Journal of Hydrology 548 (2017) 305–321

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Multi-scale quantitative precipitation forecasting using nonlinear and nonstationary teleconnection signals and artificial neural network models

Ni-Bin Chang [a,*], Y. Jeffrey Yang [b], Sanaz Imen [a], Lee Mullon [a]

[a] Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL, USA
[b] U.S. EPA, Office of Research and Development, National Risk Management Research Laboratory, Cincinnati, OH, USA

ARTICLE INFO

Article history:
Received 1 August 2016
Received in revised form 28 November 2016
Accepted 2 March 2017
Available online 9 March 2017
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Purna Chandra Nayak, Associate Editor

Keywords:
Climate change
Precipitation
Sea surface temperature
Wavelet
Teleconnection patterns
Artificial neural networks
Forecasting

ABSTRACT

Global sea surface temperature (SST) anomalies are observed to have a significant effect on terrestrial precipitation patterns throughout the United States. SST variations have been correlated with terrestrial precipitation via ocean–atmospheric interactions known as climate teleconnections. This study demonstrates how the scale effect could affect the forecasting accuracy with or without the inclusion of those newly discovered unknown teleconnection signals between Adirondack precipitation and SST anomaly in the Atlantic and Pacific oceans. Unique SST regions of both known and unknown telecommunication signals were extracted from the wavelet analysis and used as input variables in an artificial neural network (ANN) forecasting model. Monthly and seasonal scales were considered with respect to a host of long-term (30-year) nonlinear and nonstationary teleconnection signals detected locally at the study site of Adirondack. Similar intra-annual time-lag effects of SST on precipitation variability are salient at both time scales. Sensitivity analysis of four scenarios reveals that more improvements of the forecasting accuracy of the ANN model can be observed by including both known and unknown teleconnection patterns at both time scales, although such improvements are not salient. Research findings also highlight the importance of choosing the forecasting model at the seasonal scale to predict more accurate peak values and global trends of terrestrial precipitation in response to teleconnection signals. The scale shift from monthly to seasonal may improve results by 17% and 17 mm/day in terms of *R* squared and root of mean square error values, respectively, if both known and unknown SST regions are considered for forecasting.

1. Introduction

Hydrologic time series prediction has been a successful modeling tool to address the underlying physical mechanisms by a holistic approach (Coulibaly and Baldwin, 2005). A problem arises in this the type of physical system modeling due to its highly dynamic and even chaotic nature, exhibiting nonlinearity and non-stationarity. This variability means that the characteristics of the system change over time, possibly due to external forcing and internal feedback mechanisms simultaneously, which are difficult to represent by traditional numerical models with physics-based governing equations. Inputs and outputs of this type of physical system are not proportional and are therefore subject to frequent abrupt change, even in a chaotic mode as opposed to slow and gradual, during which multiple equilibrium states are the norm (Rial et al., 2004). This tendency contrasts with traditional assumptions of stationarity and linearity in many environmental systems analyses.

Climate is generally deemed the best example of nonlinear dynamics driven by interactions of its subsystems with multiple scales (e.g., the ocean, the terrestrial systems, and the atmosphere) (Gonzalez et al., 2013). Hence, the main challenge in studying the climate system is to detect the key linkages and their interactions/feedbacks that modulate the dynamic behaviour of the system (Gonzalez et al., 2013). As an example of the climate systems' functional response, we focused on the precipitation in the northeastern United States, a region with well documented, strong correlations with known climate indices or teleconnection patterns regarded as a set of index regions in the oceans, such as the Pacific North American (PNA), El-Nino Southern Oscillation (ENSO), and the North Atlantic Oscillation (NAO) Patterns. The warm El Nino phase of the ENSO has been linked with cold and stormy winter weather in the northeastern United States, whereas the cold La

* Corresponding author.
E-mail address: nchang@ucf.edu (N.-B. Chang).

http://dx.doi.org/10.1016/j.jhydrol.2017.03.003
0022-1694/

Nino cycle brings warmer and milder winter weather into the same region (Greene, 2012). The NAO, traditionally measured by sea level pressure differences between the Azores and Iceland, is the dominant pattern of Atlantic climate variability and has been shown to have broad influence over the Atlantic storm track (Marshall et al., 2001). During positive NAO phases, Iceland's low pressure systems and Azores' high pressures lock the jet stream in the northeastern zonal path across North America and cause greater maximum temperatures and fewer extreme cold days in the winter in the northeastern United States. (Wettstein and Mearns, 2002). In addition, during negative NAO phases, above-average snowfall is produced in New England as a result of eastward displacement of the typical East Coast storm track (Bradbury et al., 2003). PNA is also a dominant extratropical teleconnection pattern in the Northern Hemisphere, characterized by above average sea level pressure (SLP) or geospatial heights over the Hawaiian region and mountainous regions of western North America, as well as below average heights south of Alaska and over the southeastern United States. The positive cycle of PNA has been linked to increased precipitation over the northeastern United States by up to 15% over the last century (Beckage et al., 2008; Tang and Beckage, 2010).

Focusing on teleconnection signals generated by these known patterns in climate science, however, limits the predictive potential for hydrometeorological forecasting, mainly depending on those unstable relationships between precipitation variability and some of these known patterns (Gan et al., 2007; Jiang, 2013). The strength of these climate or teleconnection patterns, which are normally composed of some key index regions in oceans, could change in time and space, making their applications unreliable for terrestrial precipitation forecasting. Hence, one of the main efforts of this study was to explore the potential or unknown teleconnection signals affecting terrestrial precipitation over a specific geographic location, signals mainly selected as sectors of sea surface temperature (SST) not previously reported. In this context, these selected sectors of SST are termed herein as "unknown SST regions" or "unknown index regions", relative to the teleconnection power tied to these known patterns (such as NAO, ENSO, PNA, and others). These same areas are also termed "known SST regions" in the Pacific and Atlantic oceans and have a significant effect on precipitation variability over the terrestrial study site. Regardless of whether or not they are known or unknown, they are herein called "SST index regions" or simply "SST regions", all of which collectively contribute to a unique "teleconnection pattern".

Recently, artificial neural network (ANN) models have been found to be useful predictor variables for hydrologic forecasting. The ENSO pattern, for example, is an important coupled oceanic–atmospheric phenomenon with global significance to the Earth's climate and strong manifestations in temperature and precipitation worldwide (Solomon et al., 2007). Abbot and Marohasy (2012) found a strong performance enhancement to Australian precipitation forecasting when combining multiple climate-related teleconnection patterns or index regions, including the Southern Oscillation Index (SOI), Pacific Decadal Oscillation (PDO), and the Nino 3.4 index, and this scientific discovery gave rise to evidence-based support for our study.

ANNs are inherently reliant on the basic time series predictor data of the model. Early ANN-based hydrologic studies reported limited forecasting success, which has been attributed to weaknesses in the time series inputs. To solve this issue, Wang and Ding (2003) proposed pre-treating the time series data by decomposing wavelets into their fundamental signals, creating a new hybrid ANN model called the Wavelet Neural Network (WNN). Wavelet decomposition provides a mathematical process for distilling a signal into multiple levels of details while also extracting local information of the time series (Satyaji Rao and Krishna, 2009), providing a more robust representation of the time series data. WNNs also solved a common problem in ANN forecasting. Multivariate hydrologic time series data are routinely combined from many sources of varying frequencies into a single model, such as evaportransporation and streamflow data (Anctil and Tape, 2004), a process that can have a conflicting effect on the ANN, causing the model to diverge from the solution. By extracting the underlying frequency behavior of these datasets, the WNN models can more successfully interpret the interrelationships.

Mwale et al. (2004) used a WNN model to forecast seasonal precipitation across a gridded array of rainfall data at a continental scale. Using this a novel approach, Mwale et al. (2004) applied wavelet principal component analysis (WPCA), a technique that analyzes individual scale-averaged wavelet power (SAWP) datasets, to identify homogeneous zones of rainfall variability and to improve precipitation predictability in East Africa. Interestingly, Mwale et al. (2004) did not rely on the known SST regions such as ENSO and NAO in their predictive modeling, and instead included independent teleconnection patterns or index regions in the relevant oceans to obtain predictor data.

Long-term teleconnection analysis is hampered by using linear correlation directly between SST and terrestrial response when evaluating some nonlinear systems (Franzke, 2009) and when depending on some teleconnection signals with low signal-to-noise ratios. This study extends an earlier short-term study (Mullon et al., 2013) by exploring the time scale effect on the precipitation within Adirondack State Park based on a host of long-term (30-year) teleconnection signals between SST in the North Atlantic and Pacific oceans. This study does not, however, aim to show how much improvement can be gained by including some unknown teleconnection signals not previously considered because most known teleconnection patterns have been identified and proved influential for terrestrial precipitation variability. Instead, the aim of this research was to demonstrate how the scale effect from monthly to seasonal could affect the forecasting accuracy with or without the inclusion of those newly discovered unknown teleconnection signals between Adirondack precipitation and SST anomaly in the Atlantic and Pacific oceans. Based on a wavelet-based ANN forecasting model, teleconnection signals associated with specific geographic locations (e.g., the Adirondack site herein) can be comprehensively compared without relying solely on the known SST regions. The unknown teleconnection patterns were therefore derived independently and based exclusively on the SST and terrestrial precipitation databases from the National Aeronautics and Space Administration (NASA) and National Oceanographic and Atmospheric Administration (NOAA), respectively. Nevertheless, the inputs of local meteorological information associated with mesoscale and synoptic-scale weather phenomena were not included in this forecasting approach because they are associated with a more deterministic weather prediction scheme for a short-term (i.e., 5–7 day) meteorological analysis.

The following science questions are explored herein: (1) Do unknown teleconnection patterns exist between the Adirondack region precipitation and SST in the North Atlantic and Pacific oceans? (2) What is the effect of considering the time lag in temporal aggregation on detecting SST index regions that significantly affect precipitation variability over the Adirondack region? (3) How is precipitation forecasting affected by shifting the time scale from seasonal to monthly? (4) How would the performance of the developed forecasting model be affected if the unknown teleconnection patterns were not included as predictors in the model? We hypothesized that both influential known and unknown teleconnection signals exist for the study region, and including the unknown teleconnection patterns could largely improve the per-