a few studies. The prediction algorithms used in these studies are based on the theory of dynamic reconstruction of a scalar series, which is done by reconstructing the phase space using the method of delays [*Takens*, 1981]. The phase space reconstruction provides a simplified, multidimensional representation of a single-dimensional nonlinear time series. According to this approach, given the embedding dimension m and the time delay τ , for a scalar time series X_i , where $i = 1, 2, \dots, N$, the dynamics can be fully embedded in m -dimensional phase space represented by the vector

$$\mathbf{Y}_j = (X_j, X_{j+\tau}, X_{j+2\tau}, \dots, X_{j+(m-1)\tau}). \quad (1)$$

Now the dynamics can be interpreted in the form of an m -dimensional map f_T such that

$$\mathbf{Y}_{j+T} = f_T(\mathbf{Y}_j), \quad (2)$$

where \mathbf{Y}_j and \mathbf{Y}_{j+T} are vectors of dimension m , \mathbf{Y}_j being the state at current time j and \mathbf{Y}_{j+T} being the state at future time $j + T$.

[6] In all of the studies mentioned above, the determination of embedding dimension is done using any of the conventional methods such as the correlation dimension method [*Grassberger and Procaccia*, 1983], false nearest-neighbor algorithm [*Kennel et al.*, 1992], and nonlinear prediction method [*Farmer and Sidorowich*, 1987]. Similarly, the delay time is fixed from either the autocorrelation method or the mutual information method. However, an accurate estimation of these chaotic parameters depends very much on the data size, sampling frequency, and noise present in the data. *Khan et al.* [2007] commented that even though the sampling size is large, the data size may still be considered as short in a physical sense if the sampling frequency is inadequate to capture the features of the dynamical system. Chaos identification methods and prediction are also significantly affected by even small levels of noise [*Schreiber and Kantz*, 1996; *Khan et al.*, 2005]. Significant improvements were achieved in the estimates of correlation dimension for the noise-reduced river flow and rainfall series [*Porporato and Ridolfi*, 1997; *Sivakumar et al.*, 1999b]. Nevertheless, it is also observed that the suppression of certain frequencies by noise reduction methods can alter the dynamics of the filtered output signal [*Badii et al.*, 1988; *Chennaoui et al.*, 1990].

[7] The difficulties mentioned above clearly indicate the uncertainties associated with chaos identification in hydrological processes. Since almost all chaos identification methods have their own limitations, it is advisable to employ diverse techniques to examine the existence of chaotic dynamics so that the results from one method can be fortified with those from other methods [*Dhanya and Nagesh Kumar*, 2010]. Thus, while attempting prediction, instead of a single initial condition with optimum chaotic parameters, i.e., embedding dimension and delay time, the diversity of the parameter values obtained from these methods can be taken into account by adopting an ensemble approach with different initial conditions.

[8] The approximation of f_T in equation (2) can be done using either a global or a local nonlinear model. Traditionally, studies have been employing the local approximation techniques introduced by *Farmer and Sidorowich* [1987] in which the domain is broken up into many local neighborhoods and modeling is done for each neighborhood separately; that is, there will be a separate f_T valid for each neighborhood. The complexity in modeling f_T is thus considerably reduced without affecting the accuracy of prediction. Because of this advantage, most of the studies attempting prediction have adopted the local approximation method instead of the global approach. *Sivakumar et al.* [2002] made a comparative study between local polynomial models and the global artificial neural networks (ANNs) and showed that prediction errors of local models were about 4–8 times less than those of the ANN model, which is credited to the capability of local approximation methods in better captur-

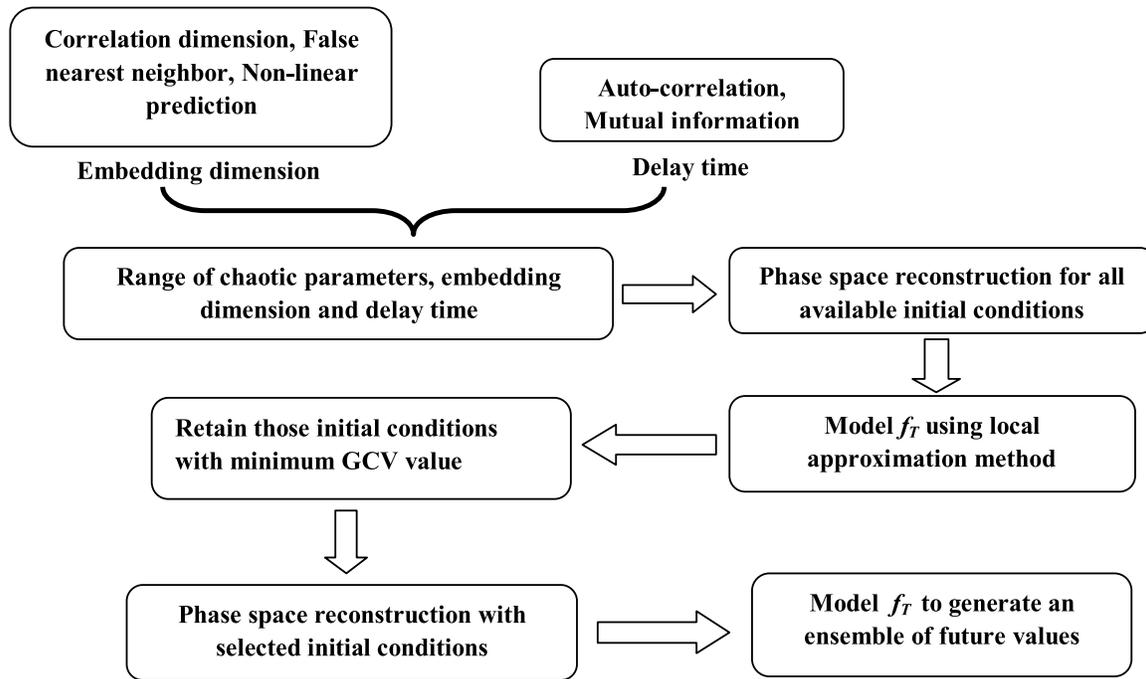


Figure 1. Schematic diagram of proposed ensemble approach.

ing the chaotic dynamics of the system. Contrary to this, *Elshorbagy et al.* [2002] have reported the superiority of the ANN model over the widely used local average (approximation) models for phase space prediction. *Karunasinghe and Liong* [2006] have also demonstrated the ability of ANNs over local approximation models to model both noise-free and noisy series. Considering all of these, a combination of local and global models featuring the benefits of both approaches may hence be adopted to improve the efficiency of the model and to increase the predictability of a chaotic system.

[9] Hence, the present study attempts to address both of the factors limiting the predictability of a chaotic series, i.e., the indefiniteness in the initial conditions, and the imperfection of the model. An ensemble of forecasts with varying combinations of embedding dimensions and delay times is generated to adequately represent the uncertainties in the initial conditions. Different combinations of embedding dimensions and delay times will lead to different initial conditions while attempting phase space reconstruction as shown in equation (1). In this study, instead of the traditional local approximation method, a relatively new technique, “wavelet network,” which is a combination of local and global approaches, is employed for the phase space prediction. First, this method is applied for the prediction of basic chaotic systems such as the Henon map and Lorenz three-variable model. Subsequently, the attention is directed toward forecasting the daily streamflow series of two stations in Mahanadi basin, India. Since much debate has been going on about the nature of streamflow regarding whether it is chaotic or not, the chaotic behavior and the predictability of the time series are analyzed initially by employing various techniques. The set of plausible parameters thus obtained is used to generate an ensemble of forecasts of the time series using the wavelet network. The efficiency of the wavelet network method is compared with the widely applied local approx-

imation method using the same set of parameters. Section 2 presents a brief overview of data used and the methodology proposed in the present study.

2. Methodology and Data

2.1. Ensemble Approach

[10] Figure 1 presents a schematic diagram of the ensemble approach developed in this study for modeling the prediction uncertainty in the chaotic streamflow series. The delay time and embedding dimension of the hydrologic variable are determined using conventional approaches. Suitable ranges of these parameters are selected for ensemble prediction, and phase space reconstruction is done for all possible combinations (which in turn leads to different initial conditions). Since all of these initial conditions may not be significant during prediction, pruning is done on the basis of the generalized cross validation (GCV) value [*Craven and Wahba*, 1978]. Studies such as those of *Lall et al.* [1996] and *Regonda et al.* [2005] had employed similar GCV values for optimization of chaotic parameters. The generalized cross validation value is given by

$$\text{GCV}(m, \tau) = \frac{\sum_{i=1}^n \frac{e_i^2}{n}}{\left(1 - \frac{p}{n}\right)^2}, \quad (3)$$

where e_i is the error, n is the number of data points, and p is the number of parameters to be determined. In the present study, since a sufficient number of combinations (around 70–100) are found falling under a GCV value of 10%, those combinations with GCV value falling under 10% of the lowest GCV value are chosen as the best ones. The phase space is reconstructed with each of these best parameter combinations. Finally, an ensemble of predictions is gener-

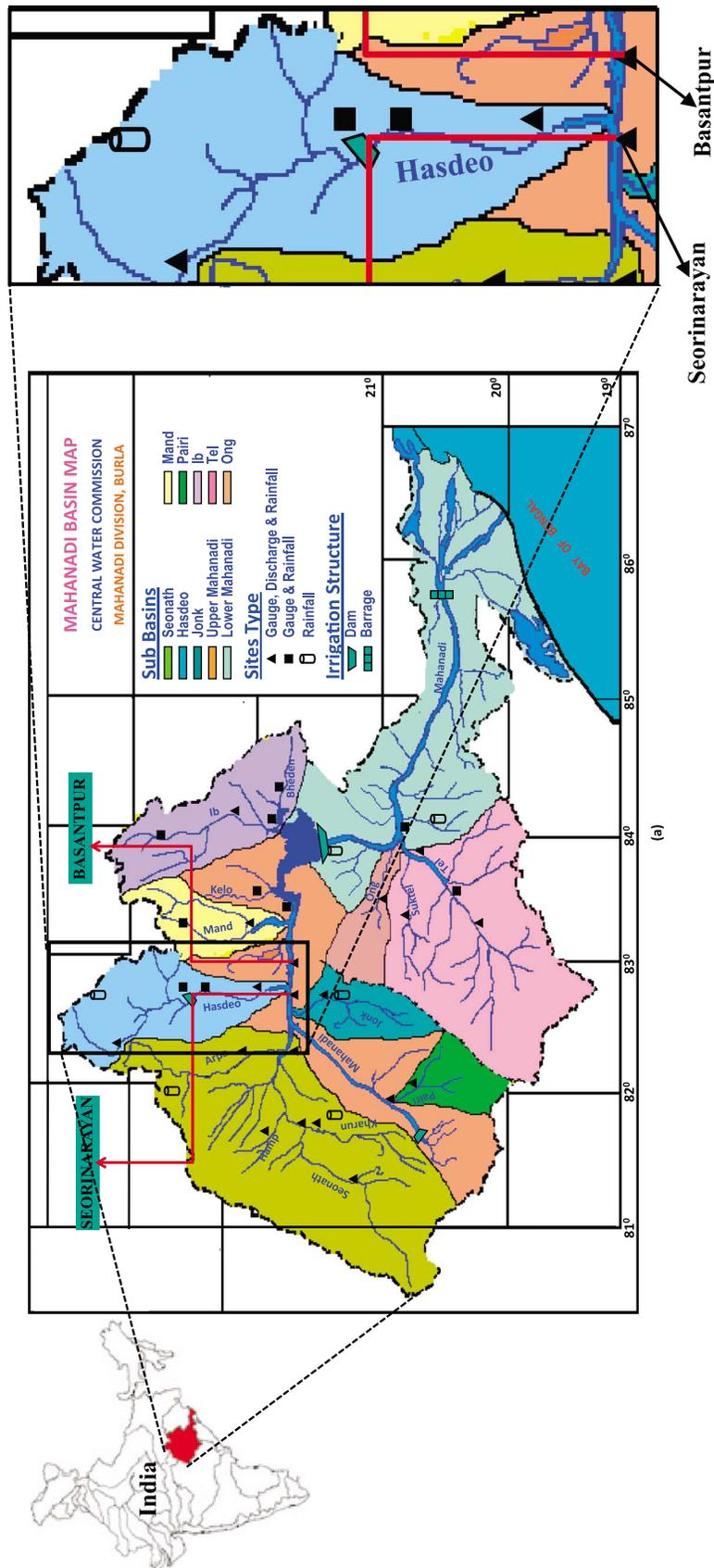


Figure 2. (a) Location map of Seorinarayan and Basantpur stations in Mahanadi basin. (b) Frequency histogram of daily streamflow and (c) box plots of the mean daily streamflow of Seorinarayan station for the period June 1986 to May 2004. (d) Frequency histogram of daily streamflow and (e) box plots of the mean daily streamflow of Basantpur station for the period June 1972 to May 2004.

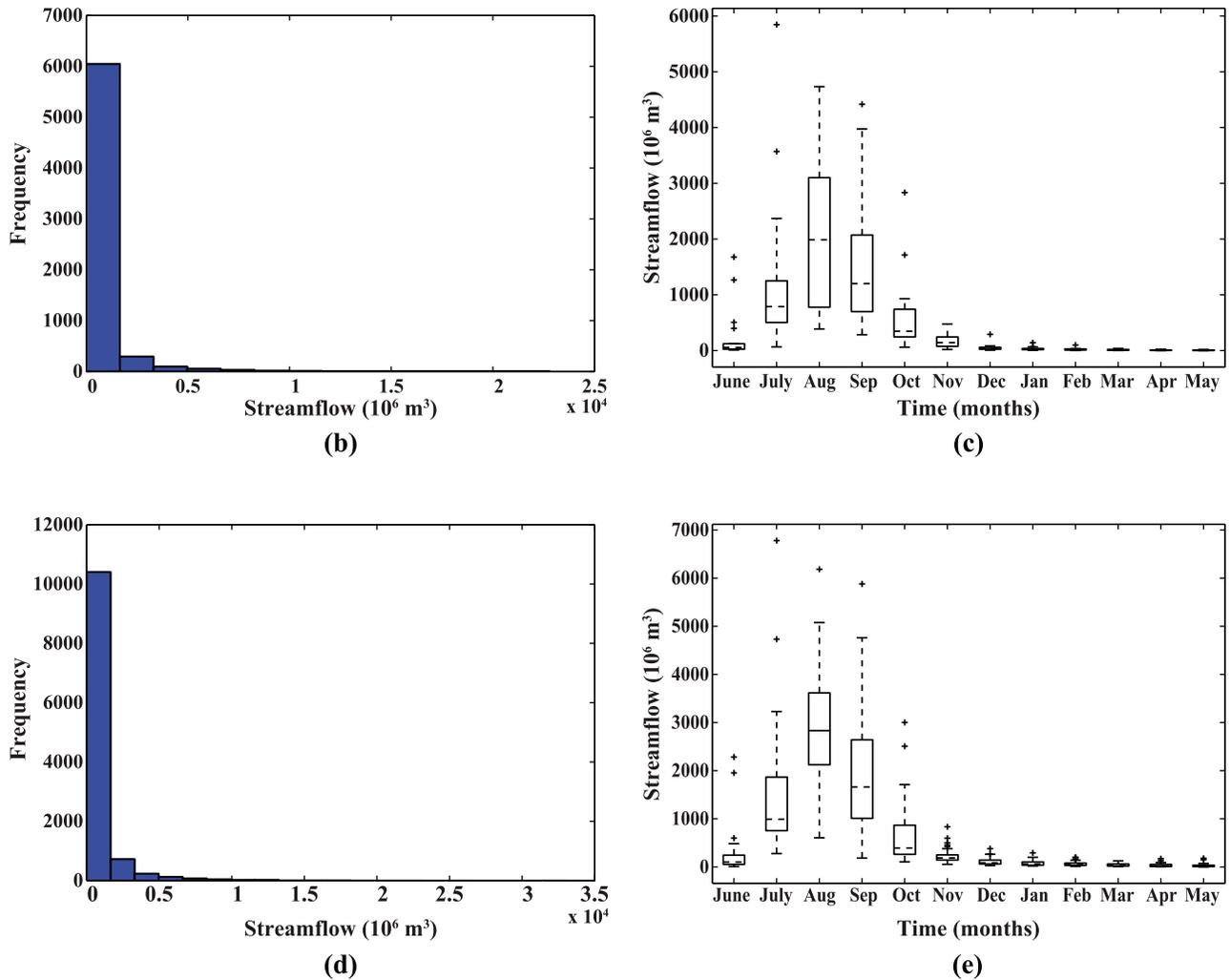


Figure 2. (continued)

ated by modeling f_T for each of these phase spaces using both wavelet network and local approximation methods. Application of this ensemble approach is demonstrated with the streamflow at two stations, Seorinarayan and Basantpur in Mahanadi basin, India.

2.2. Observed Streamflow Data

[11] The Mahanadi River of eastern India rises in the highlands of central India in Chattishgarh and flows east into the Bay of Bengal. The length of the river is 860 km. The drainage basin extends over an area of $0.141 \times 10^6 \text{ km}^2$, covering the states of Maharashtra, Jharkhand, Orissa, and Chattisgarh. Mahanadi River is a rain-fed river with maximum streamflow during June to September. The streamflow is primarily used for irrigation and power production. The data considered for this case study are the daily streamflow data at Seorinarayan and Basantpur stations in the basin (Figure 2a). Both stations are on the upstream side of the Hirakud dam and are thereby unaffected by the stochastic operations of the dam. While data from Seorinarayan station are available for the period June 1986 to May 2004 (18 years), data from Basantpur station is available for the period June 1972 to May 2004 (32 years). Among these, the first 14 years

(5114 points) of Seorinarayan streamflow data and first 28 years (10,227 points) of Basantpur streamflow data are taken as training data sets. The testing sets are the last 4 years (1461 points) of both streamflow series. Although the two stations selected are adjacent, the flow characteristics of two stations differ much because of the contribution from the major tributary Hasdeo between them (Figure 2a). Figures 2b–2e show the frequency histograms of the daily streamflow series and the box plots of the average daily streamflows for both stations. The maximum daily flow of Seorinarayan is $22,800 \times 10^6 \text{ m}^3$ and that of Basantpur station is $33,100 \times 10^6 \text{ m}^3$. The maximum frequency is falling in the range of $0\text{--}1000 \times 10^6 \text{ m}^3$. The nonmonsoonal flows are almost invariant, while the monsoon flows show large deviations from the mean. Table 1 shows the climatological mean and standard deviation of daily streamflow at both stations.

[12] On account of its wide basin area and devastating floods, numerous studies have been conducted on Mahanadi streamflow, focusing on its prediction [Maity and Nagesh Kumar, 2008; Maity et al., 2010], flood forecasting, the impact of climate change on the future flows [Asokan and Dutta, 2008; Mujumdar and Ghosh, 2008; Raje and

Table 1. Climatological Mean (10^6 m^3) and Standard Deviation (10^6 m^3) of Average Daily Streamflow of Seorinarayan and Basantpur Stations for the Study Period

Month	Seorinarayan		Basantpur	
	Mean	Standard Deviation	Mean	Standard Deviation
June	250.6	430.5	277.9	442.2
July	1279.0	1399.5	1555.2	1335.9
August	2081.3	1849.1	2856.3	2451.9
September	1564.7	1270.2	1978.6	1644.8
October	606.7	389.9	685.8	451.6
November	175.6	98.1	239.5	107.3
December	55.1	17.5	113.9	24.4
January	35.4	14.3	83.4	29.3
February	23.1	9.7	66.4	22.7
March	13.4	3.9	44.8	14.0
April	8.1	1.8	35.2	10.0
May	6.3	2.2	29.4	8.5

Mujumdar, 2009], and reservoir operation [Raje and Mujumdar, 2010]. All of these studies have assumed the existence of stochastic nature in the streamflow series. The limit of the predictability and the subsequent chaotic nature are not assessed. Dhanya and Nagesh Kumar [2010] have observed that the daily rainfall series in Mahanadi basin is inherently nonlinear and exhibits a low-dimensional chaotic nature with a predictability limit of only 40 days. Considering this and also since streamflow is a direct output of rainfall, it can be assumed that the daily streamflow series in the basin is also chaotic in nature. The chaos identification techniques employed in the Dhanya and Nagesh Kumar [2010] study on the daily rainfall series are repeated in this study also to countercheck the existence of a low-dimensional chaotic nature but without explaining the methods in detail. For details of the techniques, the readers are requested to refer to Dhanya and Nagesh Kumar [2010].

2.3. Chaotic Nature and Predictability of Streamflow Series

[13] The chaotic parameters, i.e., embedding dimension and delay time, are determined by employing various methods. The embedding dimension of the streamflow series is estimated using the correlation dimension method [Grassberger and Procaccia, 1983], false nearest-neighbor algorithm [Kennel et al., 1992], and nonlinear prediction method [Farmer and Sidorowich, 1987]. The delay time is determined using the autocorrelation method and mutual information method. Finally, the sensitiveness to initial conditions or predictability limit is assessed by checking the existence of a positive Lyapunov exponent [Kantz, 1994]. Since the streamflow series of both stations are showing similar results, the outputs of only Basantpur streamflow series are described below. The similar chaotic characteristics of both streamflow series, irrespective of the difference in their flow characteristics, may be due to their adjacency in location.

[14] A saturated correlation exponent of 5.21 at an embedding dimension $m = 18$ (Figure 3a) and a minimum (and thereafter constant) fraction of nearest neighbors at an embedding dimension of 7 (Figure 3b) suggests the possible presence of low-dimensional chaotic behavior in the streamflow series. Also, Figure 3c shows that the prediction error is minimum for a neighborhood size of 0.6–0.7 of standard deviation and thereafter it starts increasing for higher values of

neighborhoods. This is in agreement with Casdagli's test for nonlinearity [Casdagli, 1991], which states that if the prediction error decreases up to a certain number of nearest neighbors and increases for higher numbers, it shows the evidence of chaos in the data series. Further, the prediction accuracy is measured in terms of correlation coefficient, and the variation of this for various embedding dimensions using an optimum nearest-neighbor size of $0.6 \times$ standard deviation is shown in Figure 3d. The prediction efficiency is expected to increase to a value close to 1 with an increase in embedding dimension m up to an optimal m and remain constant afterward in the case of a chaotic time series. On the other hand, for a stochastic time series, there would not be any change in the prediction accuracy with an increase in the embedding dimension [Casdagli, 1989]. The maximum prediction accuracy is attained for an embedding dimension of 6 and it remains almost a constant further (Figure 3d). These again support the presence of chaos in the rainfall series, and hence, the optimum embedding dimension from the nonlinear prediction method is chosen as 6.

[15] Since the power spectrum of the daily streamflow is exhibiting a power law shape, i.e., $P(f) \propto f^{-\alpha}$ with $\alpha \approx 1.35$, the possibility of any pseudo low-dimensional chaos in the streamflow series is judged by repeating the correlation dimension method on the first derivative and the phase randomized data of the original signal [Osborne and Provenzale, 1989; Provenzale et al., 1992]. The variations of the correlation exponent with embedding dimension for the first derivative of data, phase-randomized data, and original data are shown in Figure 3e. While the variation of correlation exponent of first derivative is almost identical to that of the original data with almost the same saturation value, the correlation dimension of the phase-randomized data set is not converging at all. This eliminates the possibility of any linear correlation forcing the saturation of correlation exponent and thereby confirms the presence of a low-dimensional strange attractor in the streamflow series.

[16] The choice of the delay time τ is made using the autocorrelation method and the mutual information method. In the autocorrelation method, the lag time at which the autocorrelation function attains a zero value (Figure 4a), i.e., the 74th day, is considered as the delay time. In mutual information, the delay time is the first minimum value, which is at the 90th day (Figure 4b).

[17] The Lyapunov exponent provides a measure of the exponential growth due to infinitesimal perturbations. The maximal Lyapunov exponent is calculated employing the algorithm suggested by Rosenstein et al. [1993], which is based on the nearest-neighbor approach. The variations of $S(\Delta t)$ with time t for Basantpur station at dimensions $m = 4$ to 6 is shown in Figure 4c. A positive Lyapunov exponent (slope of the linear part of the curve) of around 0.167 confirms the exponential divergence of trajectories and hence the chaotic nature of daily streamflow. The inverse of the Lyapunov exponent defines the predictability of the system, which is only around 7 days. The results from the above methods have thus confirmed the daily streamflow series as an inherently low-dimensional chaotic system with dimension in the range 6–7 and with subsequent sensitiveness to initial conditions and limited predictability.

[18] As explained in section 1, most of the chaos identification methods are affected by data size, sampling frequency, and noise. However, the daily streamflow data from both

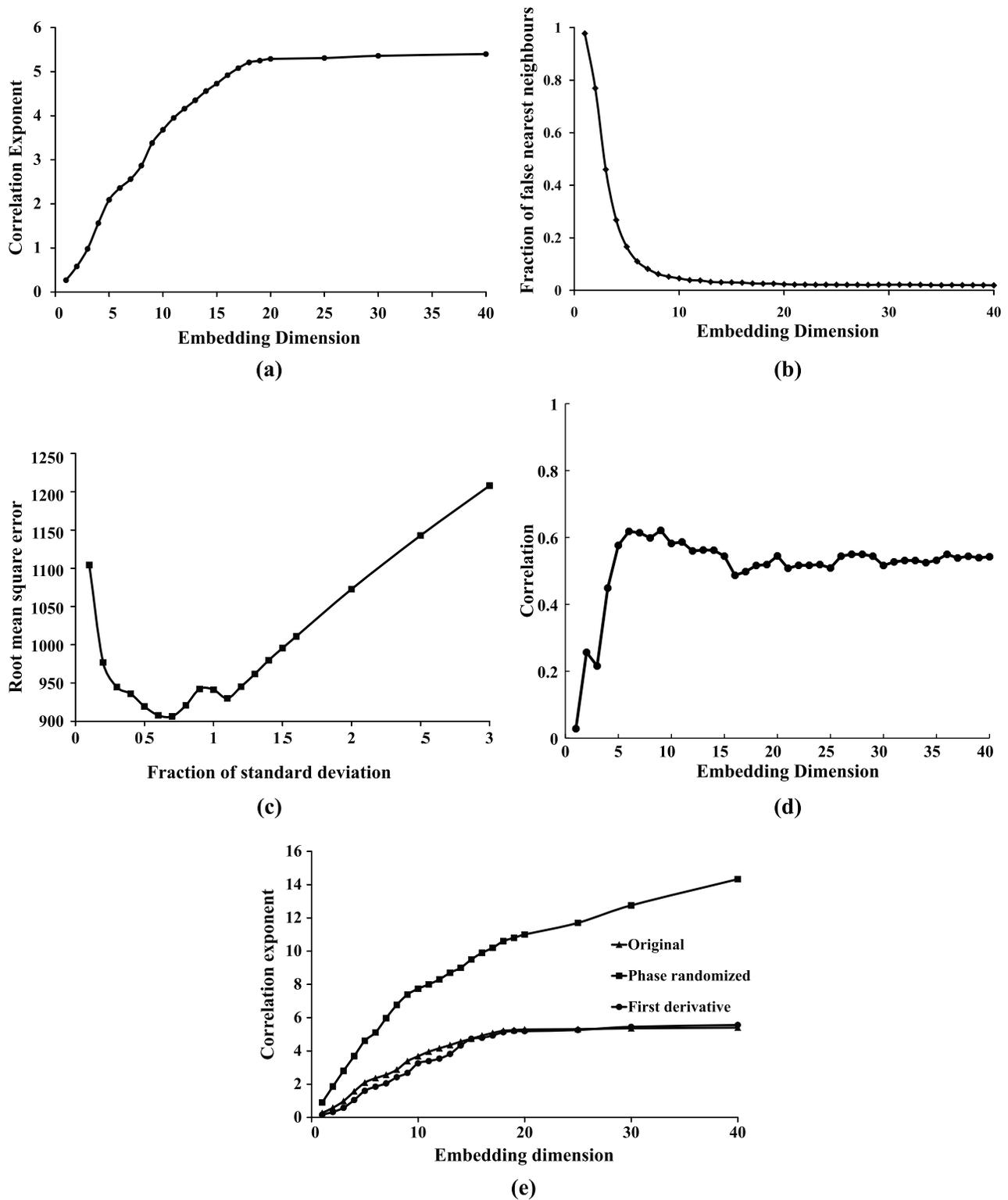


Figure 3. (a) Variation of correlation exponent with embedding dimension. (b) Variation of fraction of false nearest-neighbors with embedding dimension. (c) Variation of prediction error with neighborhood size. (d) Variation of correlation coefficient with embedding dimension. (e) Variation of correlation exponent with embedding dimension for original data, phase-randomized data, and first derivative of data.

stations used in this study have 6575 points (18 years) and 11,688 points (32 years), which are sufficient to capture the streamflow dynamics. In order to examine the effect of noise on the chaotic identification methods, a simple non-

linear noise reduction method [Schreiber, 1993] is applied on the streamflow data. The repetition of the false nearest-neighbor (FNN) method in the noise-reduced data reveals that the fraction of FNNs is falling to a minimum at an

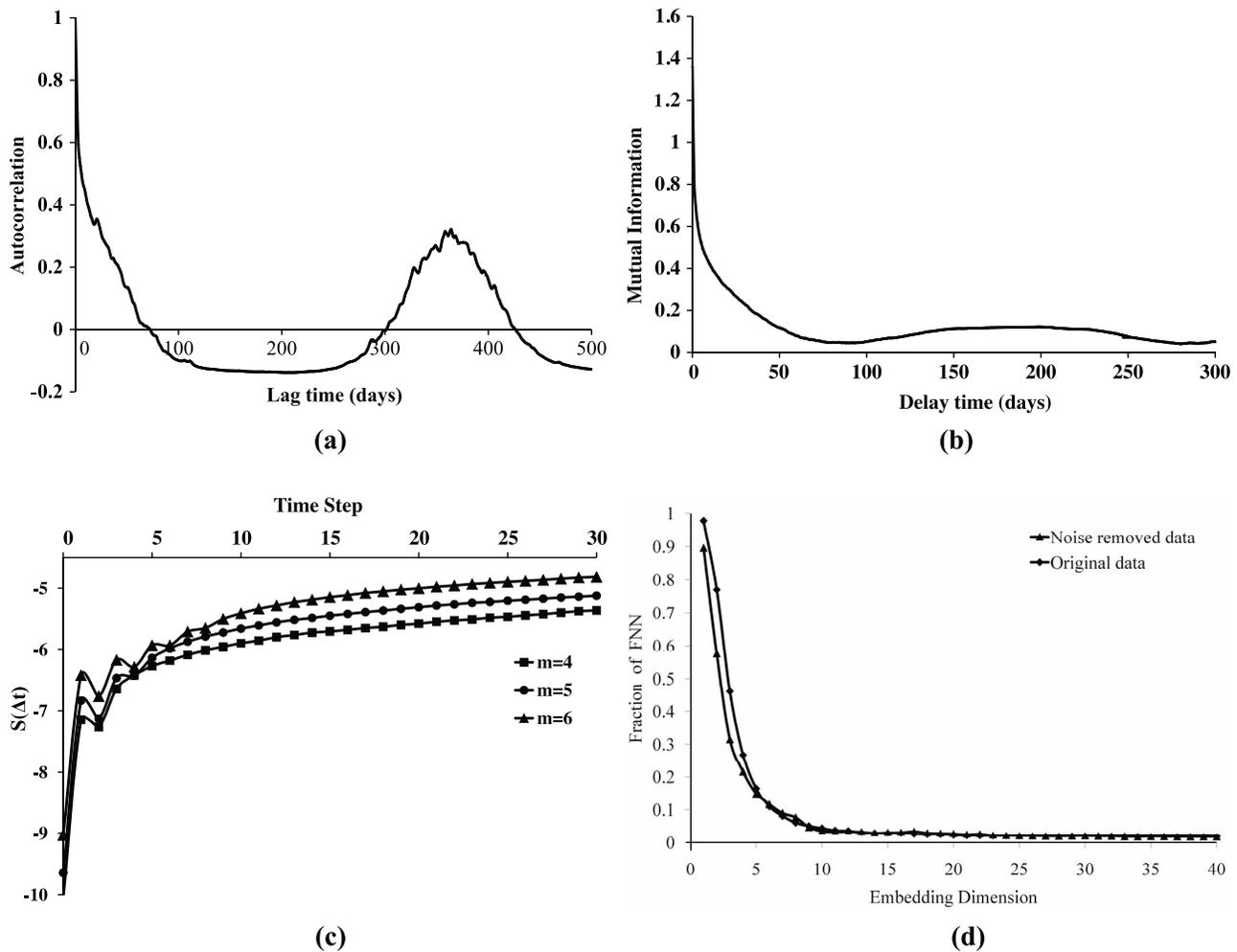


Figure 4. (a) Autocorrelation function of Basantpur streamflow. (b) Variation of mutual information with lag time. (c) Variations of $S(\Delta t)$ with time for various embedding dimensions. (d) Variation of fraction of false nearest neighbors with embedding dimension for noise-removed data and original data.

embedding dimension of 7, and the variation is almost similar to the original data (Figure 4d).

[19] Even though the embedding dimensions of original and noise-reduced data are the same, it is worthwhile to mention that for assessing the effect of noise on estimators, a more detailed study would be needed. However, in the present study, a detailed analysis of the effect of noise reduction is not attempted as it needs more investigation to find a suitable noise reduction method and also there is not much variation in the final results. Besides these, the ensemble prediction approach utilizing a range of parameters will be able to neutralize the effects of noise in data (if any) and also the shortcomings of the estimator estimation methods. Section 3 presents a brief overview of the wavelet networks and the network structure.

3. Wavelet Networks

3.1. Background

[20] Recent studies that applied neural networks for chaotic time series prediction have used regular feed-forward networks with sigmoid transfer functions. Of the two types of activation functions (global as in back propagation net-

works and local as in radial basis function networks) commonly used in neural networks, the latter can improve the convergence and efficiency because of its localized learning skill. On the other hand, global function iterates many nodes, hence slowing down the adaptation and learning procedure. It also lacks local learning and manipulation of the network [Bakshi and Stephanopoulos, 1993]. Several studies have been carried out to improve the design of neural networks by analyzing the relationship between neural networks, approximation theory, and functional analysis [Zhang and Benveniste, 1992; Bakshi and Stephanopoulos, 1993; Rying et al., 2002]. Most of the orthogonal functions, used in functional analysis to represent the continuous functions (as their weighted sum), are global approximators. The introduction of an orthogonal function with local properties can improve the results comparatively, especially while dealing with complex chaotic attractor predictions. A combination of global and local models that features the advantages of both approaches may ameliorate the predictability by reducing the uncertainty.

[21] Zhang and Benveniste [1992] proposed a new prediction technique combining the learning and feed-forward neural networks and the high-resolution wavelets. They introduced a $(1 + \frac{1}{2})$ network employing wavelets as the activation func-

tions. *Bakshi and Stephanopoulos* [1993] have introduced a second type of wavelet network (or wave-nets) in which wavelets are incorporated in a radial basis function network. Wavelets are a family of basis functions combining the powerful properties of orthogonality, localization in time and frequency, fast algorithms, and compact support [*Daubechies*, 1988; *Mallat*, 1989]. Wavelets, being both orthogonal and local provide full advantage of orthonormality and localized learning, are advisable for chaotic time series prediction. The effectiveness of this technique for short-term prediction of chaotic time series has been illustrated by many studies by demonstrating it on various basic chaotic systems like the Lorenz attractor, Ikeda map, logistic map, Rossler attractor, and Henon attractor [*Cao et al.*, 1995; *Alarcon-Aquino et al.*, 2005]. *Garcia-Treviño and Alarcon-Aquino* [2006] had shown the superiority of wavelet networks over similar back propagation networks.

3.2. Wavelet Theory

[22] Wavelet analysis is a powerful tool used for function analysis similar to Fourier analysis. While Fourier analysis approximates any periodic function as the sum of sines and cosines by addressing frequency alone, wavelet analysis represents any arbitrary function as the sum of wavelets by addressing both space and scale. Wavelet transform decomposes an arbitrary signal into elementary contributions (or wavelets) which are constructed from one single function ψ called the mother wavelet by adjusting its two parameters, dilation and translation:

$$\psi_{\alpha,\beta}(x) = \sqrt{\alpha} \psi[\alpha(x - \beta)] \quad x \in \mathfrak{R}; \quad \alpha, \beta \in \mathfrak{R}; \quad \alpha > 0, \quad (4)$$

where $\psi(x)$ is the wavelet prototype, α is the dilation parameter (scale), and β is the translation parameter (space or time).

3.2.1. Continuous Wavelet Transform

[23] If a pair of radial functions $\varphi, \psi \in L_2(\mathfrak{R}^d)$ satisfy the condition

$$\int_0^\infty \alpha^{-1} \hat{\varphi}(\alpha\omega) \hat{\psi}(\alpha\omega) d\alpha = 1, \quad \forall \omega \in \mathfrak{R}^d, \quad (5)$$

where $\hat{\varphi}$ and $\hat{\psi}$ are the Fourier transforms of φ and ψ , respectively, then for any function $f \in L_2(\mathfrak{R}^d)$, the continuous wavelet transform can be expressed as

$$w_f(\alpha, \beta) = \alpha^{d-1/2} \int_{\mathfrak{R}^d} f(x) \varphi(\alpha(x - \beta)) dx, \quad (6)$$

where $\forall \alpha \in \mathfrak{R}_+, \beta \in \mathfrak{R}^d$. This gives a set of wavelet coefficients over different widths or levels of wavelets. Reconstruction of the function by summing over the whole range of translated and dilated wavelets is done by applying the inverse wavelet transform defined by

$$f(x) = \int_{\mathfrak{R}_+ \times \mathfrak{R}^d} w_f(\alpha, \beta) \psi[\alpha(x - \beta)] \alpha^{d-1/2} d\alpha d\beta. \quad (7)$$

However, for practical application, the inverse wavelet transform needs to be discretized.

3.2.2. Discrete Wavelet Transform

[24] The inverse wavelet transform in equation (7) is discretized into

$$f(x) = \sum_i w_i \psi[\alpha_i(x - \beta_i)]. \quad (8)$$

The existence of such an inverse discrete wavelet transform is dependent on some criteria. One common method of reconstruction is to form a finite family of dilation and translation parameters (α_i, β_i) so that the full family of decomposed wavelets $\alpha_i^{d/2} \psi[\alpha_i(x - \beta_i)]$, $\forall i \in \mathbb{Z}$ constitutes an orthonormal basis. Usually, a regularly spaced grid structure is used, with constants α_0 and β_0 defining the step sizes of the dilation and translation discretizations. Even though many efficient algorithms have been developed to construct orthonormal bases, the implementation of this method becomes difficult when dealing with multidimensional wavelets [*Kugarajah and Zhang*, 1995]. An alternative method is to create a family of wavelets that constitute a frame [*Duffin and Schaeffer*, 1952]. This concept provides more freedom in the choice of wavelet function ψ .

[25] However, in the above approaches, wavelet bases or frames are constructed for all available dilations and translations without utilizing the information from the available training data. Hence, these methods are particularly suitable for problems of small input dimension. The efficiency of the wavelet estimator will be improved if the wavelet basis is constructed applying the sampled data. In other words, instead of a regular fixed grid of (α_i, β_i) , the values of (α_i, β_i) and also w_i are adaptively determined from the training data. Such an adaptive discrete inverse wavelet transform (equation (8)), which is much similar to a $(1 + 1/2)$ layer feed-forward neural network, is termed a wavelet network [*Zhang*, 1997]. In a wavelet network [*Zhang and Benveniste*, 1992], inverse wavelet transform is (re)constructed using neural networks. Instead of initializing randomly as in usual neural networks, wavelet networks are initialized with a regular wavelet lattice.

3.3. Structure of Wavelet Network

[26] Wavelet network is organized in the form of a $(1 + 1/2)$ layer neural network with wavelets as activation functions as shown in Figure 5a. The structure of a wavelet network for a function $\psi : \mathfrak{R}^d \rightarrow \mathfrak{R}$ is of the following form:

$$f(x) = \sum_{i=1}^N w_i \psi[\alpha_i * (x - \beta_i)], \quad (9)$$

where $w_i \in \mathfrak{R}$, α_i , and $\beta_i \in \mathfrak{R}^d$ correspond to the wavelet coefficient and the dilation and translation parameters respectively. The asterisk indicates componentwise multiplication of two vectors, and the variable N is the number of wavelets. Instead of using a regular wavelet lattice, wavelet network adapts the network parameters from the training data and constructs a discretized wavelet family. A brief description of the construction of the wavelet network is given below. Interested readers are referred to *Zhang* [1997] for a more detailed presentation.

[27] The best set of wavelets for the network is obtained using the following approach [*Zhang*, 1997].

[28] 1. For the construction of a wavelet library, an initial countable set of dilated and translated representation of mother wavelet ψ is created. For convenience, typically, a

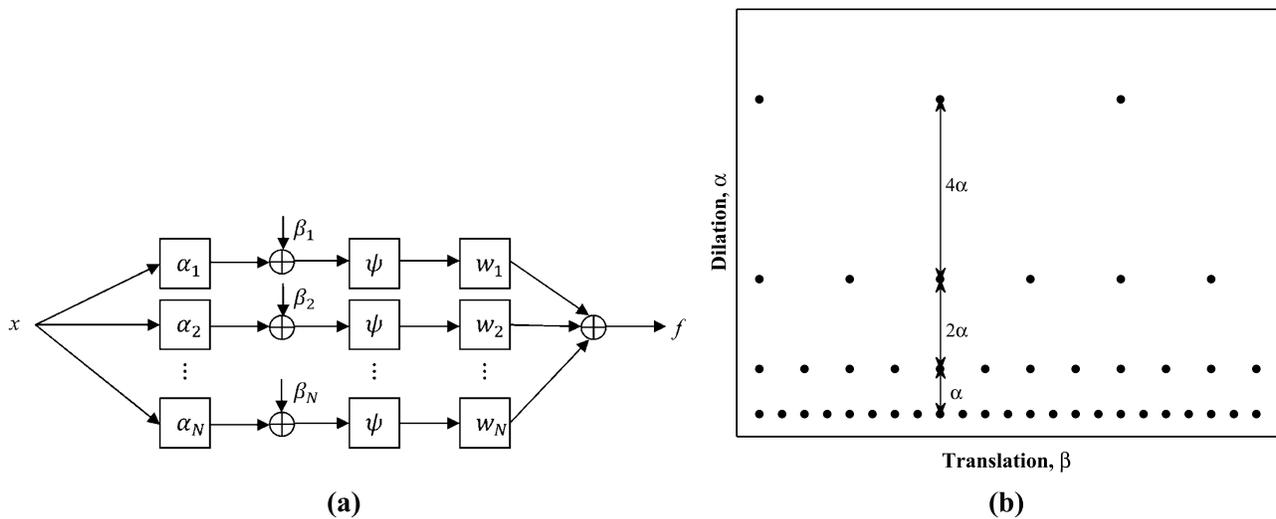


Figure 5. (a) Structure of a $(1 + \frac{1}{2})$ layer wavelet neural network. (b) Dyadic grid discretization of wavelets in the time-scale space.

dyadic grid (of the form shown in Figure 5b) is used as dilation and translation step sizes instead of the regular lattice structure. This initial set of wavelets is truncated by scanning the input data and eliminating those wavelets whose supports do not contain any sample points. Now the wavelet library \mathbf{W} of refined family of wavelets is further normalized for computational convenience. Hence, the wavelet family is

The wavelet coefficients w_i are determined using the least squares method. Now one wavelet ψ_j is eliminated from $f_N(x)$, and the increment in residual is further calculated. Finally, $f_{N-1}(x)$ is created by removing the particular wavelet ψ_j that minimizes the increment in the sum of square residuals. Likewise, more wavelets are reduced one by one until the optimum number L is reached.

$$\mathbf{W} = \left\{ \psi_i : \psi_i(x) = A_i \psi(\alpha_i * (x - \beta_i)), A_i = \left(\sum_{j=1}^M [\psi(\alpha_i * (x - \beta_i))]^2 \right)^{-1/2}, i = 1, 2, \dots, N \right\}, \quad (10)$$

where α_i and β_i correspond to the dilation and translation parameters, respectively, x_1, x_2, \dots, x_M are the training data, and N is the number of wavelets.

[29] 2. For selection of best wavelets, considering each wavelet as a regressor, a set of best wavelets from the wavelet library \mathbf{W} are selected for the regression given by

$$f_L(x) = \sum_{i=1}^L w_i \psi_i(x),$$

where $L \leq N$ is the number of best wavelets or regressors and $w_i \in \mathfrak{R}$.

[30] Since extracting all possible L element subsets of \mathbf{W} and performing a minimization of least squares problem to obtain the w_i will be exhaustive and hence infeasible, heuristic procedures are to be adopted to solve this problem. In the present study, a backward elimination method is used.

[31] In this method, regression is made considering all possible N wavelets available in the wavelet library \mathbf{W} , which can be expressed as

$$f_N(x) = \sum_{i=1}^N w_i \psi_i(x). \quad (11)$$

[32] The optimum number of wavelets L can be decided using the Akaike final prediction error (FPE) criterion written as

$$\text{FPE} = \frac{1}{2M} \left(\frac{1 + n_p/M}{1 - n_p/M} \right) \sum_{j=1}^M [\hat{f}(x_j) - y_j]^2, \quad (12)$$

where n_p is the number of parameters to be determined (all of the dilation and translation parameters considering all of the wavelets), x_j, y_j are the training input and output data respectively and M is the length of training data.

[33] After the elimination of each best contributing wavelet ψ_j , FPE is evaluated. The elimination is stopped when there is only one wavelet remaining in the network. The optimum number of wavelets is chosen as the elimination level corresponding to minimum FPE.

[34] 3. The network gets initialized from the above two steps. A back propagation algorithm is used to enhance the quality of the model further.

[35] The flowchart of the wavelet network structure is shown in Figure 6.

3.4. Final Structure of the Wavelet Network

[36] The structure of the wavelet network as given in equation (9) is slightly modified in order to take the linear

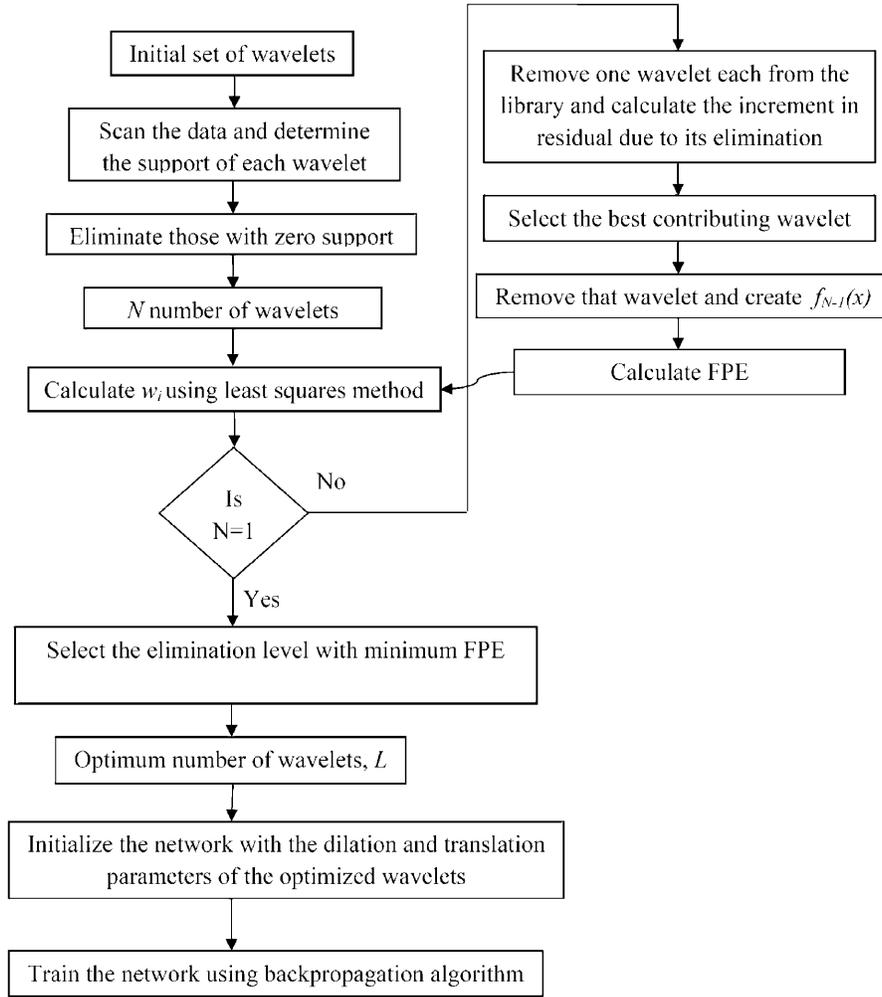


Figure 6. Flowchart of wavelet network.

regions in the time series into account in addition to the nonlinearities. Hence, the final structure of the wavelet network is

$$f(x) = \sum_{i=1}^L w_i \psi[\alpha_i * (x - \beta_i)] + c^T x + b, \quad (13)$$

where $c \in \mathfrak{R}^d$ is the linear coefficient and $b \in \mathfrak{R}$ is the bias term. The modified structure is shown in Figure 7.

3.5. Initialization of Wavelet Network Parameters

[37] The training data is reconstructed into phase space vectors using equation (1) for the chosen embedding dimension m and delay time τ . Thus, the predictive model is of the form

$$x_{j+1} = F(x_j, x_{j-\tau}, x_{j-2\tau}, \dots, x_{j-(m-1)\tau}) \quad (14)$$

in which the reconstructed vectors are the inputs and the corresponding x_{j+1} are the output. The mother wavelet function used for the current study is Mexican hat [Chui, 1992]

of the form $\psi(s) = (1 - s^2)e^{-s^2/2}$. Thus, equation (13) is modified into

$$f(x) = \sum_{i=1}^L w_i \left[1 - (\alpha_i * (x - \beta_i))^2 \right] e^{-(\alpha_i * (x - \beta_i))^2 / 2} + c^T x + b. \quad (15)$$

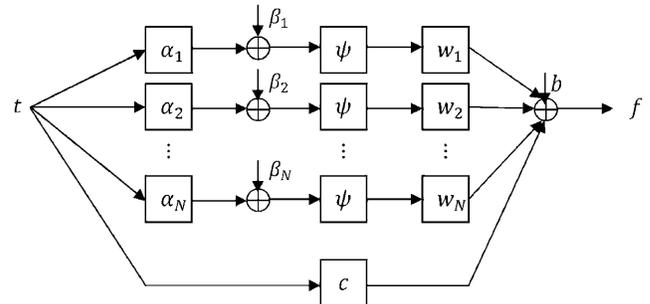


Figure 7. Structure of a modified wavelet neural network.

[38] Initialization of the dilation parameter α and translation parameter β is done by constructing the wavelet library \mathbf{W} as explained in section 3.3. For each combination of α and β (i.e., for all of the N available wavelets), the network parameters w , c , and b are estimated by computing the least squares solution:

$$\begin{bmatrix} [1 - (\alpha_1 * (x_1 - \beta_1))^2] e^{-(\alpha_1 * (x_1 - \beta_1))^2 / 2} & \dots & [1 - (\alpha_N * (x_1 - \beta_N))^2] e^{-(\alpha_N * (x_1 - \beta_N))^2 / 2} & x_1 & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \vdots \\ [1 - (\alpha_1 * (x_M - \beta_1))^2] e^{-(\alpha_1 * (x_M - \beta_1))^2 / 2} & \dots & [1 - (\alpha_N * (x_M - \beta_N))^2] e^{-(\alpha_N * (x_M - \beta_N))^2 / 2} & x_M & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ c \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix}. \quad (16)$$

The optimum number of wavelets is fixed further by employing the backward elimination method and determining the Akaike final prediction error after eliminating each wavelet. Considering one wavelet, there are one dilation parameter, m translation parameters, and one linear parameter to be determined. Hence, for calculating FPE, the total number of parameters n_p for the entire network of N wavelets, after also taking into account the $d + 1$ linear network connections, will be equal to $N(m + 2) + m + 1$. The optimum number of wavelets is chosen as the one with minimum FPE. Sections 4 and 5 describe the application of wavelet network to phase space prediction in known chaotic systems and also in observed streamflow series.

4. Application to Known Chaotic Systems

[39] Two basic chaotic systems, the Henon map and Lorenz three-variable model, are chosen as benchmark problems in order to ensure the applicability of the wavelet network algorithm to nonlinear dynamic streamflow series. The prediction accuracy and the effect of sample size are noted for both the wavelet network method and local approximation method. The Henon map is a less complex dynamic system when compared to the Lorenz model.

4.1. Henon Map

[40] The Henon map [Hénon, 1976] is a two-variable discrete time dynamical system defined by

$$\begin{aligned} x_{n+1} &= y_n + 1 - ax_n^2 \\ y_{n+1} &= bx_n. \end{aligned} \quad (17)$$

The chaotic nature of the map is dependent on the parameters a and b in equation (17). The canonical Henon map with values of $a = 1.4$ and $b = 0.3$ is chaotic in nature. The Henon map is iterated with these parameter values to generate 11,700 data points (approximately equal to the data size of the Basantpur streamflow series). The correlation dimension and maximum Lyapunov exponent of the Henon map are 1.21 and 0.418, respectively.

[41] The x variable of the Henon map is reconstructed for an embedding dimension $m = 2$ and delay time $\tau = 1$

according to equation (1). Of the 11,700 data points, the last 1460 data points are set apart as the testing set for prediction. To study the sensitivity of the prediction algorithm to data size, prediction is attempted for different training data set sizes such as 10,240, 5000, and 1500, keeping the testing data

size the same. The predictive model given by equation (14) is modeled using both the wavelet network algorithm and local approximation method for different training data sets. For wavelet networks, the optimum number of wavelets (i.e., minimum FPE) is obtained as 3. The optimum neighborhood size for local approximation method is fixed at 0.6 of standard deviation. Prediction is done for different lead times up to 10 for analyzing the predictability potential of each method.

[42] The root-mean-square error (RMSE) and correlation coefficient (CC) plots (shown in Figures 8a and 8b, respectively) show that the performance of the local approximation (LA) method is slightly better only for the first data set (with 10,240 points). For the other two data sets with lesser data points, the wavelet network (WN) algorithm is significantly more efficient than the LA method with lesser RMSEs and higher CCs for all lead times. Also, the WN algorithm is almost insensitive to the data size of training set, as can be noticed from the constant RMSE and CC for the three data sets, whereas the efficiency of LA is directly proportional to data size.

4.2. Lorenz Model

[43] The Lorenz system [Lorenz, 1963] is a dynamical continuous system exhibiting chaotic behavior which is much more complex than the Henon map. The equations governing the Lorenz model are

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{aligned} \quad (18)$$

where $\sigma, \rho, \beta > 0$; σ is called the Prandtl number; and ρ is called the Rayleigh number. Usually, $\sigma = 10$, $\beta = 8/3$, and ρ is varied. The system exhibits chaotic behavior for $\rho = 28$. The three state variables x , y , and z form an inhomogeneous “butterfly” structure in phase space. Of the two time scales, the first describes the evolution of the system around the center of each butterfly wing (i.e., unstable fixed point), and the second describes the residence time within one of the butterfly wings. According to Palmer [1993], most of the unpredictability of the Lorenz model arises when the tra-

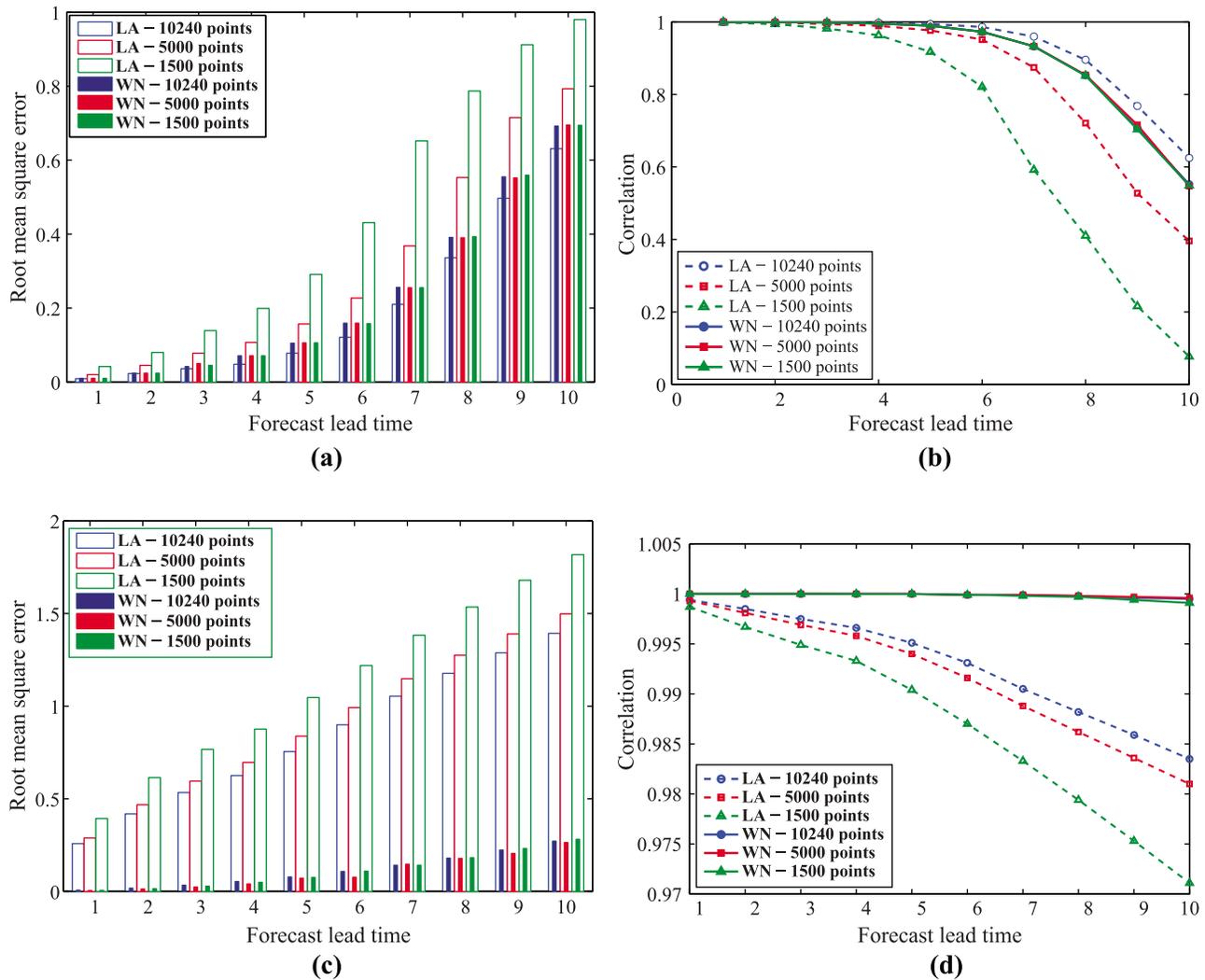


Figure 8. Variation of root-mean-square error (RMSE) and correlation coefficient (CC) over lead time for predictions from the wavelet network (WN) and local approximation (LA) method for different data sizes. (a) RMSE for the Henon map. (b) Correlation coefficient for the Henon map. (c) RMSE for the Lorenz model. (d) Correlation coefficient for the Lorenz model.

jectory switches from one wing to the other wing of the attractor, i.e., near the origin. A set of 11,700 data points is generated with $\sigma = 10$, $\beta = 8/3$, and $\rho = 28$. The correlation dimension and maximum Lyapunov exponent of the Lorenz model are 2.05 and 0.906, respectively.

[44] The x variable of the Lorenz model is reconstructed for an embedding dimension $m = 3$ and delay time $\tau = 2$ according to equation (1). Similar to the procedure followed for Henon data, three data sets of different sizes (10,240, 5000, and 1500) are chosen as training data sets to predict the last 1460 data points. Prediction is done using the LA method and WN method for different lead times up to 10. The root-mean-square error and correlation coefficient plots for different data sets and different lead times for both methods are shown in Figures 8c and 8d, respectively. The efficacy of the WN method is more distinguishable even for the largest data set. The unvarying RMSE and CC for different data sets prove the insensitivity of the WN algorithm to the data

size. The rate of increase of RMSE or decrease of CC over lead time by WN is much less when compared to the LA method, which in turn provides evidence for the increase in predictability. Hence, for a chaotic series, the WN algorithm improves the predictability and also is unaffected by the data size.

[45] Having demonstrated this, the sensitivity of the algorithm for different initial conditions is judged by making predictions for a set of embedding dimensions and delay times. The range of embedding dimension and delay time is chosen as 2–5, thus yielding a total of 16 combinations. Phase space is constructed for all of these combinations for the first data set (10,240 points), and prediction is done for all of these combinations. The box plots of 16 predictions for the first 200 points from both methods are shown in Figures 9a and 9b. The widths of box plots are insignificant for WN method predictions. A comparison of box plot widths (i.e., 75th percentile to 25th percentile) for both methods is

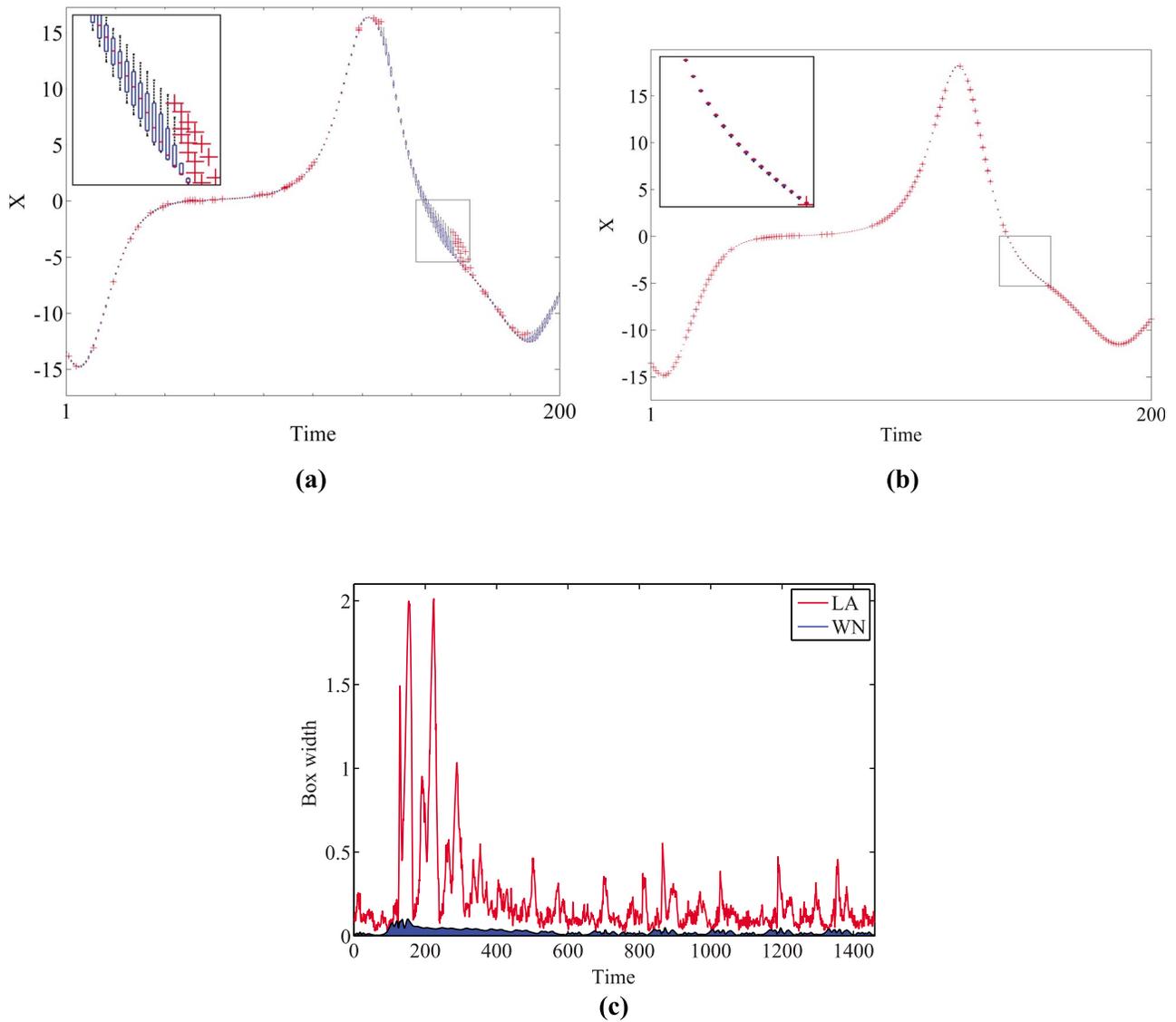


Figure 9. Ensemble prediction of the Lorenz model as box plots for the first 200 data points from (a) local approximation model and (b) wavelet network. The boxes for a particular period are enlarged and shown as insets. (c) Box width of predictions for the testing period from wavelet network and local approximation method.

Table 2. Root-Mean-Square Error (RMSE) and Correlation Coefficient (CC) Between the Observed and Mean Ensemble Daily Streamflow Values for Predictions From the Wavelet Network (WN) and Local Approximation (LA) Methods

	1 day		2 day		3 day	
	RMSE	CC	RMSE	CC	RMSE	CC
<i>Seorinarayan Station</i>						
WN	166.62	0.996	287.29	0.986	373.32	0.974
LA	476.41	0.961	586.73	0.938	633.39	0.926
<i>Basantpur Station</i>						
WN	186.46	0.994	320.30	0.982	394.69	0.971
LA	590.54	0.945	730.09	0.910	796.15	0.890

shown in Figure 9c. The sensitivity of WN to initial conditions (in terms of box width) is trivial when compared to that of the LA method.

5. Application to Observed Streamflow Data

5.1. Optimum Range of Embedding Dimension and Delay Time

[46] The chaos identification methods (described in section 2.3) have confirmed that the daily streamflow series is nonlinear and low-dimensional chaotic. An appropriate range of embedding dimension and delay time is fixed from these results to produce an ensemble of predictions. The ranges of values used for the ensemble prediction are (1) embedding dimension, 3–10 and (2) delay time, 60–100. The range is fixed the same for streamflow series of both stations. From all of these available initial conditions (i.e., combinations of embedding dimension and delay time), the optimum parameter values are selected on the basis of the minimum GCV value. The phase space reconstructed training sets of both streamflow series are fitted with a local approximation model, and the prediction error (in terms of GCV value) is calculated for each of the available combinations. All of the parameter combinations falling under 10% of the lowest GCV value are selected as the relevant ones. The best parameter combinations which give minimum GCV values are (1) embedding dimension, 4–7 and (2) delay time, 70–95. These best combinations are further employed for prediction using both wavelet network and local approximation methods.

5.2. Ensemble Prediction

[47] Now phase space reconstructions of the training data of both stations are done for each of the best combinations of embedding dimension m and delay time τ obtained. The optimum number of wavelets is obtained for both cases on the basis of the minimum FPE criterion (as described in sections 3.3 and 3.5). The optimum number of wavelets ranges between three and five wavelets for all of the relevant combinations of embedding dimension m and delay time τ .

[48] The network parameters are initialized with optimum number of wavelets in the network structure. The initialized network is trained with the training data sets (phase reconstructed) applying a back propagation algorithm. Training is done for all relevant combinations of embedding dimensions and delay times separately to extract the respective

network parameters and the nonlinear function f . This gives a set of wavelet networks with different parameters, nonlinear functions, and number of wavelets, which is further used to create an ensemble of predictions for the future.

[49] Prediction is done for the next 4 years (from June 2000 to May 2004, i.e., 1461 data points) using these optimized WN and LA methods for 1 day, 2 day, and 3 day lead times for both stations. The root-mean-square error and correlation coefficient between the observed and mean ensemble daily streamflow for both stations are shown in Table 2. Substantially low RMS errors and comparatively higher correlations for all of the lead times show the superiority of wavelet networks in predicting the chaotic series. In Figure 10, comparison is made of the variations of observed streamflow values and mean ensemble streamflow values from the LA and WN methods for a low-flow year (June 2002 to May 2003) and a high-flow year (June 2003 to May 2004) for all of the lead times for Seorinarayan station as an illustration. The extreme daily flows are better captured by the WN method, although the deviation increases with lead time. The absolute deviation between the mean ensemble and actual flow for WN is about half of that of LA as can be noted from Figure 10.

[50] Detailed analyses are done by comparing the cumulative probability distributions (CDF) of ensembles and the observed series for the two methods and are shown in Figures 11a–11f for all three lead times. In all of the cases, the ensembles from both methods were able to catch the observed streamflow probabilities well with its range. Despite this, as the lead time increases, the widths of the ensemble CDFs also seem to be increasing in both methods, indicating the limit in predictability. It is observed that the value of streamflow at which the CDF reaches the value of 1 is around $0.8-1 \times 10^6 \text{ m}^3$ for the LA approach, while that for observed flow is at around $2.3 \times 10^6 \text{ m}^3$, which shows a poor prediction of extreme high flow events by the LA method.

[51] Figures 12, 13, 14, 15, 16, and 17 show a comparison between the two methods for mean, variation or standard deviation, and skewness of daily flows over a year for both stations. The absolute deviations between the observed and the mean ensemble of the statistic are shown as bar plots in Figures 12–17. The wavelet network ensembles are able to capture both the yearly mean and yearly standard deviation of daily flows within the range of prediction, while local approximation fails to predict the unusual variations of streamflow. Even the skewness of the series is also closely captured by the wavelet networks, as can be inferred from Figures 16 and 17. The absolute deviations of all statistics (given as bar plots in Figures 12–17) for the WN method are always smaller than those for the LA method. The above results have proved WN to be a much efficient method when compared to the traditional LA approach.

5.3. Uncertainty in Initial Conditions

[52] The prediction uncertainty due to the sensitivity to initial conditions can be determined through the box plot widths of the daily flows. The width of box plots is a measure of the divergence of the trajectories starting from different initial conditions. Figures 18a, 18b, 19a, and 19b show the variation of box plot widths for a low-flow and a high-flow year, respectively, from which it is clear that the WN approach is

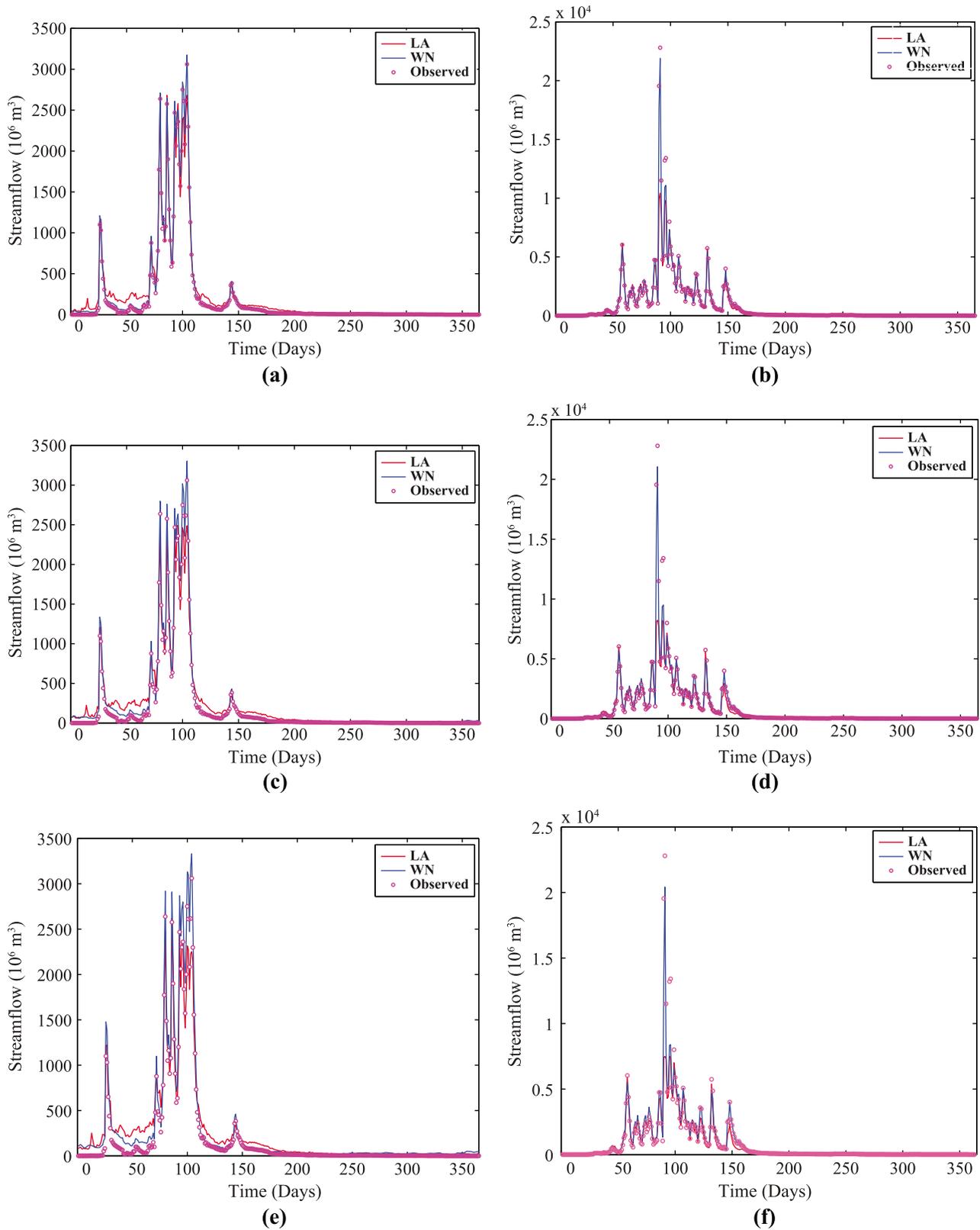


Figure 10. Comparison of mean ensemble predictions from the WN and LA approaches with observed streamflow values for a low-flow period and a high-flow period: (a) 1 day lead prediction for low-flow year, (b) 1 day lead prediction for high-flow year, (c) 2 day lead prediction for low-flow year, (d) 2 day lead prediction for high-flow year, (e) 3 day lead prediction for low-flow year, and (f) 3 day lead prediction for high-flow year. The duration of the period is taken from 1 June to 31 May.

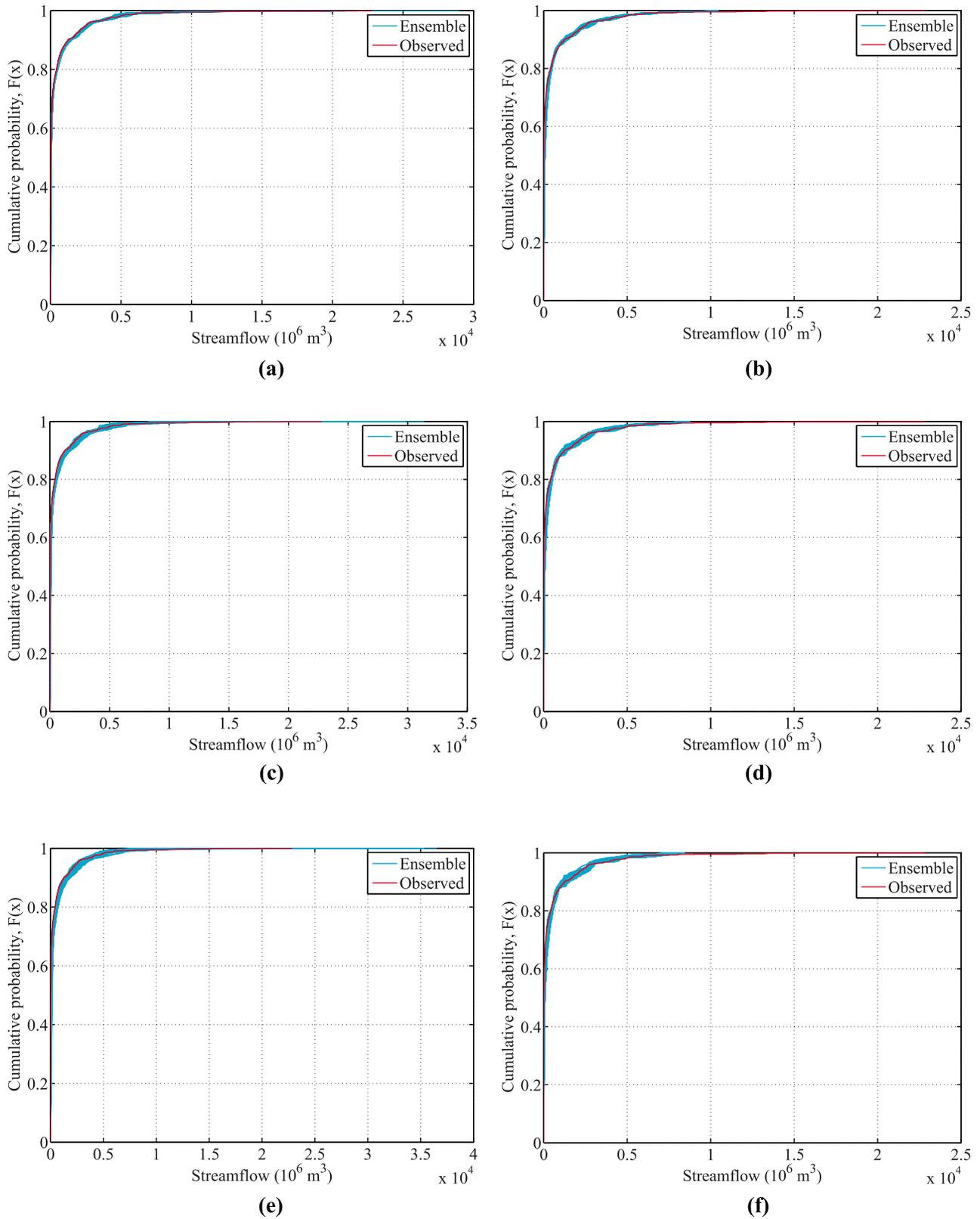


Figure 11. Cumulative density functions of Seorinarayan daily streamflow for the testing period: (a) 1 day lead prediction using WN, (b) 1 day lead prediction using LA, (c) 2 day lead prediction using WN, (d) 2 day lead prediction using LA, (e) 3 day lead prediction using WN, and (f) 3 day lead prediction using LA. The ensemble CDFs (solid blue lines) and also the observed streamflow CDF (solid red line) are shown.

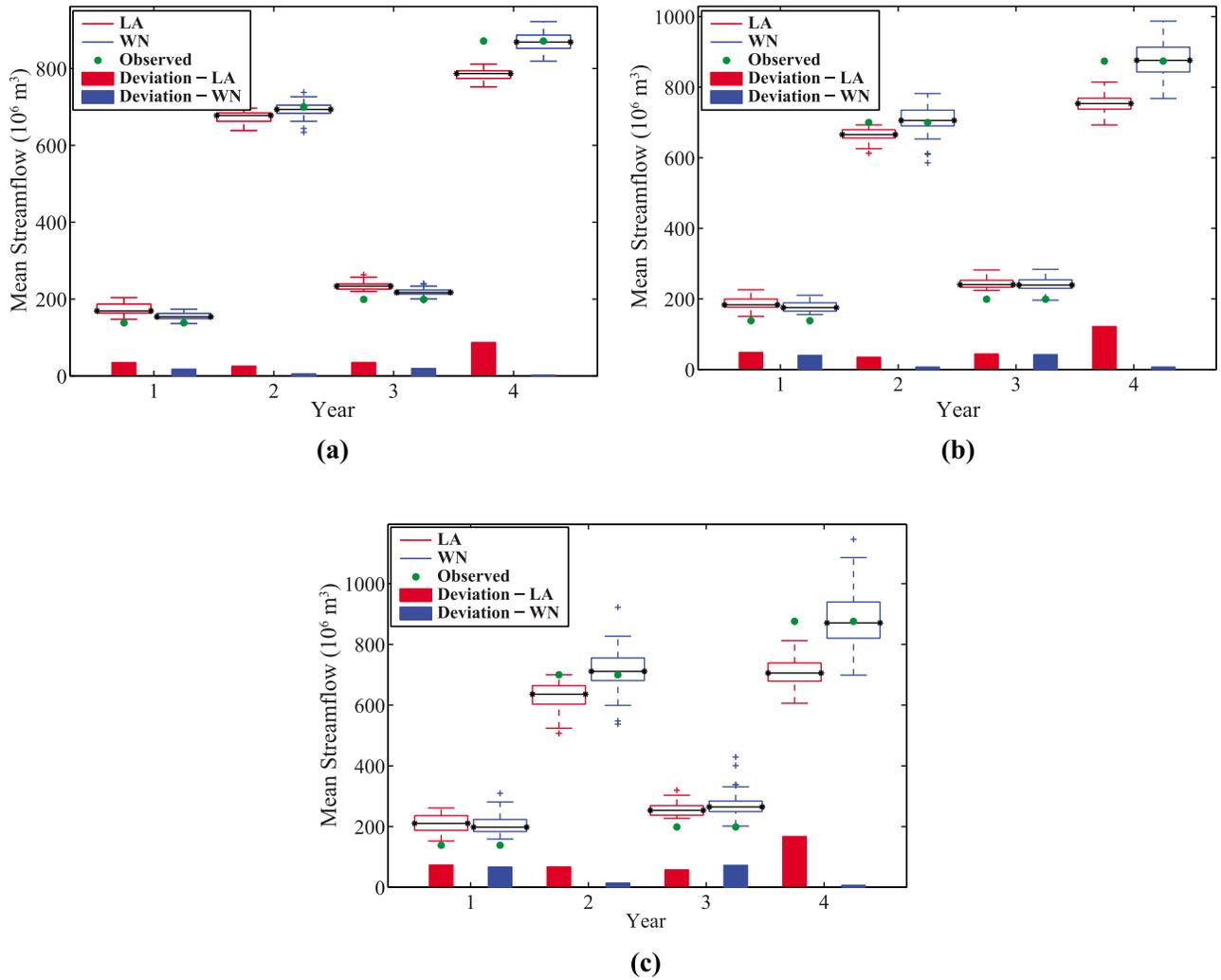


Figure 12. Comparison of box plots of the average daily streamflow values of the ensembles from the LA and WN methods for the 4 year testing period for Seorinarayan station: (a) 1 day lead prediction, (b) 2 day lead prediction, and (c) 3 day lead prediction. The observed average daily streamflow values are also shown. The absolute deviations between the mean ensemble values and the corresponding observed values are shown as bar plots.

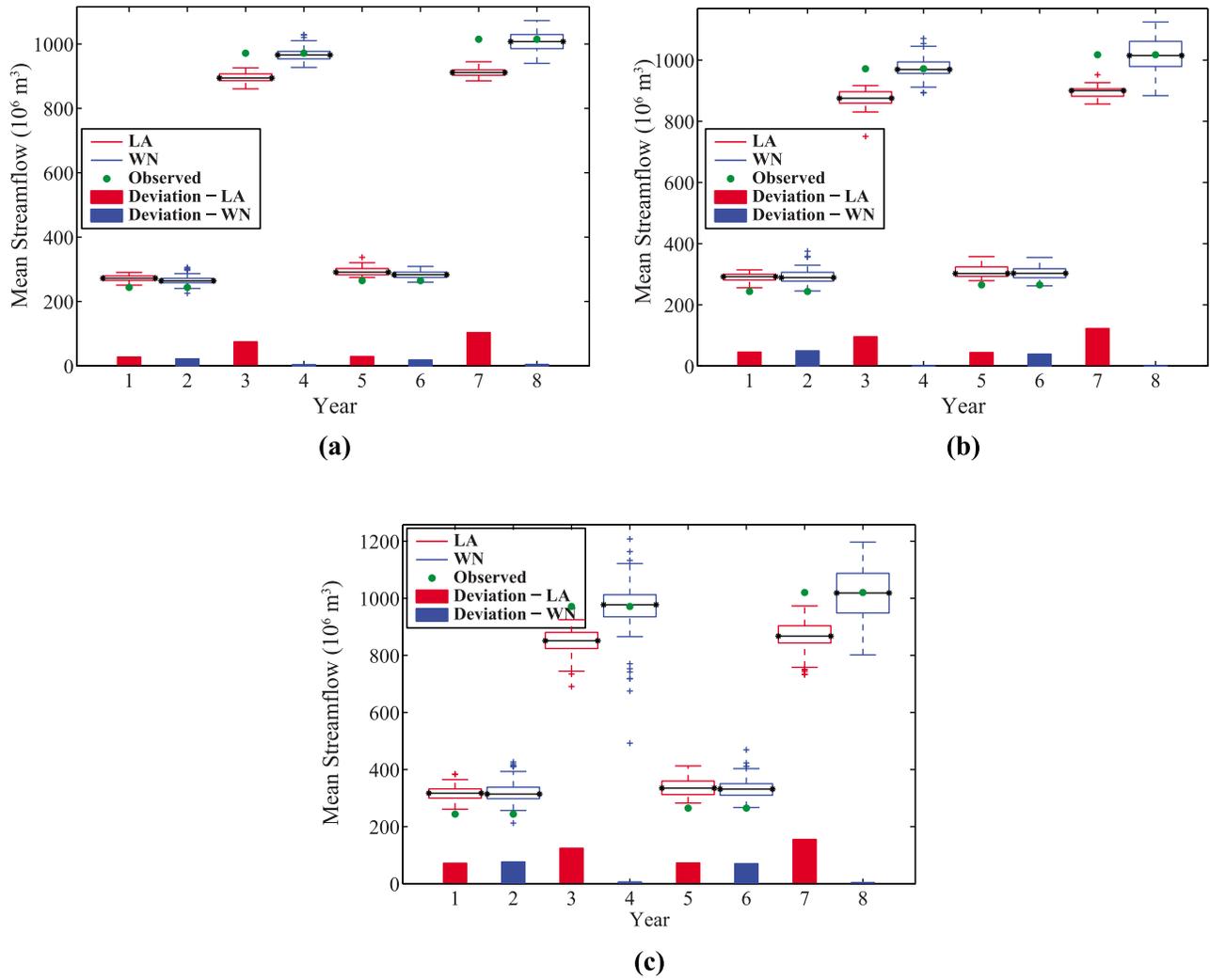


Figure 13. Same as Figure 12 but for Basantpur station.

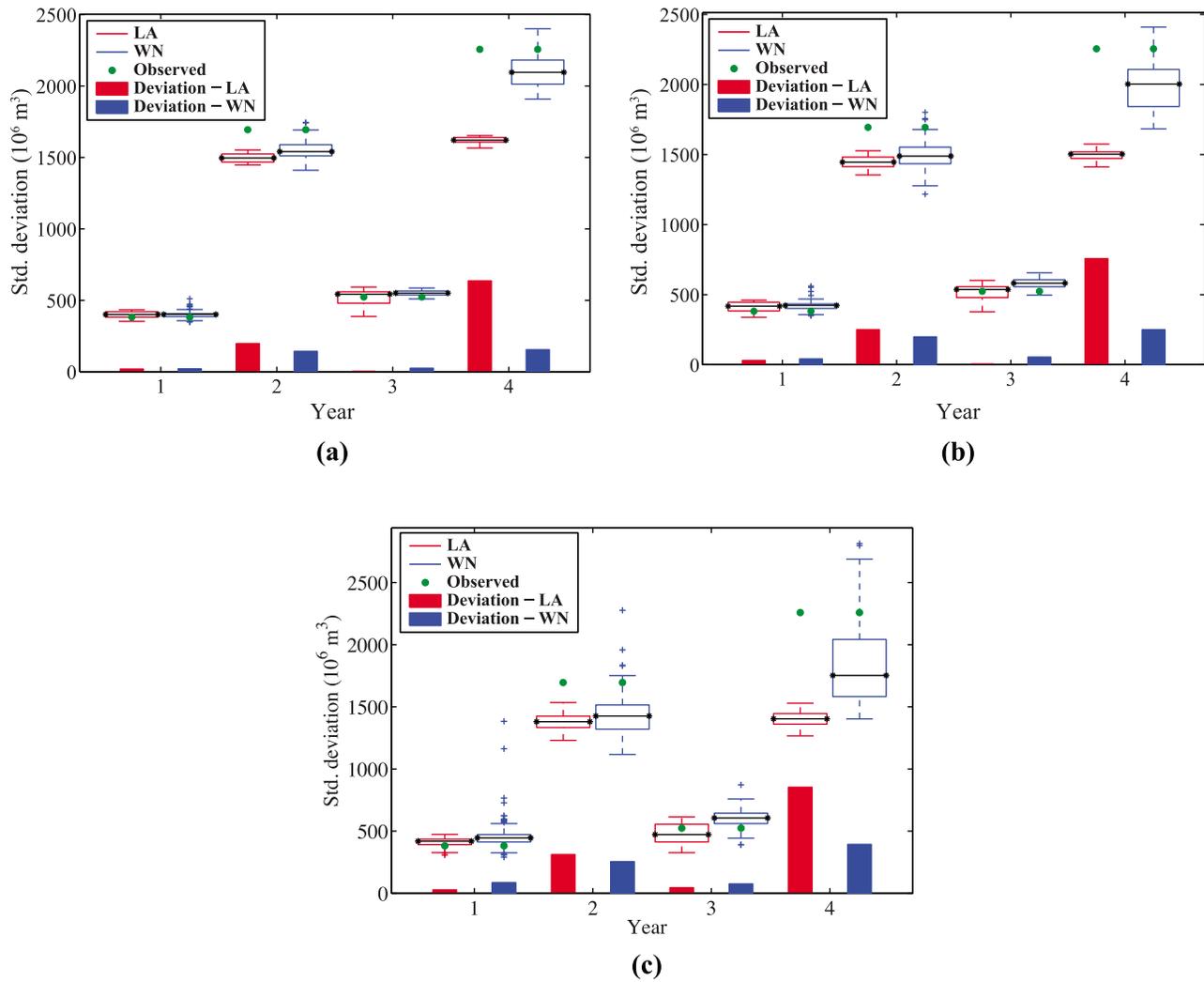


Figure 14. Comparison of box plots of the average standard deviation values of the ensembles from the LA and WN methods for the 4 year testing period for Seorinarayan station: (a) 1 day lead prediction, (b) 2 day lead prediction, and (c) 3 day lead prediction. The observed average standard deviations are also shown. The absolute deviations between the mean ensemble values and the corresponding observed values are shown as bar plots.

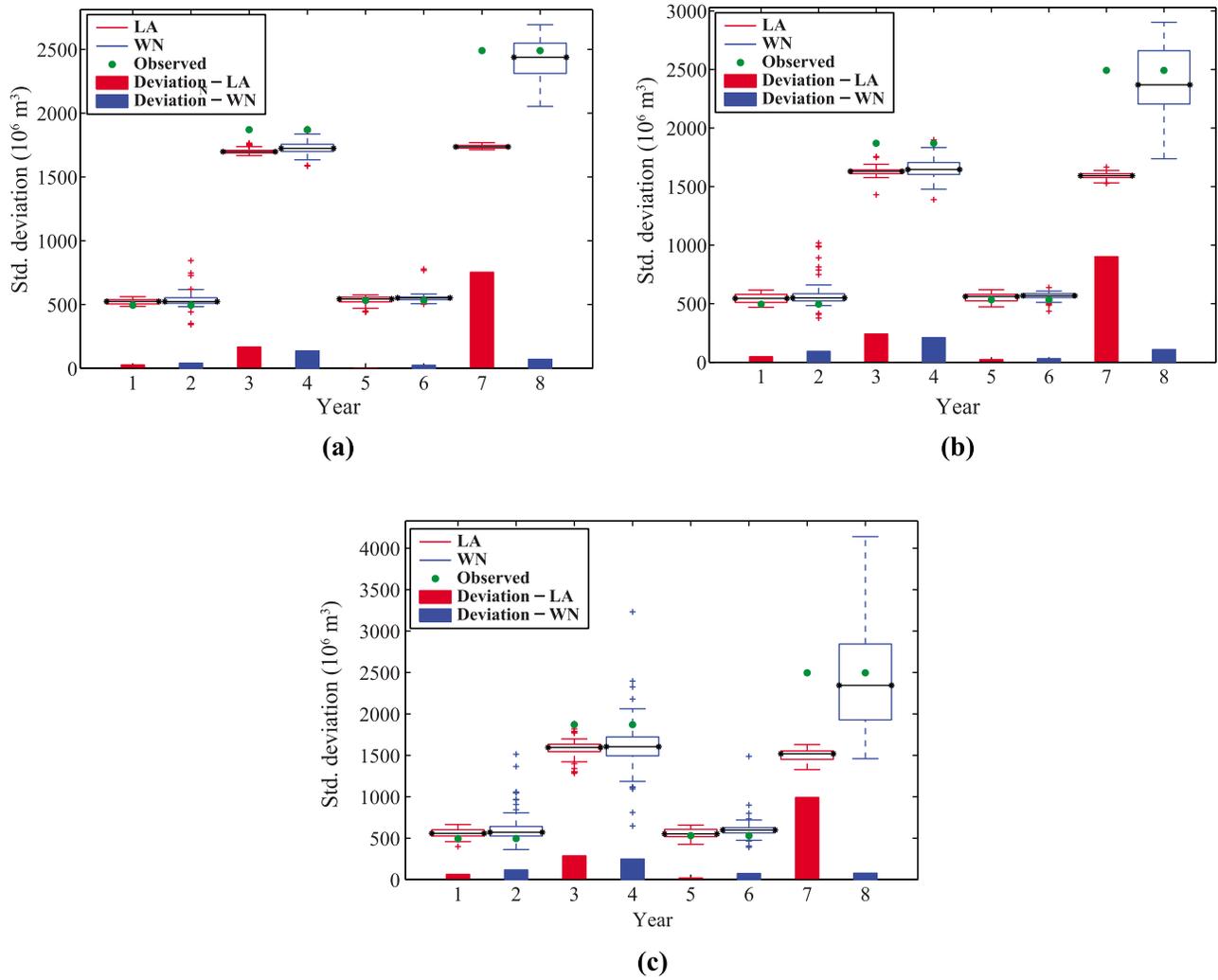


Figure 15. Same as Figure 14 but for Basantpur station.

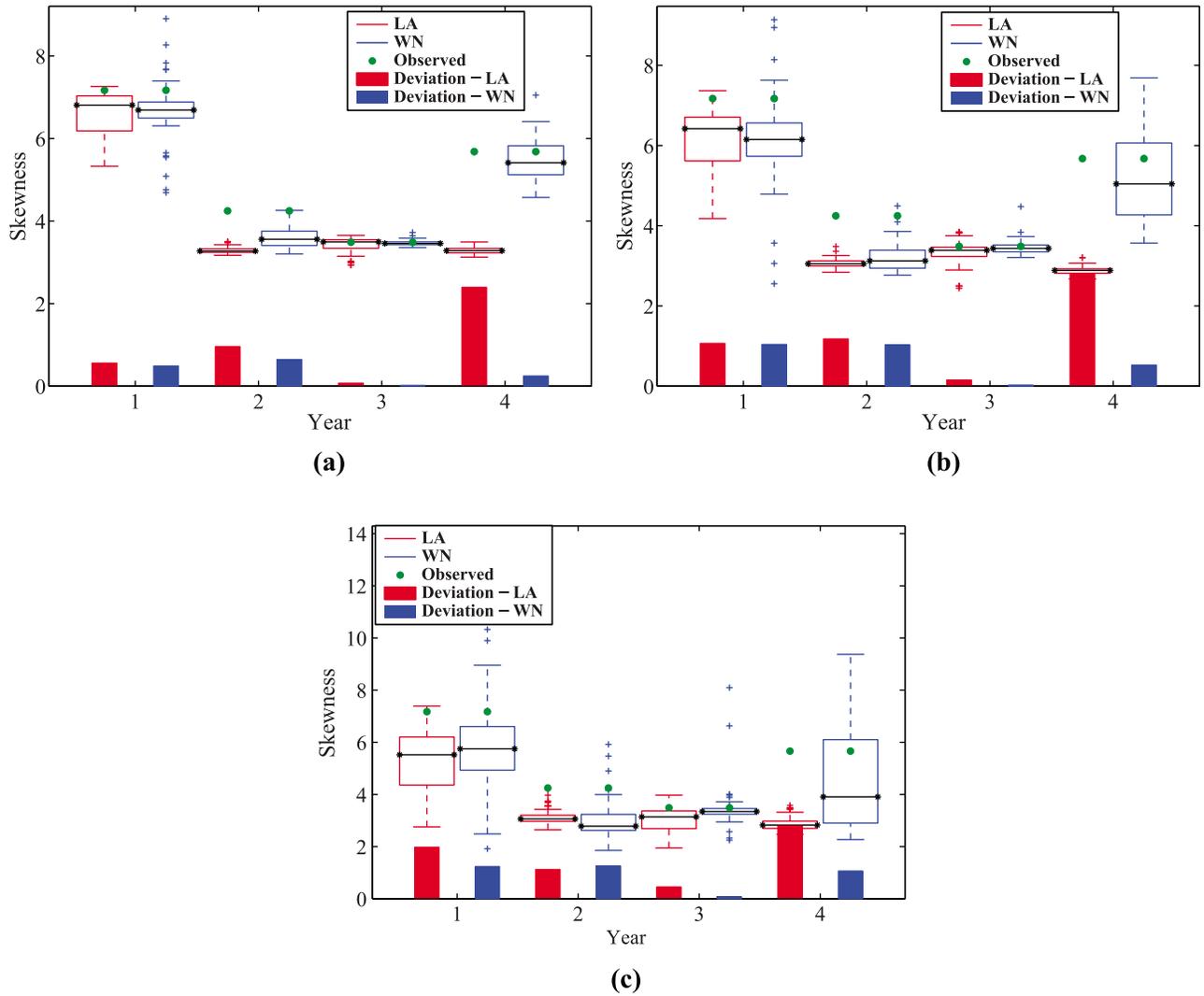


Figure 16. Comparison of box plots of the average skewness of the ensembles from the LA and WN methods for the 4 year testing period for Seorinarayan station: (a) 1 day lead prediction, (b) 2 day lead prediction, and (c) 3 day lead prediction. The observed average skewness are also shown. The absolute deviations between the mean ensemble values and the corresponding observed values are shown as bar plots.

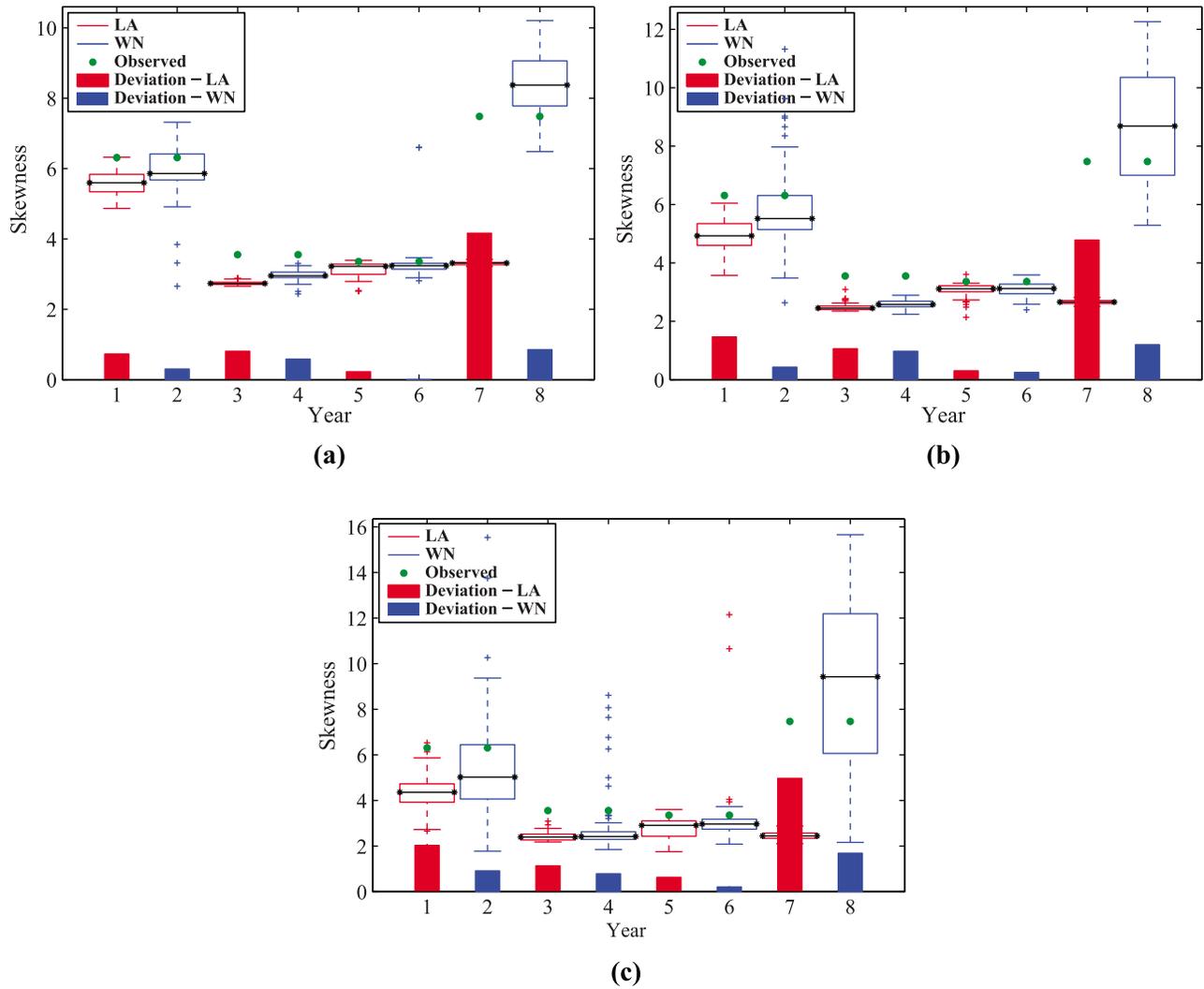


Figure 17. Same as Figure 16 but for Basantpur station.

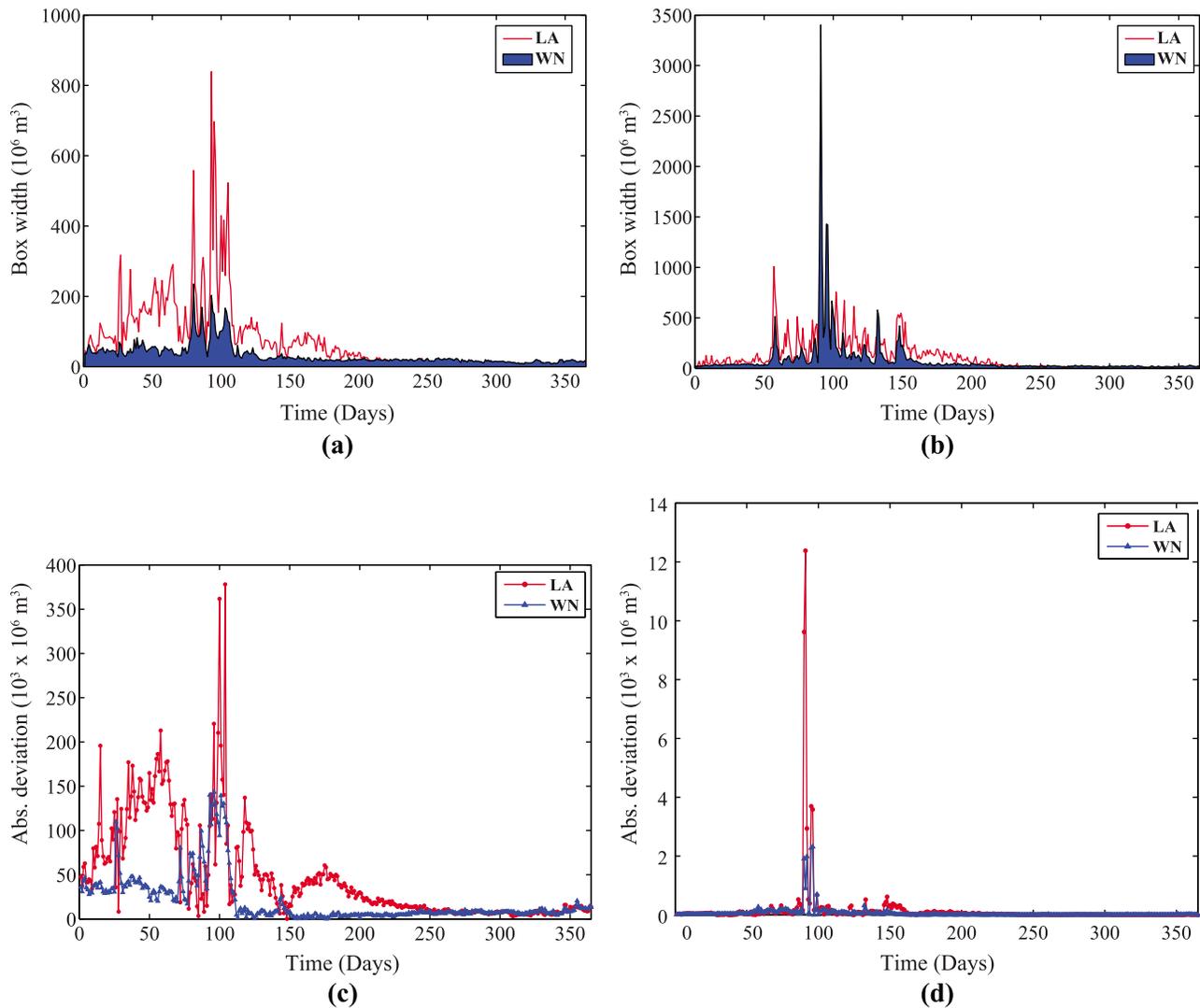


Figure 18. Comparison of box widths of predictions from the WN and LA methods for (a) low-flow year and (b) high-flow year. Comparison of absolute deviations between observed streamflow values and mean ensemble prediction from both methods for (c) low-flow year and (d) high-flow year. The results shown above are for 1 day lead prediction. The duration of the year is taken from 1 June to 31 May.

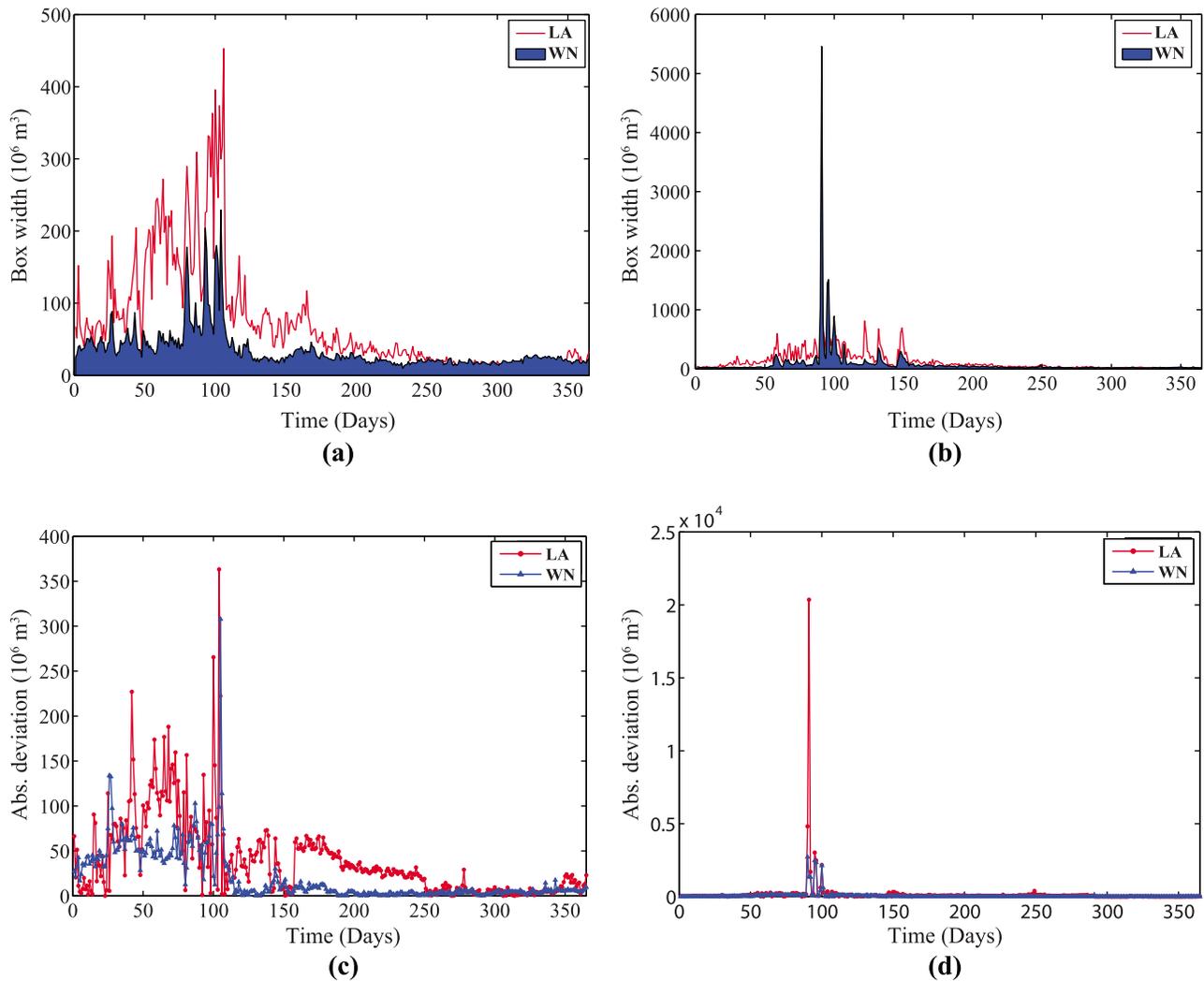


Figure 19. Same as Figure 18 but for Basantpur station.

less sensitive to the change in the initial conditions, particularly in low-flow periods. Except during a few very high flow days, the box plot widths for the WN method are always smaller than that of the LA approach. It is to be noted that, even though the prediction uncertainty is high during high-flow days, the absolute deviations between the mean ensemble and the observed streamflow for the particular days are very low for the WN method (Figures 18c, 18d, 19c, and 19d).

5.4. Discussion

[53] The predictability of the daily streamflow chosen is only about 7 days by the Lyapunov exponent. However, as demonstrated by Dhanya and Nagesh Kumar [2010], the predictability of daily rainfall (averaged over the basin) over the same basin can be increased to 40 days. The low predictability obtained for streamflow may be due to the influence of other processes such as infiltration or evaporation. Additionally, Dhanya and Nagesh Kumar [2010] have pointed out that spatial averaging alone can increase the predictability of a chaotic series. Another point to be noted here is that the 40 day predictability limit observed was for the daily rainfall series averaged over the entire

basin, whereas the daily streamflow series considered was for a single station.

[54] Since a basic assumption in the implementation of any chaos identification method is that the time series is infinite, the restriction in data size limits the accurate estimation of parameters (m and τ) and also the nonlinear prediction accuracy. While Smith [1988], Procaccia [1988], and Nerenberg and Essex [1990] pointed out that the minimum number of data points varies exponentially with the embedding dimension, i.e., 42^m , 10^m , and $10^{2+0.4m}$, respectively, Kurths and Herzel [1987] used only 640 points. Jayawardena and Lai [1994] have demonstrated on synthetic series that the number of data points should be ≈ 1200 for an accurate estimation of chaotic parameters. Sivakumar et al. [1999a] suggested that a minimum of 1500 data points is required to reveal the rainfall dynamics. However, the application of wavelet networks on the Henon and Lorenz system (Figure 8) demonstrates that the proposed method is insensitive to the number of data points since the performance measures are almost constant for the different data sets with varying number of data points.

[55] In spite of the chaotic nature with limited predictability, the predictions from both methods are capable of

Table 3. Mean Absolute Deviations (10^6 m^3) Over a Year Between the Observed and Mean Ensemble Daily Streamflow Values for Predictions From Wavelet Network (WN) and Local Approximation (LA) Methods

Year	1 day Lead		2 day Lead		3 day Lead	
	WN	LA	WN	LA	WN	LA
<i>Seorinarayan Station</i>						
1	18.3502	40.21299	39.88485	52.89304	61.8361	60.5545
2	57.0751	77.10709	101.7196	120.0655	133.3843	150.4509
3	20.02199	47.39537	42.31594	60.90499	66.1047	72.9915
4	61.00386	152.2715	116.1334	198.3819	161.6039	225.88
<i>Basantpur Station</i>						
1	23.05476	43.2012	50.04199	63.42036	78.8854	81.1098
2	56.90769	124.4384	105.2515	170.8619	138.2082	200.814
3	20.20842	38.9846	40.92355	54.1439	68.0545	74.9838
4	64.97456	168.386	113.9161	236.1052	150.7869	274.8348

catching up the variations in streamflow for all of the lead times considered. It can be noticed that, as the lead time increases, the divergence from the observed value increases, thereby reducing the reliability of predictions. Such a behavior is expected in regular chaotic systems in which the system will become almost unpredictable as the lead time increases, irrespective of the prediction method employed. However, the divergence is much lower in predictions by wavelet network method for all of the lead times as can be noted from Figure 10. Hence, the predictability also depends on the prediction algorithm employed.

[56] The local approximation method entirely depends on the past observed flows. The reason the failure of the LA approach in predicting the extremely high flows is the absence of such flows in the past data. In contrast, the WN approach has no such limitations and has the potential to predict unusually high flows also. The root-mean-square error values from WN predictions are about half of those from the LA approach. Table 3 gives a comparison of the absolute deviations of mean ensemble values of the WN and LA methods from the observed values for three lead times. The deviations of the WN approach are always lesser compared to those of the LA approach. The differences between the deviations from the two approaches (deviation from LA minus deviation from WN) are larger for high-flow years, which indicates that the WN approach is most efficient in predicting the unexpected high flows.

Table 4. Mean Box Widths (10^6 m^3) Over a Year for Predictions From the Wavelet Network (WN) and Local Approximation (LA) Methods

Year	1 day Lead		2 day Lead		3 day Lead	
	WN	LA	WN	LA	WN	LA
<i>Seorinarayan Station</i>						
1	24.3829	64.7133	48.2513	86.4905	69.0157	97.2802
2	75.4824	101.0025	136.3	138.4946	185.9075	151.2409
3	33.1983	74.6065	64.8274	101.8228	92.8164	118.8653
4	88.3658	107.5436	166.9715	151.8213	236.6322	180.5148
<i>Basantpur Station</i>						
1	30.6059	71.2748	61.8004	100.4692	97.6473	129.0564
2	69.9582	132.2318	130.6388	175.0348	184.6156	213.5233
3	34.4805	72.3271	68.7484	103.7795	107.947	142.8995
4	84.9296	116.2867	159.9328	172.5184	225.3146	217.4818

Table 5. Mean Absolute Deviations (10^6 m^3) Over a Year Between the Observed and Predicted Daily Streamflow Values From a Single Initial Condition From the Wavelet Network (WN) and Local Approximation (LA) Methods

Year	1 day Lead		2 day Lead		3 day Lead	
	WN	LA	WN	LA	WN	LA
<i>Seorinarayan Station</i>						
1	27.6734	64.9495	60.5542	89.8518	95.1653	106.8514
2	76.7445	101.1523	137.4347	148.6521	186.9305	198.546
3	30.3038	82.8967	56.2994	121.2774	80.8663	143.6372
4	99.6162	179.596	182.2908	268.1606	241.9138	322.0757
<i>Basantpur Station</i>						
1	58.8444	65.1224	122.1445	94.7069	159.1572	116.82
2	70.9586	173.0133	128.2393	242.4577	174.8071	283.2043
3	37.6463	83.5924	72.0296	120.4204	105.0406	154.9974
4	87.9809	225.9265	159.9286	308.0498	218.1132	362.9765

[57] The prediction uncertainty due to the sensitiveness to initial conditions is captured by creating ensembles with different initial conditions. A comparison of the box widths (measure of uncertainty) from ensemble predictions from both approaches is shown in Table 4. The ensembles from the WN approach are less uncertain except for the high-flow days. While prediction uncertainty (i.e., box width) is high during high-flow periods, the mean ensemble is much nearer to the observed value; that is, the absolute deviations are much lower during those periods. Hence, WN trades off between these two measures during high-flow periods.

[58] Table 5 shows the mean absolute deviations from a single prediction instead of an ensemble approach. The initial condition is created from an optimum embedding dimension of 6 (from the correlation dimension method) and a delay time of 74 (from the autocorrelation method). An educated comparison of absolute deviations of mean ensemble (Table 3) and those of single prediction (Table 5) from the observed value shows that the ensemble approach is 50% more efficient than the single prediction. Even for single prediction, the WN approach has proved to be 30%–50% more superior than the LA approach. Hence, as rightly stated by *Smith* [2000], an ensemble of predictions with different initial conditions from a perfect model can accurately reflect the likelihood of occurrence of various future conditions.

6. Concluding Remarks

[59] The predictability of a chaotic series is limited to a few future time steps because of its sensitivity to initial conditions and the exponential divergence of the trajectories. Despite this limit in predictability, which is the basic characteristic of chaotic system, the prediction skill can be considerably improved by adopting an ensemble approach with various initial conditions and with efficient methods with powerful localization properties. In wavelet networks, the benefits of both neural network (a global approach) and wavelet decomposition (a local approach) are combined. While neural network will give a good approximation of the nonlinear function, wavelet decomposition will provide the high-frequency local details. Such a network can efficiently capture the underlying dynamics of a chaotic system when compared with the traditional local approximation method, which works on the nearest-neighbor approach.

[60] The present study was aimed at generating an ensemble of predictions by employing the wavelet network method and examining whether the predictability of a chaotic series can be improved. Daily streamflow data at Seorinarayan and Basantpur stations in Mahanadi basin, India, were considered for the study. The data were analyzed for any chaotic behavior employing various techniques. A positive Lyapunov exponent of 0.167 and embedding dimension 6–7 obtained from correlation dimension, the false nearest-neighbor method, and the nonlinear prediction method confirm the existence of a low-dimensional chaotic attractor in the streamflow series. These results suggest that the seemingly irregular behavior of streamflow series can be better explained through a chaotic framework. A range of embedding dimensions and delay times were selected for the prediction on the basis of the outputs from these techniques.

[61] Predictions were done employing wavelet networks and were compared with those from the local approximation method. A good initialization of the wavelet parameters (dilation and translation), network parameters, and the number of wavelets was done using the least squares method, backward elimination, and the Akaike final prediction error. The optimum number of wavelets was in the range of 3–5 for all of the combinations of embedding dimension m and delay time τ . The trained networks with different m and τ were used to generate an ensemble of predictions with lead times 1, 2, and 3 days.

[62] A comparison of these predictions with those from the local approximation method indicates that wavelet networks provide more reliable predictions by capturing the mean, standard deviation and also skewness of the observed data. While the local approximation method failed to predict the sudden high-frequency high-flows, the microscopical behavior of the wavelet network due to its different dilations and translations facilitates the near-accurate prediction of such flows. The ensembles generated from the wavelet networks are of very less spread (except for very high flow days), hence confirming the reliability in prediction. Even so, it should be noted that as the lead time increases, the ensemble spread also increases, thereby decreasing the reliability. However, when compared with the local approximation results, the widening is much less for higher lead times.

[63] The nonlinear prediction method based on the ensemble approach with wavelet networks considerably improves the prediction efficiency and hence is typically suitable for modeling and assessing the underlying dynamics of the complex chaotic streamflow process. The predictability of a chaotic series is therefore improved by combining the global and local approaches. The predictability can be further improved using a multivariate wavelet network analysis by taking information from other climatic and atmospheric indices which dominantly influence the streamflow dynamics.

[64] **Acknowledgments.** The authors sincerely acknowledge the associate editor and three anonymous reviewers for their valuable suggestions and comments that helped us to refine the conceptual aspects of the subject and hence enabled us to improve the presentation.

References

Alarcon-Aquino, V., E. S. García-Treviño, R. Rosas-Romero, and J. F. Ramirez-Cruz (2005), Learning and approximation of chaotic time series using wavelet-networks, paper presented at Sixth Mexican International

- Conference on Computer Science, Inst. of Electr. and Electron. Eng., Puebla, Mexico.
- Asokan, S. M., and D. Dutta (2008), Analysis of water resources in the Mahanadi River basin, India under projected climate conditions, *Hydrol. Processes*, *22*, 3589–3603, doi:10.1002/hyp.6962.
- Badii, R., G. Broggi, B. Derighetti, M. Ravani, S. Ciliberto, A. Politi, and M. A. Rubio (1988), Dimension increase in filtered chaotic signals, *Phys. Rev. Lett.*, *60*, 979–982, doi:10.1103/PhysRevLett.60.979.
- Bakshi, R. B., and G. Stephanopoulos (1993), Wave-net: A multiresolution, hierarchical neural network with localized learning, *AIChE J.*, *39*(1), 57–81.
- Cao, L., Y. Hong, H. Fang, and G. He (1995), Predicting chaotic time series with wavelet networks, *Physica D*, *85*, 225–238, doi:10.1016/0167-2789(95)00119-O.
- Casdagli, M. (1989), Nonlinear prediction of chaotic time series, *Physica D*, *35*, 335–356, doi:10.1016/0167-2789(89)90074-2.
- Casdagli, M. (1991), Chaos and deterministic versus stochastic nonlinear modeling, *J. R. Stat. Soc., Ser. B*, *54*, 303–324.
- Chennaoui, A., K. Pawelzik, W. Liebert, H. G. Schuster, and G. Pfister (1990), Attractor reconstruction from filtered chaotic time series, *Phys. Rev. A*, *41*, 4151–4159, doi:10.1103/PhysRevA.41.4151.
- Chui, C. K. (1992), *An Introduction to Wavelets*, Academic, San Diego, Calif.
- Craven, P., and G. Wahba (1978), Smoothing noise data with spline functions, *Numer. Math.*, *31*, 377–403, doi:10.1007/BF01404567.
- Daubechies, I. (1988), Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, *41*, 909–996, doi:10.1002/cpa.3160410705.
- Dhanya, C. T., and D. Nagesh Kumar (2010), Nonlinear ensemble prediction of chaotic daily rainfall, *Adv. Water Resour.*, *33*, 327–347, doi:10.1016/j.advwatres.2010.01.001.
- Duffin, R. J., and A. C. Schaeffer (1952), A class of nonharmonic Fourier series, *Trans. Am. Math. Soc.*, *72*(2), 341–366, doi:10.1090/S0002-9947-1952-0047179-6.
- Elshorbagy, A., S. P. Simonovic, and U. S. Panu (2002), Estimation of missing streamflow data using principles of chaos theory, *J. Hydrol.*, *255*, 123–133, doi:10.1016/S0022-1694(01)00513-3.
- Farmer, J. D., and J. J. Sidorowich (1987), Predicting chaotic time series, *Phys. Rev. Lett.*, *59*, 845–848, doi:10.1103/PhysRevLett.59.845.
- García-Treviño, E. S., and V. Alarcon-Aquino (2006), Chaotic time series approximation using iterative wavelet-networks, in *Proceedings of the 16th IEEE International Conference on Electronics, Communications and Computers*, pp. 19–26, Inst. of Electr. and Electron. Eng., New York.
- Grassberger, P., and I. Procaccia (1983), Measuring the strangeness of strange attractors, *Physica D*, *9*, 189–208, doi:10.1016/0167-2789(83)90298-1.
- Hénon, M. (1976), A two-dimensional mapping with a strange attractor, *Commun. Math. Phys.*, *50*(1), 69–77, doi:10.1007/BF01608556.
- Islam, M. N., and B. Sivakumar (2002), Characterization and prediction of runoff dynamics: A nonlinear dynamical view, *Adv. Water Resour.*, *25*, 179–190, doi:10.1016/S0309-1708(01)00053-7.
- Jayawardena, A. W., and F. Lai (1994), Analysis and prediction of chaos in rainfall and streamflow time series, *J. Hydrol.*, *153*, 23–52, doi:10.1016/0022-1694(94)90185-6.
- Kantz, H. (1994), A robust method to estimate the maximal Lyapunov exponent of a time series, *Phys. Lett. A*, *185*, 77–87, doi:10.1016/0375-9601(94)90991-1.
- Karunasinghe, D. S. K., and S. Y. Liang (2006), Chaotic time series prediction with a global model: Artificial neural network, *J. Hydrol.*, *323*, 92–105, doi:10.1016/j.jhydrol.2005.07.048.
- Kennel, M. B., R. Brown, and H. D. I. Abarbanel (1992), Determining embedding dimension for phase-space reconstruction using a geometric method, *Phys. Rev. A*, *45*, 3403–3411, doi:10.1103/PhysRevA.45.3403.
- Khan, S., A. R. Ganguly, and S. Saigal (2005), Detection and predictive modeling of chaos in finite hydrological time series, *Nonlinear Processes Geophys.*, *12*(1), 41–53, doi:10.5194/npg-12-41-2005.
- Khan, S., S. Bandyopadhyay, A. R. Ganguly, S. Saigal, D. J. Erickson III, V. Protopoulos, and G. Ostrouchov (2007), Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data, *Phys. Rev. E*, *76*(2), 026209.
- Kugarajah, T., and Q. Zhang (1995), Multidimensional wavelet frames, *IEEE Trans. Neural Networks*, *6*(6), 1552–1556, doi:10.1109/72.471353.
- Kurths, J., and H. Herzel (1987), An attractor in a solar time series, *Physica D*, *25*, 165–172, doi:10.1016/0167-2789(87)90099-6.
- Lall, U., T. Sangoyomi, and H. D. I. Abarbanel (1996), Nonlinear dynamics of the Great Salt Lake: Nonparametric short-term forecasting, *Water Resour. Res.*, *32*(4), 975–985, doi:10.1029/95WR03402.

- Lee Drbohlav, H.-K., and V. Krishnamurthy (2010), Spatial structure, forecast errors, and predictability of the South Asian monsoon in CFS monthly retrospective forecasts, *J. Clim.*, 23(18), 4750–4769, doi:10.1175/2010JCLI2356.1.
- Liu, Q., S. Islam, I. Rodriguez-Iturbe, and Y. Le (1998), Phase-space analysis of daily streamflow: Characterization and prediction, *Adv. Water Resour.*, 21, 463–475, doi:10.1016/S0309-1708(97)00013-4.
- Lorenz, E. N. (1963), Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.
- Lorenz, E. N. (1972), Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?, paper presented at 139th Meeting, Am. Assoc. for the Adv. of Sci., Washington, D. C.
- Maity, R., and D. Nagesh Kumar (2008), Basin-scale streamflow forecasting using the information of large-scale atmospheric circulation phenomena, *Hydrol. Processes*, 22, 643–650, doi:10.1002/hyp.6630.
- Maity, R., P. P. Bhagwat, and A. Bhatnagar (2010), Potential of support vector regression for prediction of monthly streamflow using endogenous property, *Hydrol. Processes*, 24, 917–923, doi:10.1002/hyp.7535.
- Mallat, S. G. (1989), A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(7), 674–693, doi:10.1109/34.192463.
- Mujumdar, P. P., and S. Ghosh (2008), Modeling GCM and scenario uncertainty using a possibilistic approach: Application to the Mahanadi River, India, *Water Resour. Res.*, 44, W06407, doi:10.1029/2007WR006137.
- Nerenberg, M. A. H., and C. Essex (1990), Correlation dimension and systematic geometric effects, *Phys. Rev. A*, 42, 7065–7074, doi:10.1103/PhysRevA.42.7065.
- Osborne, A. R., and A. Provenzale (1989), Finite correlation dimension for stochastic systems with power law spectra, *Physica D*, 35, 357–381, doi:10.1016/0167-2789(89)90075-4.
- Palmer, T. N. (1993), Extended-range atmospheric prediction and the Lorenz model, *Bull. Am. Meteorol. Soc.*, 74(1), 49–65, doi:10.1175/1520-0477(1993)074<0049:ERAPAT>2.0.CO;2.
- Porporato, A., and L. Ridolfi (1997), Nonlinear analysis of river flow time sequences, *Water Resour. Res.*, 33(6), 1353–1367, doi:10.1029/96WR03535.
- Porporato, A., and L. Ridolfi (2001), Multivariate nonlinear prediction of river flows, *J. Hydrol.*, 248, 109–122, doi:10.1016/S0022-1694(01)00395-X.
- Procaccia, I. (1988), Complex or just complicated?, *Nature*, 333, 498–499, doi:10.1038/333498a0.
- Provenzale, A., L. A. Smith, R. Vio, and G. Murante (1992), Distinguishing between low-dimensional dynamics and randomness in measured time series, *Physica D*, 58, 31–49, doi:10.1016/0167-2789(92)90100-2.
- Puente, C. E., and N. Obregón (1996), A deterministic geometric representation of temporal rainfall: Results for a storm in Boston, *Water Resour. Res.*, 32(9), 2825–2839, doi:10.1029/96WR01466.
- Raje, D., and P. P. Mujumdar (2009), A conditional random field-based downscaling method for assessment of climate change impact on multi-site daily precipitation in the Mahanadi basin, *Water Resour. Res.*, 45, W10404, doi:10.1029/2008WR007487.
- Raje, D., and P. P. Mujumdar (2010), Reservoir performance under uncertainty in hydrologic impacts of climate change, *Adv. Water Resour.*, 33, 312–326, doi:10.1016/j.advwatres.2009.12.008.
- Regonda, S., B. Rajagopalan, U. Lall, M. Clark, and Y.-I. Moon (2005), Local polynomial method for ensemble forecast of time series, *Nonlinear Processes Geophys.*, 12(3), 397–406, doi:10.5194/npg-12-397-2005.
- Rodriguez-Iturbe, I., B. Febres De Power, M. B. Sharifi, and K. P. Georgakakos (1989), Chaos in rainfall, *Water Resour. Res.*, 25(7), 1667–1675, doi:10.1029/WR025i007p01667.
- Rosenstein, M. T., J. J. Collins, and C. J. De Luca (1993), A practical method for calculating largest Lyapunov exponents from small data sets, *Physica D*, 65, 117–134, doi:10.1016/0167-2789(93)90009-P.
- Rying, E. A., G. L. Bilbro, and J.-C. Lu (2002), Focused local learning with wavelet neural networks, *IEEE Trans. Neural Networks*, 13(2), 304–319, doi:10.1109/72.991417.
- Saha, S., et al. (2006), The NCEP climate forecast system, *J. Clim.*, 19(15), 3483–3517, doi:10.1175/JCLI3812.1.
- Sangoyomi, T. B., U. Lall, and H. D. I. Abarbanel (1996), Nonlinear dynamics of the Great Salt Lake: Dimension estimation, *Water Resour. Res.*, 32(1), 149–159, doi:10.1029/95WR02872.
- Schertzer, D., I. Tchiguirinskaia, S. Lovejoy, P. Hubert, H. Bendjoudi, and M. Larchevêque (2002), Discussion of “Evidence of chaos in the rainfall-runoff process” Which chaos in the rainfall-runoff process?, *Hydrol. Sci. J.*, 47(1), 139–148, doi:10.1080/02626660209492913.
- Schreiber, T. (1993), Extremely simple nonlinear noise-reduction method, *Phys. Rev. E*, 47(4), 2401–2404, doi:10.1103/PhysRevE.47.2401.
- Schreiber, T., and H. Kantz (1996), Observing and predicting chaotic signals: Is 2% noise too much?, in *Predictability of Complex Dynamical Systems*, Springer Ser. Synergetics, vol. 69, edited by Y. A. Kravtsov and J. B. Kadtko, pp. 43–65, Springer, Berlin, Germany.
- Shang, P., X. Na, and S. Kamae (2009), Chaotic analysis of time series in the sediment transport phenomenon, *Chaos Solitons Fractals*, 41(1), 368–379, doi:10.1016/j.chaos.2008.01.014.
- Shukla, J. (1998), Predictability in the midst of chaos: A scientific basis for climate forecasting, *Science*, 282(5389), 728–731, doi:10.1126/science.282.5389.728.
- Sivakumar, B., S.-Y. Liang, C.-Y. Liaw, and K.-K. Phoon (1999a), Singapore rainfall behavior: Chaotic? *J. Hydrol. Eng.*, 4(1), 38–48, doi:10.1061/(ASCE)1084-0699(1999)4:1(38).
- Sivakumar, B., K.-K. Phoon, S.-Y. Liang, and C.-Y. Liaw (1999b), A systematic approach to noise reduction in chaotic hydrological time series, *J. Hydrol.*, 219, 103–135, doi:10.1016/S0022-1694(99)00051-7.
- Sivakumar, B., S. Sorooshian, H. V. Gupta, and X. Gao (2001a), A chaotic approach to rainfall disaggregation, *Water Resour. Res.*, 37(1), 61–72, doi:10.1029/2000WR900196.
- Sivakumar, B., R. Berndtsson, J. Olsson, and K. Jinno (2001b), Evidence of chaos in the rainfall-runoff process, *Hydrol. Sci. J.*, 46(1), 131–145, doi:10.1080/02626660109492805.
- Sivakumar, B., A. W. Jayawardena, and T. M. K. G. Fernando (2002), River flow forecasting: Use of phase-space reconstruction and artificial neural networks approaches, *J. Hydrol.*, 265, 225–245, doi:10.1016/S0022-1694(02)00112-9.
- Smith, L. A. (1988), Intrinsic limits on dimension calculations, *Phys. Lett. A*, 133, 283–288, doi:10.1016/0375-9601(88)90445-8.
- Smith, L. A. (2000), Disentangling uncertainty and error: On the predictability of nonlinear systems, in *Nonlinear Dynamics and Statistics*, edited by A. Mees, pp. 31–64, Birkhäuser, Boston, Mass.
- Smith, L. A., C. Ziehmann, and K. Fraedrich (1998), Uncertainty dynamics and predictability in chaotic systems, *Q. J. R. Meteorol. Soc.*, 125, 2855–2886.
- Takens, F. (1981), Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Lect. Notes Math.*, vol. 898, edited by D. A. Rand and L. S. Young, pp. 366–381, Springer, Berlin.
- Wang, Q., and T. Y. Gan (1998), Biases of correlation dimension estimates of streamflow data in the Canadian prairies, *Water Resour. Res.*, 34(9), 2329–2339, doi:10.1029/98WR01379.
- Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano (1985), Determining Lyapunov exponents from a time series, *Physica D*, 16, 285–317, doi:10.1016/0167-2789(85)90011-9.
- Zhang, Q. (1997), Using wavelet network in nonparametric estimation, *IEEE Trans. Neural Networks*, 8(2), 227–236, doi:10.1109/72.557660.
- Zhang, Q., and A. Benveniste (1992), Wavelet networks, *IEEE Trans. Neural Networks*, 3(6), 889–898, doi:10.1109/72.165591.

C. T. Dhanya and D. Nagesh Kumar, Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, India. (dhanya@civil.iisc.ernet.in; nagesh@civil.iisc.ernet.in)