Journal of Hydrology 374 (2009) 338–350

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Nonlinear multivariate rainfall–stage model for large wetland systems

Alaa Ali *

Hydrologic and Environmental Systems Modeling Department, South Florida Water Management District, 3301 Gun Club Rd., Stop Code 7520, West Palm Beach, FL 33406, USA

ARTICLE INFO

Article history:
Received 30 December 2008
Received in revised form 4 June 2009
Accepted 22 June 2009

This manuscript was handled by P. Baveye, Editor-in-Chief, with the assistance of C. Corradini, Associate Editor

Keywords:
Everglades
Wetland restoration
Multivariate time series analysis
Nonlinear model
Rainfall water-level model
Rainfall Driven Operations

SUMMARY

Wetland restoration is often measured by how close the spatial and temporal water level (stage) patterns are to the pre-drainage conditions. Driven by rainfall, such multivariate conditions are governed by nonstationary, nonlinear, and nonGaussian processes and are often simulated by physically based distributed models which are difficult to run in real time due to extensive data requirements. The objective of this study is to provide the wetland restorationists with a real time rainfall–stage modeling tool of simpler input structure and capability to recognize the wetland system complexity. A dynamic multivariate Nonlinear AutoRegressive network with eXogenous inputs (NARX) combined with Principal Component Analysis (PCA) was developed. An implementation procedure was proposed and an application to Florida Everglade's wetland systems was presented. Inputs to the model are time lagged rainfall, evapotranspiration and previously simulated stages. Data locations, preliminary time lag selection, spatial and temporal nonstationarity are identified through exploratory data analysis. PCA was used to eliminate input variable interdependence and to reduce the problem dimensions by more than 90% while retaining more than 80% of the process variance. A structured approach to select optimal time lags and network parameters was provided. NARX model results were compared to those of the linear Multivariate AutoRegressive model with eXogenous inputs. While one step ahead prediction shows comparable results, recursive prediction by NARX is far more superior to that of the linear model. Also, NARX testing under drastically different climatic conditions from those used in the development demonstrates a very good and robust performance. Driven by net rainfall, NARX exhibited robust stage prediction with an overall Efficiency Coefficient of 88%, Mean Square Error less than 0.004 m^2, a standard error less than 0.06 m, a bias close to zero and normal probability plots show that the errors are close to normal distributions.

© 2009 Elsevier B.V. All rights reserved.

Introduction

Wetland hydrology is perhaps the most important key element in wetland restoration projects. Wetlands have been drained and destroyed by human activities such as drainage, filling, dam construction, water diversions, groundwater pumping, canal dredging, and levee delineation. Such activities altered the spatial and temporal hydropattern characteristics including timing, amplitude, frequency, and duration of high and low waters. For example, the implementation of water management measures in Florida's Everglades in the late 1940s has destroyed many tree islands by either inundation due to prolonged high-water levels or by peat fires due to prolonged low-water levels. Improvement of such management practices requires proper prediction, as a first step, of restoration targets based on pre-drainage wetland response to current weather conditions.

The pre-drainage wetland hydrology is dominated by surface water processes and is driven by rainfall and evapotranspiration where surface water generally moves slowly downstream in response to the low land surface water gradient. Storm events within this system have an immediate local impact and a fading but persistently prolonged effect on areas downstream. Such a highly nonlinear stage is typically simulated by two dimensional physically based models that are typically used in planning studies and are difficult to run in real time applications due to extensive input data preparation and processing requirements. A real time modeling tool of such a complex environment with a simpler input structure to predict pre-drainage stage target time series is the focus of this study. Real time optimization of the managed system inflows and outflows, as a second step, to achieve the predicted stage targets is deferred to another manuscript.

System theoretic approach often serves as a viable alternative to physical modeling in real time applications. In this approach, difference or differential equations are used to characterize mapping of the input to the output directly with less emphasis on the internal structure driving the physical process. An example of this approach is the linear time series models (Salas et al., 1980; Bras and Rodriguez-Iturbe, 1985) where the emphasis has been the rainfall–runoff (R–R) modeling for flooding prediction. Although

* Tel.: +1 561 682 6513; fax: +1 561 697 7219.
E-mail addresses: aali@sfwmd.gov, aali@members.asce.org

0022-1694/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2009.06.033

these models produced reasonable predictions in many applications, they did not recognize the inherent nonlinearity of the R–R relationship. Those studies found R–R mapping to be complex, nonlinear, and nonstationary both spatially and temporally.

While a considerable research has been conducted on R–R studies, none or little was done to address the rainfall–stage (R–S) process. Van Lent (1995) provided the first attempt to model the R–S process in the Florida Everglades wetlands. He used weekly data at three rainfall stations and one Potential Evapo-Transpiration, PET, station to predict weekly stage target at three locations. In the absence of historical pre-drainage observations, the data used for the stage targets are taken from the 1965–1995 output of the South Florida Water Management District (SFWMD) physically based distributed Natural System Model (NSM), V4.6, which is based on extensive ecological conceptualization (SFWMD, 2000). Van Lent focused on modeling rainfall, PET, and stage residuals' covariance structure using a linear Autoregressive Moving Average model (ARMA). Based on his analysis, rainfall residuals were uncorrelated and nonGaussian while both PET and stage were correlated but Gaussian. He concluded that any stochastic model for this relationship must take nonGaussian white noise and Gaussian correlated signals to produce a Gaussian correlated signal. While the model results for one step ahead prediction were very reasonable, the results for the recursive prediction were unsatisfactory. Van Lent (1995) concluded that a linear model is not adequate for recursive prediction of the R–S relationship.

Artificial neural network and nonlinear modeling

As a system theoretic model, ANNs are a mathematical scheme with interconnected nodes and layers that is capable of mapping complex nonlinear processes from the input side to the output side. They are typically composed of three parts: inputs, one or many hidden layers and an output layer. Hidden and output neuron layers include the combination of weights, biases, and transfer functions. A neuron on a given layer is a hub that receives weighted contributions from the preceding layer's neurons and it sends weighted contributions to the succeeding layer's neurons. The weights are connections between neurons on one layer and another while the transfer functions are linear or nonlinear algebraic functions. When a pattern is presented to the network, weights and biases are adjusted so that a particular output is obtained. Such a learning is often achieved by means of backpropagation where such weights and biases are updated in the direction in which the performance function (e.g., Mean Square Error, MSE) decreases most rapidly (steepest descent). Because this does not necessarily lead to the fastest convergence, numerical nonlinear optimization techniques such as Conjugate Gradient and Levenberg–Marquardt Algorithms are often employed (Masters, 1995).

A satisfactory level of ANN training is the one that results in a good network generalization (i.e., satisfactory network performance on input data that was not part of the training). To help the network generalize, two data sets during training (modeling) are utilized; one to develop (train) the network and one to validate (verify) the performance. The training is stopped early (despite a continuous result improvement based on the training data set) if the network performance on such verification data failed to improve, remained the same, or deteriorated for a number of consecutive iterations.

ANNs can be classified into static feedforward networks and dynamic feedback networks. The feedforward networks are the most common form of ANNs. The architecture of this network consists of neurons connected by links across the input, hidden, and output layers. In this network, learning is based on a purely feedforward input to output mapping. The resulting weights are fixed where the state of neurons at a given time is determined by the input–output pattern for that time only without any consideration of the previous inputs, outputs, and states of network and hence called static network. Such networks are easily constructed with simple optimization algorithm and are of wide use in hydrology. These networks, however, are not efficient for highly dimensional; time dependent problems because of the slow convergence with a likely freeze in local minima, system memory modeling deficiency, and the need for a large training data set. The dynamic network has feedback connection(s) from the output layer (time delayed outputs); and/or hidden layers (previous states). The feedback feature provides a powerful learning capability for the network when memory is important in the system being modeled. The resulting network weights are adjustable to account for the previous neuron states, input and output. The dynamic network requires less training data sets for the same problem size (compared to static network). However, there are always network stability issues due to the dynamic nature of weights. A comparison between static and dynamic networks in an application to rainfall–runoff (R–R) modeling is provided by Chiang et al. (2004).

R–R modeling receives a major, if not the most major, share among ANN applications in hydrology (Hsu et al., 1995; Sajikumar and Thandaveswara, 1999; Govindaraju and Ramachandra Rao, 2000; Chang and Chen, 2001; Chang et al., 2002, 2004; Rajurkar et al., 2004; Ali et al., 2006, and Lin and Wang, 2007). In these studies, ANN architecture, learning algorithms and parameters in addition to R–R memory structure and memory length were the main issues of interest to address the nonlinear dynamics and complexity of the R–R process. ANN studies to address the R–S relationship are limited in the literature. One of these models reported by Chang and Chen, 2003 used a "novel" Radial Basis Function ANN to develop a one time step forecasting model for water stage at one station in an estuary subject to riverine and marine processes as a function of lunar calendar and measured stage at six locations. They used a two step process: (1) unsupervised training using fuzzy min–max clustering and (2) supervised learning using multivariate linear regression. The one hour forecasting results demonstrate satisfactory performance. No recurrent and/or multi stage concurrent prediction was reported in their study.

In this study, we provide a framework for the practitioner to develop and apply an ANN based model to predict stage target time series in response to rainfall and PET (aka Rainfall Driven Formula (RDF)) that recognizes the nonlinear, multivariate and temporal dynamics of large wetland systems and to demonstrate its efficacy in addressing system spatial and temporal nonstationarities. An application of this framework to Florida's Everglades will be presented.

Development of Rainfall Driven Formula

The development of the nonlinear RDF is based on the original linear AutoRegressive model with eXogenous input (ARX). To understand the nonlinear model, we first present the linear case as developed by Van lent (1995) followed by a full presentation of the RDF based on a Nonlinear AutoRegressive Dynamic Network with eXogenous variable (NARX).

Multivariate AutoRegressive model with eXogenous variable (ARX)

Van Lent (1995) applied the moving average coefficient matrix to rainfall and PET residual vector in lieu of the error vector which is, in essence, equivalent to AutoRegressive model with eXogenous input (ARX) (Box and Jenkins, 1970). The construction of the ARX model for this study is important to: (1) understand the development of the nonlinear model and (2) perform comparison between the linear and nonlinear cases. For theoretical description and limitations of the linear ARX, the reader is referred to Shumway and Stoffer (2000). The multivariate R–S residual time series relationship can be expressed as follows:

$$\boldsymbol{s}_t = \sum_{i=0}^{q-1} \boldsymbol{\Gamma}_i \boldsymbol{u}_{t-i} + \sum_{j=1}^{p} \boldsymbol{\Phi}_j s_{t-j} + \boldsymbol{\nu} + \boldsymbol{\omega}_t \tag{1}$$

where q and p are the input and output maximum time delays (model orders), n and m are state and exogenous variable observations at any time t, $\boldsymbol{u}$ is the mxq residual exogenous input matrix, $\boldsymbol{s}$ is the $p \times n$ residual state variable matrix, $\boldsymbol{\Gamma}$ is the nxm exogenous coefficient matrix, $\boldsymbol{\Phi}$ is the $p \times p$ state transition matrix, and $\boldsymbol{\nu}$ is the bias vector, and $\boldsymbol{\omega}$ is a random vector with zero mean and covariance matrix. Input and output data are residuals of the original time series after subtracting the corresponding periodic mean based on the development data period of record. Parameter Matrices $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}$, ν, and $\boldsymbol{\omega}$ are estimated using stepwise regression according to Neumaier and Schneider (2001). The model optimal orders (p and q) is selected based on the minimum Akaike Information Criteria (AIC) score formulated by Bedrick and Tsai (1994). An application of this model is presented and compared to the nonlinear case in the application section. The ARX assumes linearity where the system can be adequately described by the principle of superposition where input individual contribution to the output is independent and additive. In a nonlinear system, coupling across all possible connections must be properly modeled. A nonlinear rainfall–stage model in wetland systems presented in a multivariate recursive framework as follows.

Nonlinear AutoRegressive Network with eXogenous variables (NARX or the RDF)

A nonlinear form of the ARX model is represented by a recurrent dynamic network with feedback connections enclosing several layers of the network. This two-layer network is used to approximate a nonlinear characteristics function $\wp$ in a nonlinear multivariate framework. Given Eq. (1) variables, the mathematical representation of a noise free nonlinear system can be expressed as:

$$s_t = \wp\left(s^1_{t-1} \quad \cdots \quad s^n_{t-p} \quad \cdots \quad u^1_t \quad \cdots \quad u^m_{t-1-q}\right) \tag{2}$$

Large values of n, m, q, and p can significantly increase the problem dimensionality resulting in insufficient data size for sound training. Also, input vector elements "s" and "u" may be interdependent violating the ARX model assumption. To address these two issues, we use Principal Component Analysis, PCA (Jolliff, 2002) to transform the input data to orthogonal components. In this analysis, we compute the Eigenvectors, V, of the input covariance matrix, C, such that, $\boldsymbol{V^{-1}CV = D}$, where D is the Eigen value diagonal matrix. We create the Principal Component (PC) coefficient matrix "$\boldsymbol{A}$" by: (1) rearranging columns of matrix V and Eigen value matrix D in order of descending Eigen value, and then (2) retaining the minimum number of Eigenvectors, r, corresponding to cumulative Eigen values that explain most of the data variance. The $1 \times (np + mq)$ input vector of Eq. (2) can be transformed to its first "r" PCs using the $(np + mq) \times r$ "$\boldsymbol{A}$" PC coefficient matrix. Eq. (2) can be re-written as follows:

$$s_t = \Re\left(\left[s^1_{t-1} \cdots s^n_{t-p} \cdots u^1_t \cdots u^m_{t-1-q}\right]^T_{1\times(np+mq)} \times [A]_{(np+mq)\times r}\right) \tag{3}$$

A diagram of the resulting network is depicted in Fig. 1. The ANN is composed of one input layer, a hidden layer, and an output layer. The input layer takes time delayed exogenous input signal, u, and time delayed output signal, s, processes the first "r" PCs and passes them on to each neuron of the hidden layer which approximates the nonlinear characteristic function $\Re$ using a hyperbolic tangent sigmoid transfer function. The output layer takes a linear combination of the hidden layer signals and connects them to each output neuron. The output signal is then fed-back into the input layer for the next time step simulation.

Bayesian regularization backpropagation as part of the Levenberg–Marquardt Optimization is used as the training function to speed the backpropagation convergence rate. Bayesian regularization minimizes a linear combination of squared errors and weights. It also modifies the linear combination so that at the end of training, the resulting network has good generalization qualities. For more detailed discussions of Bayesian regularization, the reader is referred to MacKay (1992).

Model implementation

The Comprehensive Everglades Restoration Plan (CERP) is one of the largest wetland restoration projects in the world's history (Fig. 2). The operation of CERP components aims at improving the amount and timing of water delivery to/from the Florida's 2.4 million acre Everglades to attain stage targets based on pre-drain-

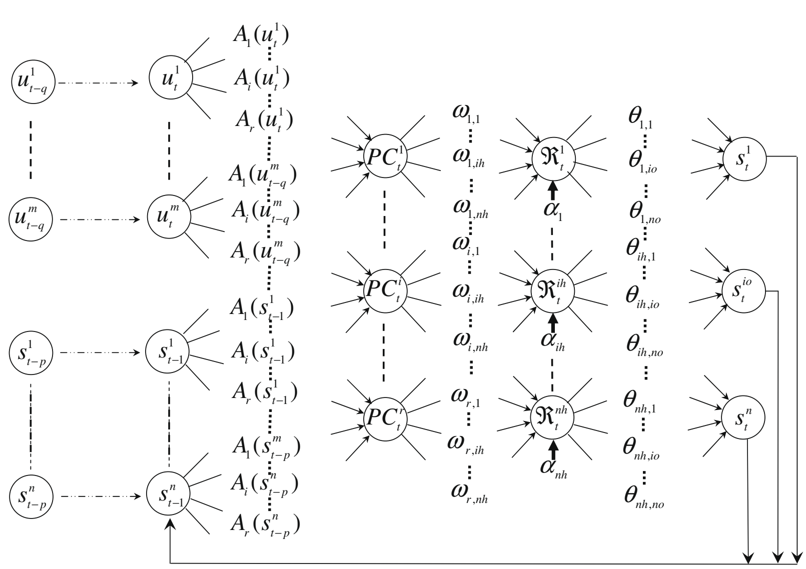

Fig. 1. Schematic diagram of for the Rainfall Driven Formula.

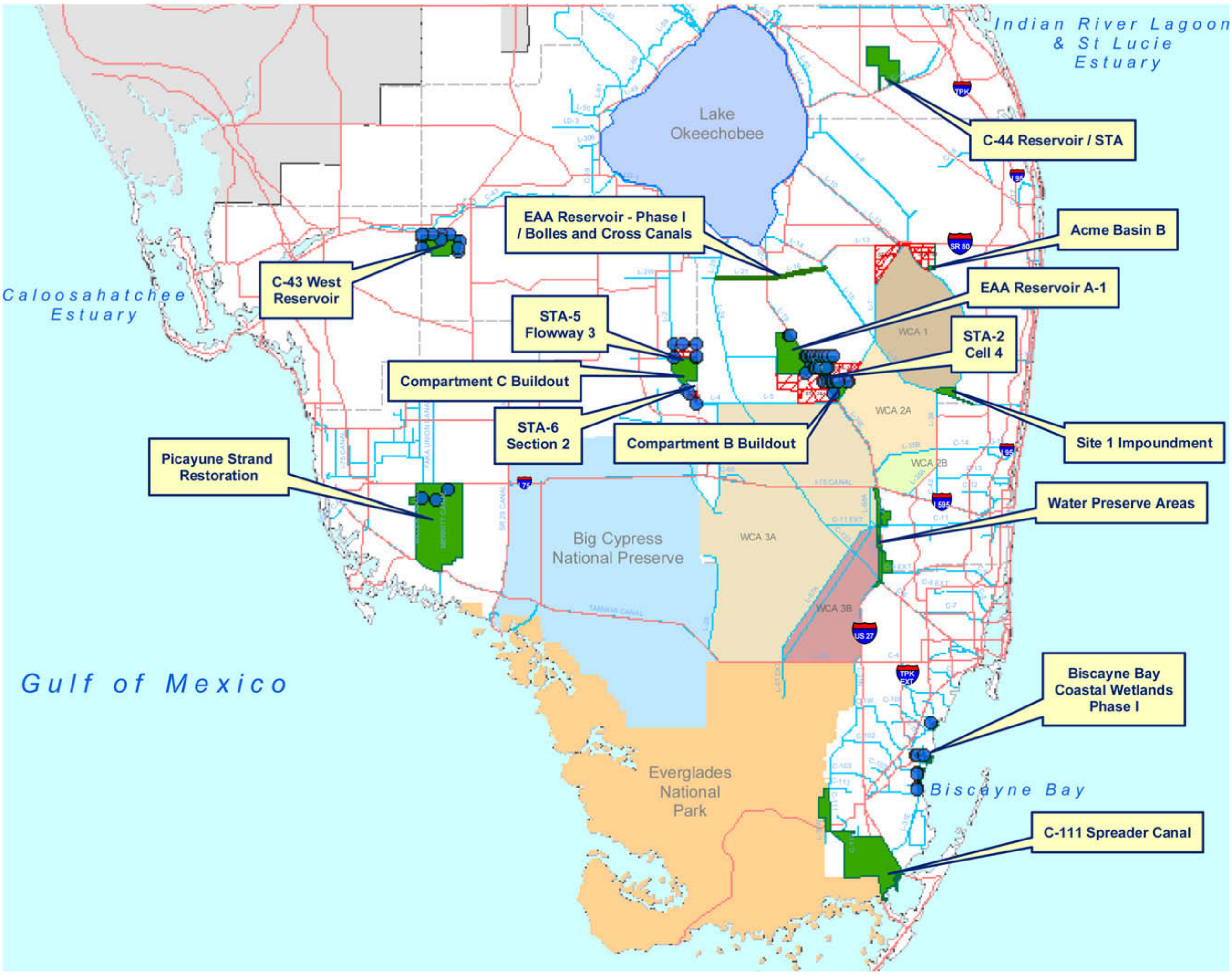

Fig. 2. Project area with overall Comprehensive Everglades Restoration Plan (CERP) components.

age conditions and as modeled by SFWMD Natural System Model (NSM), V4.6. Real time observations from such a natural system are not possible and hence recursive prediction of stage target time series must rely on previously predicted values representing a challenge to model robustness against error propagation. In this section we provide procedures to implement ARX and NARX models presented in the preceding section to predict stage targets in the half million acre Water Conservation Areas 3A (WCA-3A), and the one million acre Everglades National Park (ENP) using a limited data of rainfall and PET. Data is analyzed first to understand the system dynamics, select rainfall data locations, and determine preliminary time lags and to explore climatic nonstationarity. Model data and parameterization is then presented for ARX and NARX models.

Exploratory analysis

Historical data used to develop and test the ARX and NARX models are rainfall and PET as input and pre-drainage stage at five locations as output (solid triangles in Fig. 3). In the absence of historical pre-drainage observations, simulated stage target data based on Natural System Model (NSM), V4.6, is used. Two PET stations at "Belle Glades" and "TAMIAMI Trail" locations are selected for their proximity to the study area (light squares in Fig. 3). To select a limited but effective rainfall data set out of the SFWMD rainfall network, we first identified the spatial extent of all rainfall stations whose cross correlation with at least one stage trigger is 0.66 (the light circles in Fig. 3). A subset of 24 rainfall satiations was then selected (solid circles in Fig. 3) based on current operation, data quality, value of correlation, and many preliminary trial and error model runs.

Weekly time step for long term restoration has been an acceptable time scale for CERP planning, modeling, and operations of South Florida hydrologic systems. System dynamics of smaller than weekly scale is local in nature and is not in the interest of wetland long term restoration while larger than weekly scale smears pattern periodic variability that is of such a restoration interest. Weekly stage Auto Correlation and Partial Auto Correlation Functions (ACF and PACF) describe the system linear temporal behavior, provide insight about the system's memory and help identify preliminary time lags for both models. Fig. 4 shows stage ACF and PACF at the five stage trigger locations. The ACF shows a long slowly fading memory in the system that converges at lag 26 after which the correlation spuriously increases due to the seasonality. The PACF shows that most of the system's memory is stored in the most recent 2–4 weeks. Stage below ground surface behaves differently from that above the ground surface at all locations (Fig. 5 shows representative behavior at WCA3A central location "WC3A4"). The analysis in this section suggests preliminary model orders in the range of 2–4 weeks and also indicates spatial nonstationarity of stage responses if stage across the sites has mixed positions with respect to ground surface.

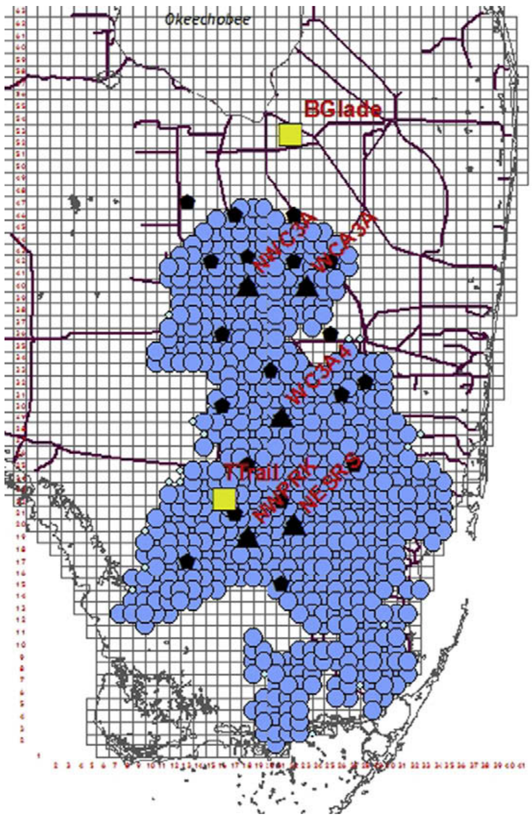

Fig. 3. Overall high rainfall–stage correlation zone (>0.66) (denoted by light blue circles) and data locations. Rainfall stations are denoted by black circles, PET stations are denoted by yellow boxes and stage trigger locations are black triangles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Climatic global trend

The Central and South Florida hydrologic system was significantly wetter during the 1991–2000 period, as compared to the 1965–1990 period. A spatially averaged stage time series across the five stage target sites shows this trend clearly (Fig. 6). It is effortless to notice the distinct pattern differences between the post-1991 and pre-1991 stage target variation. The pre-1991 pattern exhibits lower stage levels reflecting dryer conditions with larger amplitude reflecting more variable weather with severe dry weather in the dry season. The post-1990 pattern exhibits higher stage levels reflecting wetter conditions with smaller amplitude reflecting persistent wet conditions that sustained the troughs above certain levels. A traditional filtering of periodic means (e.g., spectral analysis) of these data is not sufficient to produce stationary residuals.

Normal probability plots for the two periods' stage (Fig. 7) depict distribution differences. The post-1991 stage is nearer to Gaussian distribution compared to the pre-1991 period. Historical weekly averages for the post-1991 and pre-1991 periods exhibit different patterns (Fig. 13, upper portion, solid black and white traces). Not only did the climatic trend result in wetter conditions, but it also caused a 4 week seasonality shift creating a longer, wetter than average dry season. Model application in such a temporally nonstationary environment is a challenge.

ARX model data and parameterization

In this study, 24 rainfall stations and two PET stations as exogenous input data and NSM simulated stage target at five stations in WCA-3A and ENP as output data were selected (Fig. 3). Daily data were converted into a weekly time step and were partitioned into two sets: modeling set (1965–1990) and testing set (1991–2000). Rainfall, PET, and stage residuals are obtained by subtracting the weekly data from its respective weekly periodic mean using Fast Fourier Transform (Duhamel and Vetterli, 1990). Global mean and variance of the stage residuals for the development data set is significantly different from the corresponding measures of the testing data set with the mean being higher and the variance being lower (Table 1). Such an observed nonstationarity is a challenge for the linear ARX application. AIC score exhibited a flat trend where the model order could not be uniquely defined. According to the aforementioned ACF and PACF analyses and according to subsequent lag selection for the nonlinear model, rain, PET, and stage maximum lags were selected to be 4, 4, and 2 weeks, respectively. For space limitation, we refer the reader to Van Lent (1995) for full details of the model parameter selection of the ARX model (Eq. (1)).

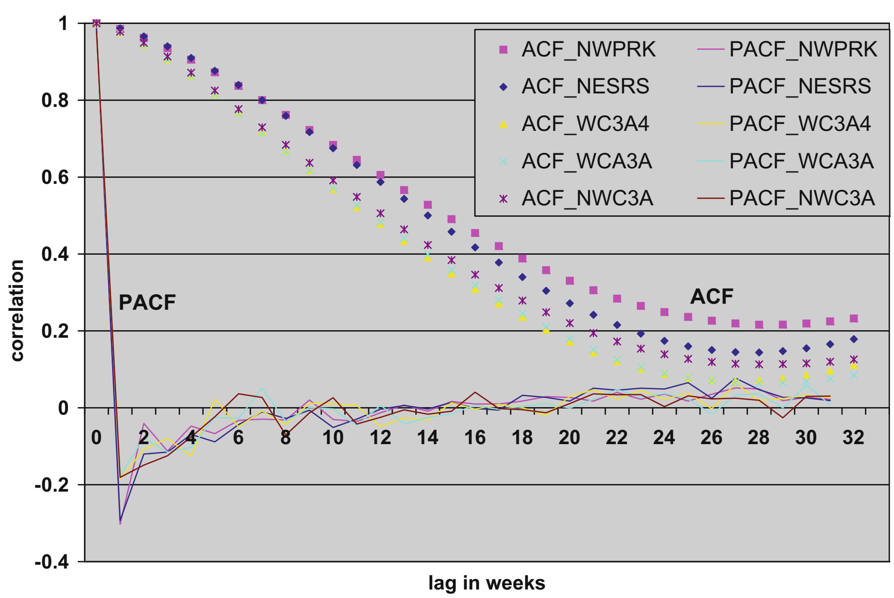

Fig. 4. Stage target auto and partial auto correlation functions at the five trigger locations.

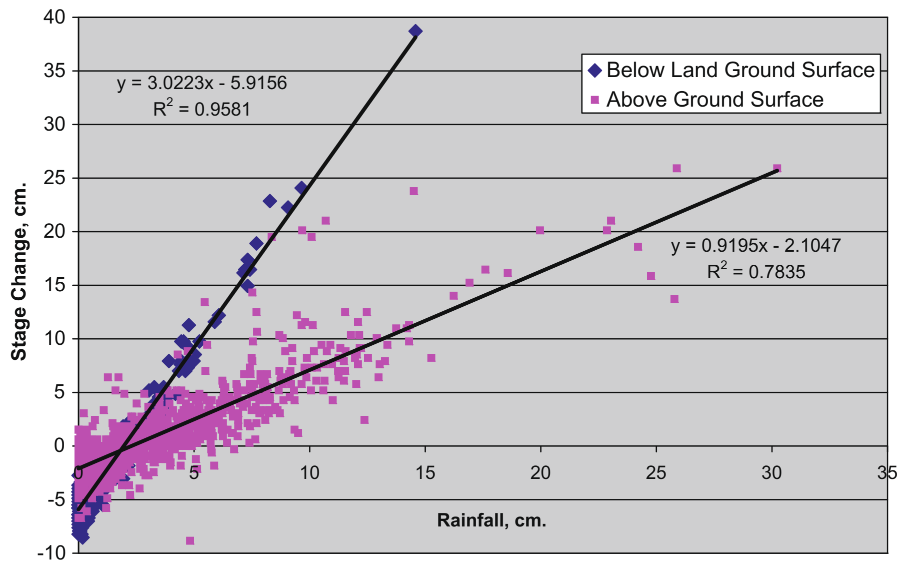

Fig. 5. Effect of land surface elevation on stage response to rainfall at gage location "WC3A4" with land surface elevation = 2.54 m.

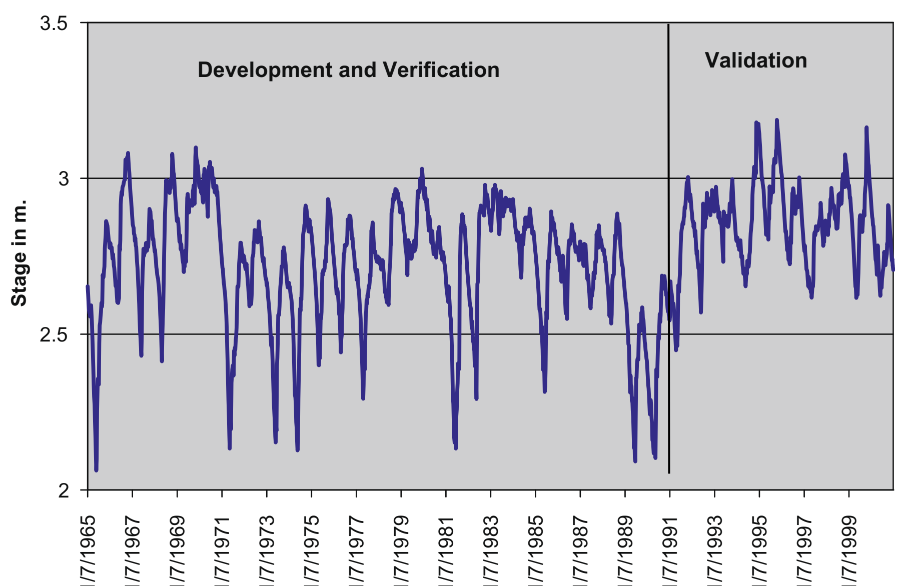

Fig. 6. Stage time series for 1965–2000 averaged across the five trigger locations.

NARX model data and parameterization (the RDF)

The 26 year weekly modeling data set used to develop the ARX model was randomly sub-sampled without replacement to produce development (18 years) and validation (8 years) subsets adopting a technique that preserves the associated serial sequence for any data entry. Two hundred partition scenarios were obtained to allow for the exploration of different combinations of development/validation subsets that result in a good selection of optimal parameters and a good NARX network generalization independently from the testing data set. To eliminate disparity across rainfall, PET, and stage data, each time series in each of the three data subsets is standardized by subtracting the respective mean and dividing by the respective standard deviation of the development data subset.

Given 26 rain and PET stations and five stage output variables with time lag sensitivity range of 4 weeks, the model input can be as large as $26 * 4 + 5 * 4 = 124$ variables. PCAs on such variables for different lag combinations show that the majority of data variance is retained by the first 10 components. A pareto plot of a typical PCA shown in Fig. 8 displays bar chart of % variance explained by each component in a descending order along with cumulative curve of variance explained. The first 10 components explain about 82% of the variance with diminishing contribution beyond the 10th PC.

Given the first 10 PCs, and Bayesian regularization backpropagation training function, the stage and rain/PET time lags (model orders), the number of hidden layers and hidden nodes are selected based on numerous NARX sensitivity runs looping through the parameter ranges presented in Table 2 for each of the 200 development/validation partition scenarios. For each partition scenario, the least MSE score and the associated four parameter values are recorded. The MSE scores across the 200 partitions were more sensitive to time lags more than NARX hidden layers or hidden nodes.

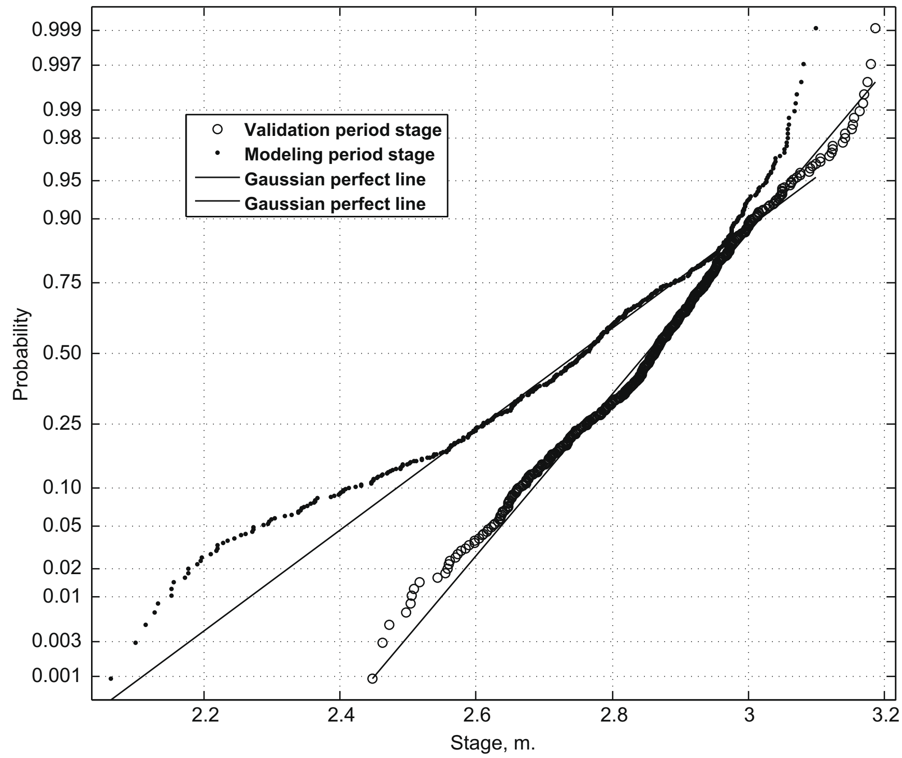

Fig. 7. Normal probability plot for the NSM (historical) stage for the modeling and testing periods.

Table 1
Global mean and variance of stage residuals for both modeling and testing periods.

	NWC3A	WCA3A	WC3A4	NESRS	NWPRK
Modeling					
Mean, m	0.000	0.000	0.000	0.000	0.000
Variance, m^2	0.228	0.356	0.266	0.340	0.665
Testing					
Mean	0.313	0.412	0.447	0.494	0.581
Variance, m^2	0.179	0.190	0.119	0.182	0.211

Table 3 shows the relative frequency of each lag combination that achieved the least MSE across the 200 partition runs. Two week stage and 4 week rain/PET S2R4 lag combination has a fairly higher relative frequency than the rest of the other lag combinations. In this exercise we record for each partition scenario, the minimum MSE, the associated time lag combinations, standard error (SE), and bias (AvgE). We also record the maximum "worst" MSE and the associated time lag combinations (Fig. 9). Lag combinations left to S2R4 on the *x*-axis exhibited spurious jumps in the maximum MSE and bias measures. The poor maximum MSE performance may be attributed to insufficient stage/rain predictors to model the process based on drastically nonstationary development/validation data random partition scenario. The error performance stabilizes for all stage–rain combinations right to S2R3 on the *x*-axis (for all partition scenarios regardless of nonstationarity). Among such combinations, S2R4 recorded the lowest bias and the lowest "minimum" MSE scores. It ranked the 4th lowest standard error, SE, and the 5th lowest "maximum" MSE. The other lag combinations with reasonable error scores are S3R3 and S3R4 with low SE but high "minimum" MSE and high error bias. Given the order of importance of these statistics, the stage–rain S2R4 lag combination appears to be the best selection for the RDF model associated with one hidden layer, eight hidden nodes, and five output nodes.

Results

The overall results for ARX and NARX (RDF) models for one step ahead and recursive predictions, respectively are presented in Section "Overall one step ahead prediction", and part of Section "Overall recursive prediction performance". Detailed results of NARX model is presented in the remainder of Sections "Overall recursive prediction performance" and "NARX (RDF) performance at central WCA3A (WC3A4 site)". The reader is referred to the original report of this study (Ali, 2007) for full results presentation.

Overall one step ahead prediction

The ARX and NARX models are applied using rain, PET, and the previous time step true stage to predict stage at the current time step. The use of the previous time step true data makes the prediction much more accurate than the recursive case where the previous time step simulated stage is used. While such a prediction is not important for long term restoration, it reflects the model capability in modeling the change in the system state within one time step. Table 4 presents lumped statistical measures such as Efficiency Coefficient (EC) (Nash and Sutcliffe, 1970), MSE, SE and bias between the observed and predicted stage for ARX and NARX, respectively. The overall results for both models are very good with ARX being slightly better. The EC of ARX is almost 1 during the modeling period and it deteriorated by an average of 6% during the testing period with EC score below 0.9 for WCA3A4 location. The other statistics exhibit similar performance. The EC of NARX on the other hand is in the range of 0.91–0.94 for the modeling period and has improved in the testing period by 2% on the average. Both MSE and SE exhibit similar slight improvement during the testing period while the bias exhibits slight deterioration. Such counterintuitive results are attributed to the wetter conditions with lower variability during the testing period. The statistics show better performance for ARX in both modeling and testing periods.

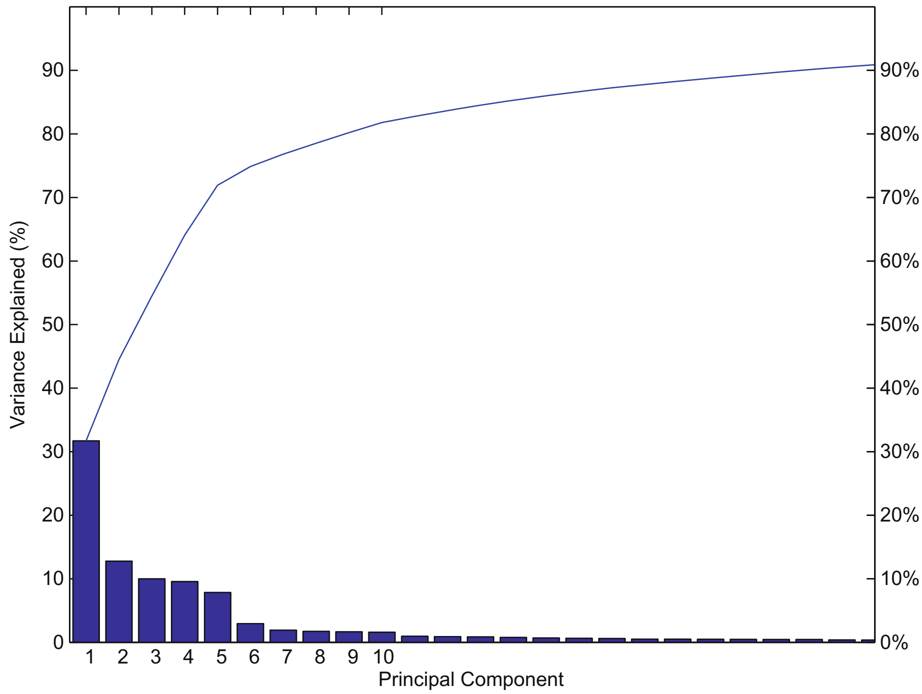

Fig. 8. Pareto plot for percentage of variance explained by individual; and cumulative principal components, PCs.

Table 2
Parameter ranges for the Rainfall Driven Formula (RDF) model.

Parameter	Range	Value selected
Stage lag	1–4 weeks	2
Rain/pet lag	1–4 weeks	4
# Hidden layers	1–3 layers	1[a]
# Hidden nodes	4–12	8[a]

[a] Solution was not sensitive to this parameter and the value selected was incidental to a low MSE score.

However, the relative performance in both periods suggests that NARX generalizes better than ARX.

Overall recursive prediction performance

Table 5 presents the same measures as in Table 4 for recursive prediction where the previous output is fed-back into the input for the current time step prediction. ARX statistics show fair performance during the modeling period and unsatisfactory performance during the testing period with EC, MSE, SE, and bias being multiple folds worse indicating ARX failure under recursive environment given its good performance in one step prediction. This deterioration is partially attributed to the significantly different statistics exhibited in the testing period (Table 1) where ARX is strictly stationary model. The NARX, on the other hand, shows good performance in the testing period that shows EC above 0.8, MSE ranging between 0.002 and 0.004 m^2, the SE ranging between 0.04 and 0.06 m, and the bias is in the .005–.02 m range. Given the results presented thus far, it is preferred to reserve the remaining space of this section for NARX results only.

The above lumped statistic tables summarize the global model performance during the modeling and testing periods. The results were presented for temporally averaged statistics that do not provide any information about the RDF temporal performance. Fig. 10 shows the RDF and NSM stage time series for three locations (Fig. 3) NWCA3A, WC3A4, and NESRS representing the Northern, Central, and Southern Everglades, respectively. Despite the global similarity among the three graphs, local differences are observed and are captured individually. The peaks and troughs are reasonably captured except during the 1997–1998 water year especially in the North. The dry season for that year received significant rainfall events in the North compared to the Central area and South and compared to its typical average. The NSM stage spatial/temporal variation for this dry season of the year was exceptionally unusual. The stage amplitude within that dry season was 0.35 m in the North while it was less than 0.12 m in the Central area and in the South. Fig. 11 shows the RDF performance for each week of the year where spatially averaged stage data are averaged for that week over the years of for the modeling and testing periods. The figure shows very good matching between the RDF and the NSM data for the two periods with about .012 average bias in the modeling period and it shows also a 4 week seasonality shift in the post-1991 period which is very well captured with as a good or better performance as that of the modeling (pre-1991) period. Figs. 12a and 12b show other error graphical representations. Fig. 12a is a scatter plot exhibiting a tight dispersion around the 45° line indicating a strong correlation of 0.89 between predicted and NSM averaged stage restoration target. Fig. 12b shows a close to a straight line error distribution on a normal probability paper indicating that the global errors probability distribution is close to normal.

Table 3
Relative frequency of lag combination occurrence corresponding to the least mean square error (MSE) scores across the 200 data partitions.

Stage lag	2	2	4	4	1	3	3	1	2	4	1	3	2	3	4
Rain/PET lag	4	3	2	4	3	3	4	4	2	3	2	2	1	1	1
Percentage of the least MSE occurrences (%)	18	11	10	9	9	7	6	6	6	5	4	4	<1	<1	<1

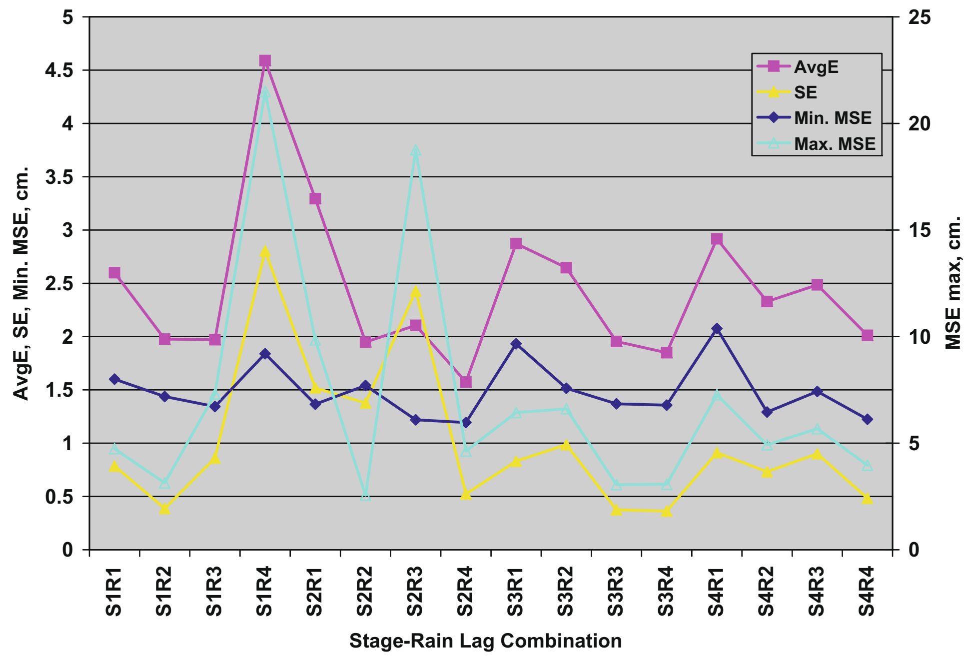

Fig. 9. Performance measure statistics as function of stage–rain lag combination (AvgE is the error average or bias, SE is the standard error, Min. MSE is the least mean square error recorded and the Max. MSE is the highest mean square error recorded).

Table 4
One step ahead prediction performance for ARX and RDF models. Performance measures used are model efficiency, EC, mean square error, MSE, standard error, SE, and bias.

	NWC3A		WCA3A		WC3A4		NESRS		NWPRK	
	ARX	RDF	ARX	RDF	ARX	RDF	ARX	RDF	ARX	RDF
Modeling										
Efficiency	0.997	0.909	0.995	0.928	0.995	0.939	0.998	0.944	0.997	0.931
MSE, m^2	0.000	0.003	0.000	0.003	0.000	0.002	0.000	0.002	0.000	0.005
SE, m	0.010	0.051	0.015	0.057	0.014	0.047	0.010	0.049	0.014	0.074
Bias, m	0.000	−0.002	0.000	0.001	0.000	−0.001	0.000	0.001	0.000	0.000
Testing										
Efficiency	0.964	0.921	0.949	0.921	0.887	0.962	0.985	0.970	0.915	0.963
MSE, m^2	0.001	0.002	0.001	0.002	0.002	0.001	0.000	0.001	0.002	0.002
SE, m	0.025	0.042	0.028	0.048	0.038	0.032	0.020	0.036	0.050	0.046
Bias, m	−0.002	0.018	−0.004	0.013	−0.007	−0.015	−0.002	−0.013	−0.003	−0.011

Table 5
Recursive prediction performance for ARX and RDF models. Same measures as those used in Table 4.

	NWC3A		WCA3A		WC3A4		NESRS		NWPRK	
	ARX	RDF	ARX	RDF	ARX	RDF	ARX	RDF	ARX	RDF
Modeling										
Efficiency	0.483	0.748	0.48	0.779	0.505	0.788	0.49	0.828	0.357	0.742
MSE, m^2	0.015	0.007	0.023	0.01	0.018	0.008	0.022	0.007	0.051	0.02
SE, m	0.121	0.083	0.153	0.098	0.133	0.085	0.147	0.084	0.225	0.141
Bias, m	−0.003	−0.016	−0.003	−0.018	−0.003	−0.017	−0.004	−0.017	−0.005	−0.019
Testing										
Efficiency	−2.056	0.833	−2.24	0.857	−5.332	0.899	−2.694	0.916	−11.832	0.909
MSE, m^2	0.054	0.004	0.05	0.004	0.085	0.002	0.096	0.003	0.374	0.005
SE, m	0.23	0.059	0.218	0.062	0.277	0.044	0.295	0.053	0.599	0.065
Bias, m	−0.027	0.011	−0.049	0.005	−0.09	−0.02	−0.095	−0.023	−0.127	−0.019

NARX (RDF) performance at central WCA3A (WC3A4 site)

Of interest in the remainder of the results section is to depict the RDF's capability in capturing certain features of the stage target signals at WC3A4 site by comparing the simulated versus the NSM restoration targets for the modeling and testing periods. The WC3A4 site is located in the center of the Water Conservation Area 3A. As seen in Fig. 10 middle plots, the RDF captured the variation trend very well and reproduced most of the peaks and troughs with the exception of early dry season peak late 1994 and early 1995. The RDF performance at this location was better than that in the Northern NWCA3A. Fig. 13 presents the predicted and

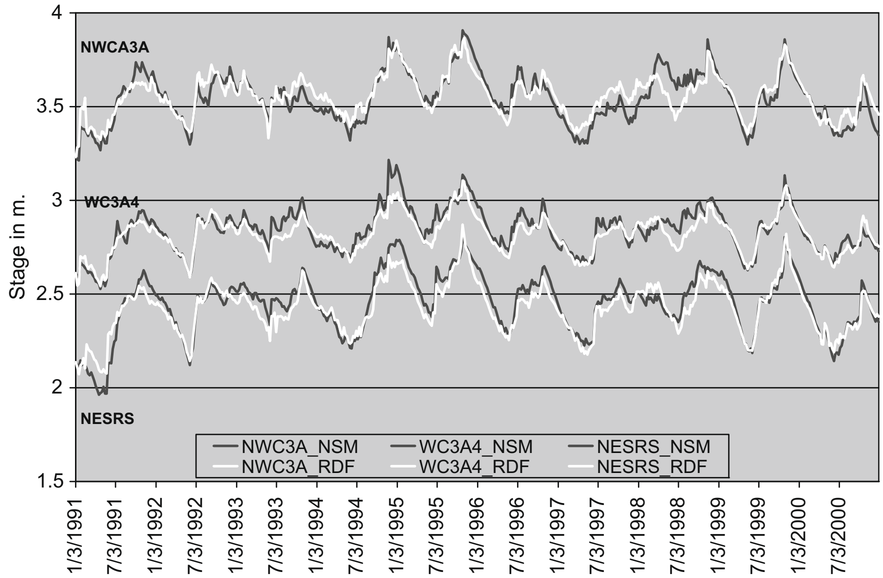

Fig. 10. NSM and RDF stage target hydrograph for the testing period for three sites.

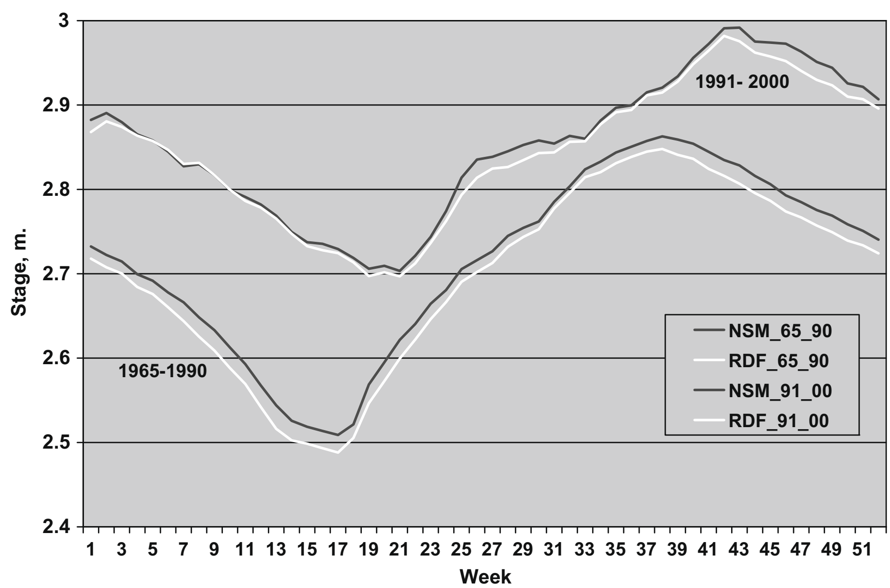

Fig. 11. NSM and RDF overall stage weekly average for both the modeling and testing periods.

NSM weekly average and standard deviation of stage target for the modeling and testing periods. The figure shows very good matching between the RDF and the NSM weekly average during the modeling period and the testing period with less bias than that observed in the overall plots in Fig. 11. Fig. 13 shows also a very good matching between the RDF and NSM weekly standard deviation for the modeling period and a fair performance during the testing period where the standard deviation trend is well captured. For full presentation of the RDF performance at all stage target sites, the reader is referred to Ali, 2007.

RDF sensitivity to the number of rainfall stations

In the application of this study, we used twenty four rainfall stations. This number may not be practically feasible for real time data acquisition by many agencies. To illustrate the efficacy of the RDF model with smaller number of rainfall stations, we performed two sets of RDF model runs using: (1) 50% of the original rainfall stations and (2) rainfall stations at the five stage gage locations. For space limitation, we only present lumped results in terms of Efficiency Coefficient only (Table 6). From Table 6, the performance of RDF for five stations seems reasonable and it improves monotonically as the number of stations increases. During the testing period, RDF generalizes reasonably well with monotonic improvement as the number of stations increases. RDF performance for five and 12 stations at the downstream gages is very similar due to the information conveyed from the upstream gages.

Summary

Pre-drainage wetland stage is essential measure for any wetland restoration project. In this study, a nonlinear multivariate

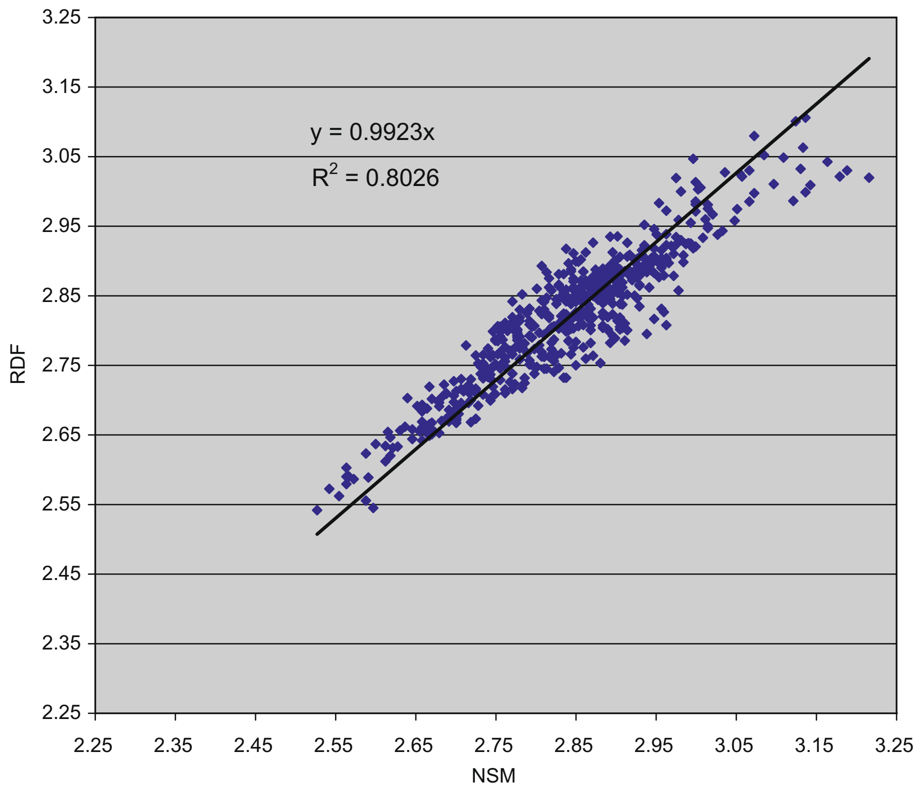

Fig. 12a. Scatter plot for overall RDF versus NSM stage targets.

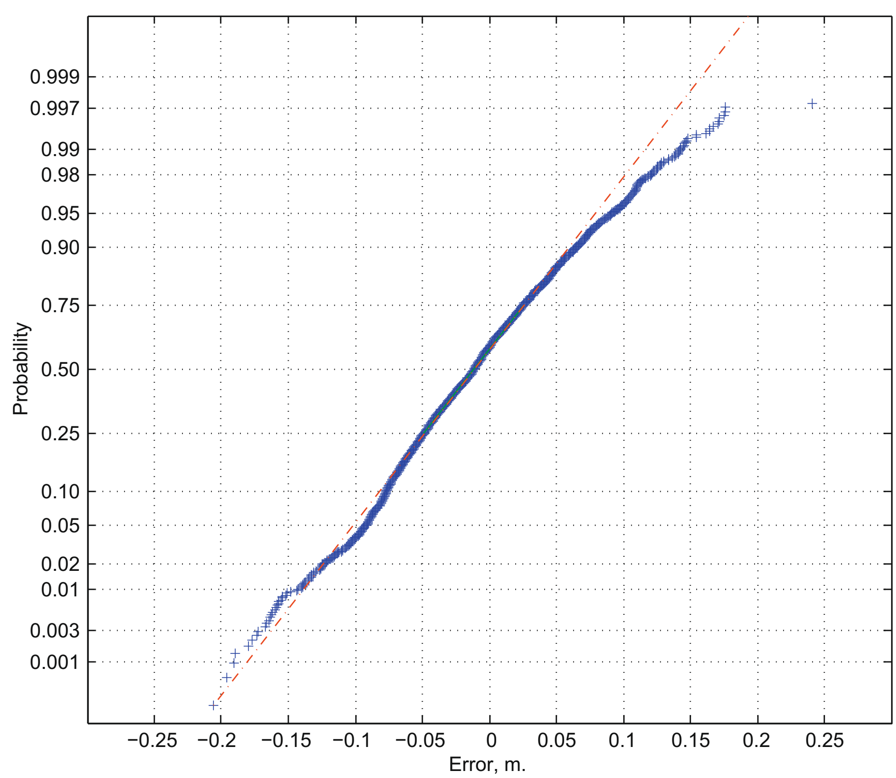

Fig. 12b. Normal probability plot for the overall residuals.

Rainfall Stage model (RDF) was developed to provide a real time recursive prediction for stage restoration targets in a complex wetland system with an application to Florida's Everglades. Although the techniques adopted are not new, the algorithm to utilize such techniques provided the first R–S model in complex wetland systems. The RDF is based on a dynamic Nonlinear AutoRegressive

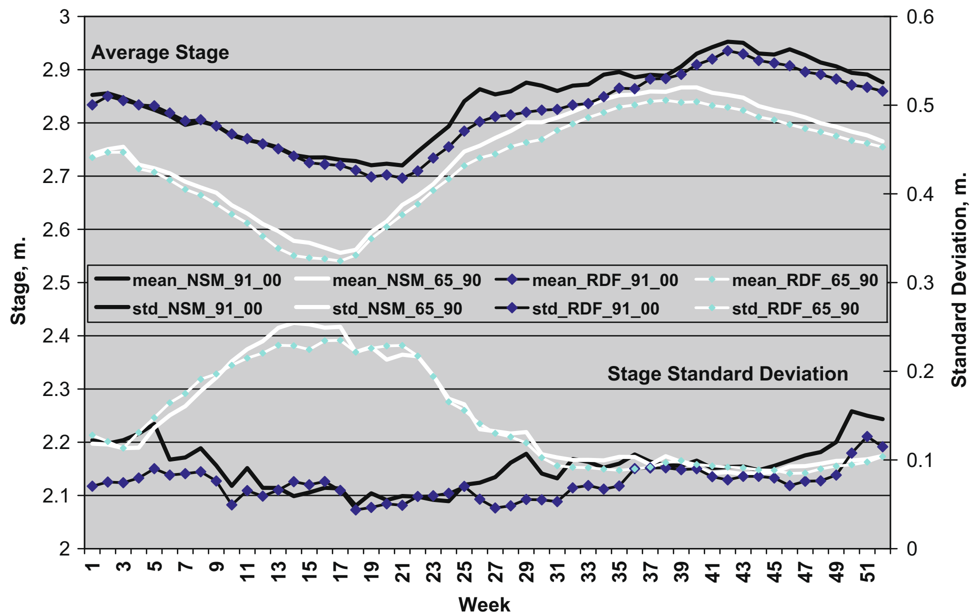

Fig. 13. NSM and RDF stage weekly average and standard deviation for the modeling and testing periods at gage location "WC3A4".

Table 6
Effect of number of input rainfall stations on the model efficiency during the modeling and testing periods.

# Stations	NWC3A	WCA3A	WC3A4	NESRS	NWPRK
Modeling					
5 Stations	0.623	0.695	0.887	0.869	0.881
12 Stations	0.708	0.779	0.861	0.864	0.878
24 Stations	0.833	0.857	0.899	0.916	0.909
Testing					
5 Stations	0.556	0.534	0.626	0.687	0.568
12 Stations	0.621	0.651	0.679	0.659	0.595
24 Stations	0.748	0.779	0.788	0.828	0.742

ANN with eXogenous inputs, NARX, with feedback connection and with certain architecture. Spatial correlation analysis was first performed to select effective number of rainfall and PET stations, and temporal correlation analysis was performed to select preliminary time delays for rainfall/PET and stage. Further analysis is performed to explore system spatial and temporal nonstationarity. PCA was performed to eliminate interdependence and to reduce the problem dimensions into manageable few representative PCs. A structured approach was designed to combine PCA with NARX training to identify optimal time delays, significant PCs, and number of ANN hidden layers and hidden nodes.

The above procedure has been applied to the 1.5 million acre Everglades wetland WCA-3A and ENP. The implementation takes the raw data consisting of rain/PET at the 26 designated locations over the most recent 4 weeks and previously simulated stage at the five stage trigger locations over the preceding 2 weeks; transforms the data set into its first ten PCs, train, validate and test NARX. The training proceeds forward with the stage output of the last two time steps are back-fed into the input of the current time step. The training stops when the ANN performance on the validation data set stops improving or deteriorating for five consecutive iterations. For comparison purposes, a classic ARX model was constructed and applied using the same data set.

Driven by only rainfall and PET, the ARX and RDF models provide very good performance for one step ahead prediction. For recursive prediction, however, the RDF performance was significantly better especially during the testing period. The RDF efficiency coefficient is 0.75 or higher during the modeling period and 0.83 or higher during the testing period. The RDF other statistics exhibit the same high performance with more improvement in the testing period. The improvement during the testing period is attributed to lower data variance.

The stage quantiles, weekly average and weekly standard deviation are all well captured at each trigger location during the modeling and testing periods. Individual error plots indicate Gaussian distribution and tight scatter plots indicating a sound unbiased prediction. The ability of the RDF to capture the hydrology wetter conditions in the 1990s and adapt to the smaller amplitude (characteristics that were never seen during training) is a powerful indicator that the RDF development has come very close to the unique solution of this problem.

The RDF parameters were selected based on a nested loop system through several parameters where thousands of ANN runs are generated and parameter sets corresponding to improved results are recorded. Identification of unnecessary variables/parameters/weights to achieve parsimony was not an easy process. The PCA capability in exploiting the system spatial and temporal dependence allowed for filtering out the components with insignificant representation of the data variability. In the Everglades, the PCA has reduced the input vector from 114 entries into 10 PCs who represented 82% of the input data variance. There has been noticeable improvement and generalization to the model performance with the increase of the number of PCs being added beyond the 10 components. Such an improvement is on the expense of model parsimony and simplicity of subsequent model application. Such a trade off between model parsimony and model adequacy in the context of ANNs structure deserves more research effort which is outside the scope of this paper.

The RDF model exhibited modest sensitivity to the number of rainfall stations being used with improved results as more rainfall stations being added. Rainfall stations limited to the stage gage locations provided reasonable Efficiency Coefficient results indicating model efficacy using limited number of data.

There are numerous sources of uncertainty associated with the data selection, period of record partition, process and ANN parameter selection leading to uncertainty in the model prediction. An improvement to the real time prediction is often achieved by an adaptive component (e.g., Kalman filter) where subsequent prediction is improved based on the relative performance of previous predictions against observations that are made available. Due to the lack of current observations of pre-drainage stage and the difficulty to characterize the non Gaussian structure of the signal noise in such a nonlinear system, adaptive prediction was not feasible in this study. In the subsequent optimization of system operation where real time stage is adjusted to achieve predicted stage targets, an extended Kalman filter (Julier and Uhlmann, 2004) for nonlinear non Gaussian system may be utilized. Without adaptive prediction component, a quantification of uncertainty becomes particularly important to guide subsequent system operation and to guide future selection of additional data locations and improvement of model parameters and structure.

Acknowledgements

The author wishes to thank Akin Owosina for useful feedback and wishes to thank the anonymous reviewers whose comments have significantly improved the presentation of this work.

References

Ali, A., 2007. Everglades Stage Based Rainfall Driven Formula, Acceler8 Rainfall Driven Operation Project, Deliverable 1, Internal Publication, IMC-HESM, South Florida Water Management District.
Ali, A., Neidrauer, C., Cadavid, L., Linton, p., Tarboton, K., Obeysekera, J., 2006. New rainfall-based management plan for more natural water deliveries to the everglades shark river slough, operating reservoirs in changing conditions. In: Proceedings of the Operations Management 2006 Conference, August 14–16, 2006, Sacramento, CA.
Bedrick, E.J., Tsai, C.-L., 1994. Model selection for multivariate regression in small samples. Biometrics 50, 226–231.
Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis, Forecasting, and Control. Holden-Day, Oakland, CA.
Bras, R.L., Rodriguez-Iturbe, I., 1985. Random Functions and Hydrology. Addison-Wesley, Reading, Mass.
Chang, F.J., Chang, L.C., Huang, H.-L., 2002. Real time recurrent learning neural network for stream-flow forecasting. Hydrological Processes 16, 2577–2588.
Chang, F.J., Chen, Y.C., 2001. A counter propagation fuzzy-neural network modeling approach to real time stream flow prediction. Journal of Hydrology 245 (1–4), 153–164.
Chang, F.J., Chen, Y.C., 2003. Estuary water-stage forecasting by using radial basis function neural network. Journal of Hydrology 270, 158–166.
Chang, L.C., Chang, F.J., Chiang, Y.M., 2004. A two-step ahead recurrent neural network for streamflow forecasting. Hydrological Processes 18, 81–92.
Chiang, Y.M., Chang, L.C., Chang, F.J., 2004. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. Journal of Hydrology 2004 (290), 297–311.
Duhamel, P., Vetterli, M., 1990. Fast Fourier transforms: a tutorial review and a state of the art. Signal Processing, vol. 19, April 1990, pp. 259–299.
Govindaraju, R.S., Ramachandra Rao, A., 2000. Artificial Neural Networks in Hydrology. Kluwer Academic Publisher.
Hsu, K., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of the rainfall–runoff process. Water Resources Research 31 (10), 2517–2530.
Jolliff, I.T., 2002. Principal Component Analysis, second ed. Springer Series in Statistics. Springer-Verlag, New York, Berlin, Heidelberg.
Julier, S., Uhlmann, J., 2004. Unscented filtering and nonlinear estimation. Proceedings of the IEEE 92 (3), 401–422. 2004.
Lin, G.F., Wang, C.M.A., 2007. Nonlinear rainfall–runoff model embedded with an automated calibration method – Part 1: the model. Journal of Hydrology 341 (3-4), 186–195.
MacKay, D.J.C., 1992. Bayesian interpolation. Neural Computation 4 (3), 415–447.
Masters, T., 1995. Advanced Algorithms for Neural Networks: A C^{++} Sourcebook. John Wiley and Sons, NY, ISBN 0-471-10588-0.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. Journal of Hydrology 10 (3), 282–290.
Neumaier, A., Schneider, T., 2001. Estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Transactions on Mathematical Software 27, 27–57.
Rajurkar, M.P., Kothyari, U.C., Chaube, U.C., 2004. Modeling of the daily rainfall–runoff relationship with artificial neural network. Journal of Hydrology 285 (1-4), 96–113.
Sajikumar, N., Thandaveswara, B.S., 1999. A nonlinear rainfall runoff model using an artificial neural network. Journal of Hydrology 216 (1-2), 32–55.
Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L., 1980. Applied Modeling of Hydrologic Time Series. Water Resour. Publ., Littleton, Colo.
Shumway, R.H., Stoffer, D.S., 2000. Time Series Analysis and its Applications. Springer, New York.
Staff at South Florida Water Management District, SFWMD, 2000. Natural System Model Release 4.6, South Florida Water Management District, West Palm Beach, 72 p.
Van Lent, T., 1995. Using the Natural System Model as A Basis for Water Management in Shark Slough, Everglades National Park, CA-5280-49014, Department of Civil Engineering, South Dakota State University.