



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

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

Download English Version:

<https://daneshyari.com/en/article/4578767>

Download Persian Version:

<https://daneshyari.com/article/4578767>

[Daneshyari.com](https://daneshyari.com)