AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020482

Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information

Tiantian Yang[1,2], Ata Akbari Asanjan[1], Edwin Welles[2], Xiaogang Gao[1], Soroosh Sorooshian[1], and Xiaomang Liu[3]

[1]Department of Civil and Environmental Engineering, Center for Hydrometeorology and Remote Sensing [CHRS], University of California-Irvine, Irvine, California, USA, [2]Deltares USA Inc., Silver Spring, Maryland, USA, [3]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Key Points:
- Artificial intelligence and data mining (AI&DM) techniques are powerful regression tools in developing reservoir monthly inflow forecasts
- Climate phenomenon indices have a complex relationship with hydrological conditions, and provide useful information for reservoir operations
- Different AI & DM techniques have strengths and limitations and are suggested to use with proper parameterization and prior examination

Supporting Information:
- Supporting Information S1

Correspondence to:
T. Yang,
tiantiay@uci.edu

Citation:
Yang, T., A. A. Asanjan, E. Welles, X. Gao, S. Sorooshian, and X. Liu (2017), Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information, *Water Resour. Res.*, *53*, 2786–2812, doi:10.1002/2017WR020482.

Received 25 JAN 2017
Accepted 6 MAR 2017
Accepted article online 9 MAR 2017
Published online 7 APR 2017
Corrected 15 MAY 2017

Abstract Reservoirs are fundamental human-built infrastructures that collect, store, and deliver fresh surface water in a timely manner for many purposes. Efficient reservoir operation requires policy makers and operators to understand how reservoir inflows are changing under different hydrological and climatic conditions to enable forecast-informed operations. Over the last decade, the uses of Artificial Intelligence and Data Mining [AI & DM] techniques in assisting reservoir streamflow subseasonal to seasonal forecasts have been increasing. In this study, Random Forest [RF], Artificial Neural Network (ANN), and Support Vector Regression (SVR) are employed and compared with respect to their capabilities for predicting 1 month-ahead reservoir inflows for two headwater reservoirs in USA and China. Both current and lagged hydrological information and 17 known climate phenomenon indices, i.e., PDO and ENSO, etc., are selected as predictors for simulating reservoir inflows. Results show (1) three methods are capable of providing monthly reservoir inflows with satisfactory statistics; (2) the results obtained by Random Forest have the best statistical performances compared with the other two methods; (3) another advantage of Random Forest algorithm is its capability of interpreting raw model inputs; (4) climate phenomenon indices are useful in assisting monthly or seasonal forecasts of reservoir inflow; and (5) different climate conditions are autocorrelated with up to several months, and the climatic information and their lags are cross correlated with local hydrological conditions in our case studies.

1. Introduction

Reservoirs are the vital human-built infrastructures that collect, store, and deliver fresh surface water for many purposes in a timely manner. Efficient reservoir management is capable of providing society with resilience against hydrological extremes, water-supply sustainability, flood protection for urban areas, and clean, renewable power production. Over the past century, much effort has been made by dam operators, policy makers, and scientists to better understand reservoir operations, and develop optimal reservoir operation strategies. According to *CDWR* [2014a] and [2014b], the primary focus of reservoir operations in developed regions, such as California, is how to validate the operation strategies of existing facilities based on hydrological information, and improve subseasonal to seasonal forecasts in mitigating changing climatic conditions at different temporal scales, i.e., real-time, subseasonal to seasonal, and single to multiple years [*CDWR*, 2014a, 2014b]. For example, it is widely acknowledged that climate conditions can significantly impact water supply and many climate phenomenon indices can be used as predictors in supporting water resources management [*Pagano and Garen*, 2003; *Montoya et al.*, 2014; *Kalra et al.*, 2013; *Gutiérrez and Dracup*, 2001; *Garen*, 1993]. The research question for the reservoir systems in developed countries is how to utilize different types of auxiliary information to support reservoir management and to develop forecast-informed operations for existing facilities. However, in developing countries, such as China, besides the requirements stated above for existing reservoirs, many new reservoirs are under construction or being planned and a number of water diversion projects have recently begun, i.e., the center route of the China's South-to-North Water Diversion Project. Advanced modeling and decision support tools, such as the AI & DM techniques, are therefore needed in developing countries for efficient management and operation purposes.

In recent years, with the advances in computer sciences, the AI & DM techniques have become more and more popular in the field of streamflow forecasts, reservoir operation planning and scheduling [*Hejazi and*

Cai, 2009]. Among all kinds of Artificial Intelligence & Data Mining (AI & DM) techniques, the Artificial Neural Networks (ANN), the Decision Tree (DT), and Support Vector Machine or Regression (SVM or SVR) methods are three of the most popular techniques in developing streamflow forecasts at different temporal scales around the world [*Schnier and Cai*, 2014; *Zealand et al*., 1999; *Yaseen et al*., 2015; *Cheng et al*., 2015; *Kumar et al*., 2013; *Erdal and Karakurt*, 2013; *Maity et al*., 2010; *Guo et al*., 2011; *Asefa et al*., 2006].

The Artificial Neural Network is a robust, nonlinear machine learning approach, which has been extensively applied for many classification and regression problems in various fields. In the field of streamflow and reservoir inflow forecasts, *Cheng et al*. [2015] compared ANN and SVM in forecasting monthly inflow of the Xinfengjiang Reservoir in China and proved both methods have satisfactory performances. *Thirumalaiah and Deo* [1998] used ANN in real-time forecasting of water levels based on upstream gauging station information and historical records in a river system in Jagdalpur in India. *Lima et al*. [2016] used a simple ANN to incorporate newly arrived meteorological data and to produce daily forecasts for two small watersheds in British Columbia, Canada. *Wu et al*. [2009] used both SVR and ANN to predict the streamflow timeseries for two river outlets located in China, and they examined the predictive skills for different lead times, including 1, 3, 6, and 12 month-ahead. *Wang et al*. [2006] compared different hybrid ANNs with regard to their capability of streamflow prediction at the headwater region of the Yellow River, China, at a daily scale. *Linares-Rodriguez et al*. [2015] demonstrated the flexibility of ANNs on adding additional runoff indices to enhance 1 day-ahead streamflow forecast in the Northeast Guadalquivir basin in southern Spain. *Jain et al*. [1999] presented the usefulness of ANNs not only in reservoir inflow prediction, but also in the optimal reservoir scheduling and management of a diversion reservoir located in the Godavari basin and Mahanadi basin of India. *Ashaary et al*. [2015] also demonstrated the application of ANN in developing the short-term forecasting model for the change in the reservoir water level in the Timah Tasoh Reservoir located in Northern Peninsular Malaysia. A recent summary of using ANN and other AI methods in hydrological applications and forecasting can be found in *Yaseen et al*. [2015].

Similarly, there are also extensive studies and applications of using DT methods to assist streamflow forecasts. *Erdal and Karakurt* [2013] compared SVR with a DT algorithm in predicting monthly streamflow in the Çoruh River in the Eastern Black Sea Region in Turkey and concluded that DT methods were able to produce better results than SVR. *Kumar et al*. [2013] tested the performances of MLR, ANN, fuzzy logic, and DT algorithms in predicting streamflow at an upstream reservoir in the Sutlej Basin in northern India. They determined the DT methods performed well when compared to other methods. *Galelli and Castelletti* [2013] assessed the predictive performances of multiple DT methods and ANN in forecasting streamflow of the Marina catchment in Singapore and the Canning River in Western Australia. In addition, *Galelli and Castelletti* [2013] demonstrated the DT method is superior over ANN due to its nonparametric characteristics, which makes DT methods suitable for large computationally intensive problems. *Wei* [2012] compared two popular DT algorithms (C5.0 and CART) in predicting reservoir releases in northern Taiwan during typhoon events and concluded that the DT methods are skillful in discharge simulation. *Cheng et al*. [2008] used a DT approach as a predictive model to determine the optimal reservoir releases before the onset of typhoons at the Shihmen Reservoir System in Taiwan and justified its capability in assisting streamflow prediction during flood conditions. In a more recent study, *Yang et al*. [2016] compared a standard DT algorithm (CART) with a Random Forest algorithm in predicting the daily reservoir discharges for nine different river basins in California and tested the suitability of DT methods for a generalized discharge simulation problem.

Besides ANN and DT methods, the Support Vector Machine (SVM) is another popular method in streamflow forecasts. *Asefa et al*. [2006] demonstrated a case study in Sevier River Basin in Utah, USA, and produced a promising streamflow prediction results at both seasonal and hourly temporal scales. *Maity et al*. [2010] applied a Support Vector Regression (SVR) method to predict monthly streamflow in the Mahanadi River Basin in the State of Orissa, India, and showed the superior performance of SVR over autocorrelation regression. *Lin et al*. [2006] presented an enhanced SVR model to predict long-term reservoir discharges from the Manwan Hydropower Reservoir in China. *Guo et al*. [2011] developed an adaptive SVR model to conduct a monthly streamflow prediction on the Three Gorges Area in the Yangtze River basin in China. *Guo et al*. [2011] also concluded an SVR model is capable of producing accurate predictions of streamflow and the SVR model has good generalization characteristics in solving streamflow prediction problems.

As a summary of the uses of AI & DM techniques in developing streamflow forecasts as listed above, the ANN, DT, and SVR are all proven to be powerful tools in predicting streamflow and reservoir inflows. However, the investigation on different parameterization among those models and comprehensive comparison study are rarely reported, which impedes the practical uses of these methods in water resources management and planning. As communicated with the Snow Survey Office from the California Department of Water Resources, many operational streamflow regression models are not as complex as AI & DM methods. The decades-long experiences from many hydrologists and engineers have been instrumental in determining the regression coefficients of the operational models. The AI & DM tools provide the mechanism for enhancing the existing hydrometerological forecasts for reservoir management, given the fact that more and more types of data have become available (so called "Big Data Era") and the increasing possibilities of taking advantages of those auxiliary information to support reservoir operations. This comparison study aims to provide a baseline, and test the applicability and robustness of different AI & DM tools in support of reservoir operations and hydrological forecasts. As a part of our development process, we conducted a generalized comparison of AI & DM methods, including (1) a benchmark three-layer feed-forward Artificial Neural Network, (2) the Random Forests method, and (3) a Support Vector Regression technique, which are all commonly used approaches in the literature.

To test the robustness of different AI & DM methods, we selected two headwater reservoirs in USA and China, and conducted a comparison experiment using different AI & DM methods with various parameterizations to simulate reservoir inflows. In USA, the Trinity Lake, also known as the Clair Engle Lake (CLE), is selected. The CLE reservoir is one of the water supply sources in the Central Valley Project (CVP) a federally funded water distribution project. The goal of CVP is to divert water from the northern part of California, where water supply is relatively abundant, to the water-scarce areas in the central and southern parts of the state for irrigation and municipal water supply purposes. The CVP is jointly operated in coordination with the California State Water Project (SWP). SWP is operated by the California Department of Water Resources, and directly transports water from the northern parts to the southern parts of California for residential water uses. In China, we selected the Danjiangkou (DJK) Reservoir as another study case. The DJK reservoir is the headwater reservoir for the central route of the China's South-to-North Water Diversion Project which also transports water from the water-abundant, in this case southern parts (Han River and Yangtze River), to the water-limited northern areas of China, including the highly populated areas of Beijing and Tianjin, and Henan and Hebei provinces.

In addition, another focus of this paper is to test and quantify the predictability of different hydrological and climatic information for reservoir inflow forecasts. For example, in order to develop better reservoir inflow subseasonal and seasonal forecasts, it may be possible to identify the climate phenomenon that dominate the local hydrology, and then to incorporate climate phenomenon indicators into a given modelling framework. As pointed out by *Burley et al*. [2012] and *Turner and Galelli* [2016], the shifts in climate conditions and streamflow should be emphasized in future studies on water resources management and planning. According to many other studies [*Montoya et al*., 2014; *Kalra et al*., 2013; *Gutiérrez and Dracup*, 2001; *Garen*, 1993; *Hamlet and Lettenmaier*, 1999], certain climate phenomenon variables or indices, such as the ENSO and PDO, are potentially useful in supporting water supply planning in the western regions of United State, and these indices are predictable with lead times up to 6 months or 1 year. However, there are two key steps necessary to make operational use of climate indices for reservoir operations. The first one is how to incorporate the signals of known climate phenomenon indices into subseasonal and seasonal prediction modeling framework. Demonstrating methods to use raw climate phenomenon indices directly in regression models of streamflow makes this potentially valuable information available to decision makers and dam operators. The authors believe the AI & DM techniques are suitable tools to address the issue of using climate phenomenon indices in regression models thus assisting water-supply planning and prediction. The second issue is how to automatically identify and select the climate phenomena indices as predictors for supporting reservoir planning and scheduling, given the fact that it is unclear to decision makers which climate phenomenon indices are effective predictors, and, hence, representative for the regional climate variability. Many previous studies were conducted on using one or two climate phenomenon indices, which are already known as useful predictors in a specific region. Inspired by a recent study by *Yang et al*. [2016], in which many decision variables of operating reservoirs are automatically ranked and selected using a Gini diversity index, the climate phenomenon indices can also be ranked and compared with regard to

their predictability for reservoir operation and seasonal forecasts. In this study, we focus on developing 1 month-ahead inflow forecasts.

In a summary, the goals of this study are (1) to apply AI & DM methods to reservoir inflow predictions and investigate the signals of climate phenomenon indices in two headwater reservoirs located in the western regions of the United States, and the southern part of China, respectively; (2) to compare the predictive performances of three popular AI & DM methods, namely RF, ANN, and SVR, in assisting subseasonal to seasonal water-supply planning; and (3) to explore the sensitivity analysis of different AI & DM methods and inform decision makers regarding the usefulness and capability of various regression approaches. Specifically, the sensitivity analysis of AI & DM methods includes (i) the default stopping criterion in ANN, i.e., maximum iteration, and the number of hidden nodes as one of the structural parameters; (ii) the maximum features in developing a Random Forest, and (iii) different kernel functions and associated penalty terms in the SVR models. The methodology and approaches employed in the current work are universally applicable to other study cases and are not limited to the case studies and sites in the current work, which are the source reservoirs for the U.S. federal Central Valley Project, and the China's South-to-North Water Diversion Project, respectively.

The organization of rest of the paper is as follows: The methodologies of ANN, DT, and SVR are summarized in section 2; section 3 introduces the two study cites, reservoir operation data, climate phenomenon indices, and model inputs; The results are given in section 4; section 5 provides the discussion with regard to methodologies, parameterization, seasonal patterns, and the implications of using climate phenomenon indices; Finally, the major findings, conclusions, and future works are summarized in section 6.

2. Methodology

2.1. Artificial Neural Network (ANN)

The ANN is a powerful classification and regression algorithm that is inspired by the neurological structure of the human brain [*Jain et al*., 1996; *Hopfield*, 1988]. The concept in ANN is to interconnect input data with output data using multiple neurons as hidden layers, which are able to extract the explicit information and relationship between input and output data. ANN is extensively used in many fields of study, such as biological memory [*Kohonen*, 2012], pattern recognition [*Jain et al*., 2000], image processing [*Egmont-Petersen et al*., 2002; *Cichocki and Amari*, 2002], precipitation estimation from satellites [*Hsu et al*., 1997], ecological modeling [*Lek and Guégan*, 1999], and reservoir operation [*Cheng et al*., 2015; *Jain et al*., 1999; *Ashaary et al*., 2015; *Shamim et al*., 2016; *Coulibaly et al*., 2001]. In this study, a Three-Layer Feed-Forward Neural Network (TLFFNN) is used in combination with a backpropagation learning algorithm [*Werbos*, 1974] (Figure 1).

A typical TLFFNN consists of an input layer, a hidden layer, and an output layer. The inputs data $\vec{x}(x_1, x_2, \ldots, x_{n_0})$ and output data $\vec{z}(z_1, z_2, \ldots, z_{n_2})$ are connected by a hidden layer $\vec{h}(h_1, h_2, \ldots, h_{n_1})$, where n_0, n_1 and n_2 represent the total number of inputs, hidden neurons, and outputs. The inputs are connected to the hidden layer by a transfer function (f), which is typically based on the weighted sum of inputs as shown in equation (1).

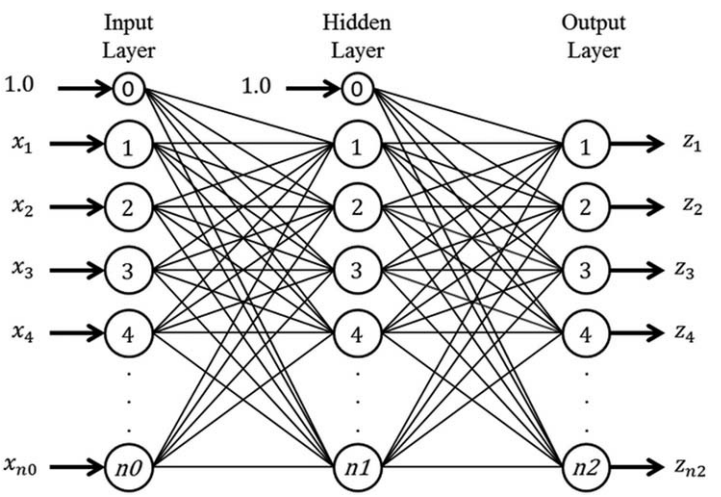

Figure 1. A three-layer feed-forward neural network (TLFFNN)

$$h_j = f\left(\sum_{i=1}^{n_0} w_{ij}x_i + w_{0j}\right) \quad (1)$$

where h_j is the j-th neuron in the hidden layer; x_i is the i-th input; w_{ij} represents the weight assigned to the i-th input in order to calculate the j-th hidden neuron; and $j \in (1, 2, \ldots n_1)$ and $i \in (1, 2, \ldots n_0)$. Similarly, another

transfer function (g) is used to connect hidden neurons to the outputs as shown in equation (2).

$$z_k = g\left(\sum_{j=1}^{n_1} v_{jk} h_j + v_{0k}\right) \tag{2}$$

where z_k is the k-th value in the output layer; h_j is the j-th neuron in the hidden layer; v_{jk} represents the weight assigned to the j-th neuron in the hidden layer; and $k \in (1, 2, \ldots n_2)$ and $j \in (1, 2, \ldots n_1)$. In TLFFNN, functions f and g are called activation functions. In this study, we use the Hyperbolic Tangent function (equation (3)) and linear function (equation (4)) as the activation functions to connect inputs to the hidden layer and hidden neurons to outputs, respectively.

$$y_j = \frac{2}{1 + e^{-2\sum_{i=1}^{n_0} w_{ji} x_i + w_{j0}}} - 1 \tag{3}$$

$$z_k = \sum_{j=1}^{n_1} v_{jk} h_j + v_{0k} \tag{4}$$

The weights w_{ij} and v_{jk} are obtained by minimizing the Sum of Square Errors (SSE) between model output and the target variable (equation (5)). The optimization problem (equation (5)) is solved using a Gradient Descent method, and backpropagation scheme [*Kişi*, 2007]. The detailed error propagation from the output layer to the hidden layer is shown in supporting information Appendix A1.

$$\text{argmin}(\text{SSE}) = \text{argmin}\left[\frac{1}{2}\sum_{k=1}^{n_2} (t_k - z_k)^2\right] \tag{5}$$

where t_k is the value of node k in the output layers, and z_k is the corresponding target value for node k in the ANN training phase.

2.2. Random Forests (RF)

Among different types of DT methods, including the Classification and Regression Tree (CART) [*Breiman et al*., 1984], ID3 [*Quinlan*, 1986], ID4.5 [*Quinlan*, 1990], Bagging-Tree [*Breiman*, 1996], and RF [*Breiman*, 2001], etc., we chose the RF algorithm, because of its wide and successful application in reservoir operations [*Yang et al*., 2016; *Wei and Hsu*, 2009; *Li et al*., 2010; *Goyal et al*., 2013a, 2013b] as well as in many other fields [*Araújo and New*, 2007; *Chebrolu et al*., 2005; *Elith and Leathwick*, 2009; *Fayyad et al*., 1996]. Differing from ANN as a black-box regression model, RF is a nonparametric, white-box classification and regression model first introduced by *Breiman* [2001]. According to *Yang et al*. [2016], *Breiman* [2001], and *Hancock et al*. [2005], the key concept of the RF algorithm is a random selection of decision variables combined with an ensemble approach with regard to the final classes using multiple standard regression trees, or the standard CART algorithm [*Breiman et al*., 1984]. The mechanism in a standard tree-growing process is to recursively carry out a binary split of a data set ("Parent Node") into two subdata sets ("Child Nodes") based on any user-defined splitting rule. In RF, each split is carried out using a random selection of decision variables instead of all decision variables [*Liaw and Wiener*, 2002]. There are several steps required to build a RF model. Given a set of decision variables (inputs or predictors) and target variables (outputs), the RF algorithm first searches all values of a randomly selected decision variables and identifies a threshold for one decision variable that can maximize the homogeneity, or in other words gives the lowest misclassification rate, in target variable space. The maximization of homogeneity is defined by the splitting rules, such as Root-Mean-Square Errors, Mean Square Errors, or the Gini Diversity Index. Then, the target variables and the associated decision variables are reorganized into two subdata sets, called "Left Child Node" and "Right Child Node" respectively, using this identified decision variable threshold. The same procedure of splitting a "Parent Node" into two "Child Nodes" is carried out on both of the subdata sets (the "Left Child Node" and "Right Child Node") to further divide the subdata sets into smaller data sets. This step is recursively conducted to expand the size of a tree. When certain stopping criteria are met, such as the maximum tree size or the minimum of members in a Child Node is reached, the tree-growing process is terminated. Finally, when multiple single trees are built, the trees with low predictive performances are eliminated from the candidate group, and the ensemble of elite candidates constitutes the final RF model. In this study, we primarily use the regression tree, which takes the average of target variable values in each class as predicted value, given predictors that satisfy the identified classification rules. The counterintuitive strategies of randomly selecting

decision variables and the ensemble of elite candidate trees employed in RF turn out to be robust against overfitting [*Breiman*, 2001].

2.3. Support Vector Regression (SVR)

The Support Vector Regression (SVR) belongs to the regression scheme of the Support Vector Machine (SVM). In general, the SVR has the identical methodology as the SVM with only slight differences. The SVM is one type of the supervised classification methods being developed from convex optimization [*Vapnik*, 1999; *Smola and Schölkopf*, 2004]. The methodology in SVM involves two steps. The first step is to project the input data (or features) to a high-dimensional feature space, in which the classification task becomes easier than it is in the original input (feature) space. The second step is to find a global optimal hyperplane to split the data by evaluating the offsets of each data point to this hyperplane [*Vapnik*, 1999]. The philosophy of the second step is to maximize the margin of each class of data points to the optimal hyperplane. Given the entire data set $\{x_i, y_i\}_{i=1}^{I}$, where I is the number of data points, we assume that there exists a function $f(x)$ describing the nonlinear relationship between features x_i and target values y_i, as shown in the following equation (6):

$$f(x)=(w\cdot\varphi(x))+b \tag{6}$$

In the function above, $\varphi(x_i)$ is the transformation function projecting the input data into a high-dimensional feature space. w is the weight vector and b is called the offset factor [*Vapnik*, 1999]. The goal is to maximize the margin between any two classes with respect to w and b, as shown in equation (7):

$$\min_{w,b}\left(\frac{1}{2}\|w\|^2+C\sum_{i=1}^{I}(\xi_i+\xi_i^*)\right) \tag{7}$$

Subject to

$$\begin{cases} y_i-(w\cdot\varphi(x_i)+b)\le\varepsilon+\xi_i \\ (w\cdot\varphi(x_i)+b)-y_i\le\varepsilon+\xi_i^* \\ \xi_i,\ \xi_i^*\ge 0,\ i=1,\ldots I \end{cases} \tag{8}$$

where, C is called the penalty coefficient which is a user-defined constant representing the amount of trade-off between dispersion of weights and objective function. The ξ_i and ξ_i^* are termed as the slack variables that describe how much the exceeding of data from a small tolerance variable ϵ [*Smola and Vapnik*, 1997]. A Lagrangian multiplier is further applied the regression function (equation (5)) to replace the weight vector w and the transformation function $\varphi(x_i)$ as shown in equation (9):

$$f(x)=\sum_{i=1}^{I}(a_i-a_i^*)K(x,x_i)+b \tag{9}$$

where a_i and a_i^* are the Lagrangian multipliers and K is called the kernel function. The quadratic form of equation (9) can be expressed as:

$$W(a_i,a_i^*)=\sum_{i=1}^{I}y_i(a_i-a_i^*)-\varepsilon\sum_{i=1}^{I}(a_i+a_i^*)-\frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{I}(a_i-a_i^*)(a_j-a_j^*)K(x_i,x_j) \tag{10}$$

with the constraints of equation (11):

$$\begin{cases} \sum_{i=1}^{I}(a_i-a_i^*)=0 \\ 0\le a_i\le C,\ i=1,\ldots,I \\ 0\le a_j^*\le C,\ i=1,\ldots,I \end{cases} \tag{11}$$

The kernels tested in this study include the Linear, Polynomial, Gaussian, and Sigmoid kernels, which all belong to the commonly used kernel functions in SVR as shown in equation (12)

$$\begin{cases} K(x,x_i)=x\cdot x_i \\ K(x,x_i)=(\gamma(x\cdot x_i)+r)^d \\ K(x,x_i)=\exp\left(-\gamma|x-x_i|^2\right) \\ K(x,x_i)=\tanh(\gamma(x\cdot x_i)+r) \end{cases} \tag{12}$$

, where d is the degree of the polynomial terms, r represents the residuals, and γ is structural parameter in the polynomial, Gaussian and Sigmoid kernels. Different d, γ, and penalty coefficient C will be tested in the experiments.

3. Data and Settings

3.1. Reservoir and Operation Data

The locations of the CLE and DJK reservoirs are shown in Figure 2. The total storage capacities for CLE and DJK reservoir are 30×10^8 m^3, and 290×10^8 m^3, respectively. CLE reservoir is the second largest reservoir in the CVP, and the DJK reservoir the largest freshwater lake in Asia. Water from CLE reservoir is primarily used for hydropower generation and irrigating farm lands, and the water from DJK reservoir is primarily used for residential consumption in the northern part of China. The daily inflow, point accumulated precipitation, and computed daily lake evaporation for CLE reservoir are retrieved from the California Data Exchange Center (CDEC). The length of CLE data is over 50 years from 1964 to 2015. The daily inflow to the DJK reservoir is retrieved from the Bureau of Hydrology, Changjiang Water Resources Commission (http://www.cjh.com.cn/), and aggregated into monthly scale. Other hydrological information for the DJK reservoir are obtained from the China Meteorological Administration, and a number of local water resources agencies through the senior author's personal communication. The DJK data covers from the period of 1961–2010.

3.2. Climate Phenomenon Indices

Besides the local hydrology information and reservoir operation data, 17 climate phenomenon indices have been selected, and the monthly values are retrieved from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA-ESRL) (http://www.esrl.noaa.gov/psd/data/climateindices/list/). The selection criterion is based on the temporal coverage of a climate phenomenon index, in which an index must be continually updated from at least 1960s to the present, and those with only limited temporal coverage are neglected. Table 1 lists the detailed information about the selected climate phenomenon indices. In general, the selected climate phenomenon indices are associated with many climate and

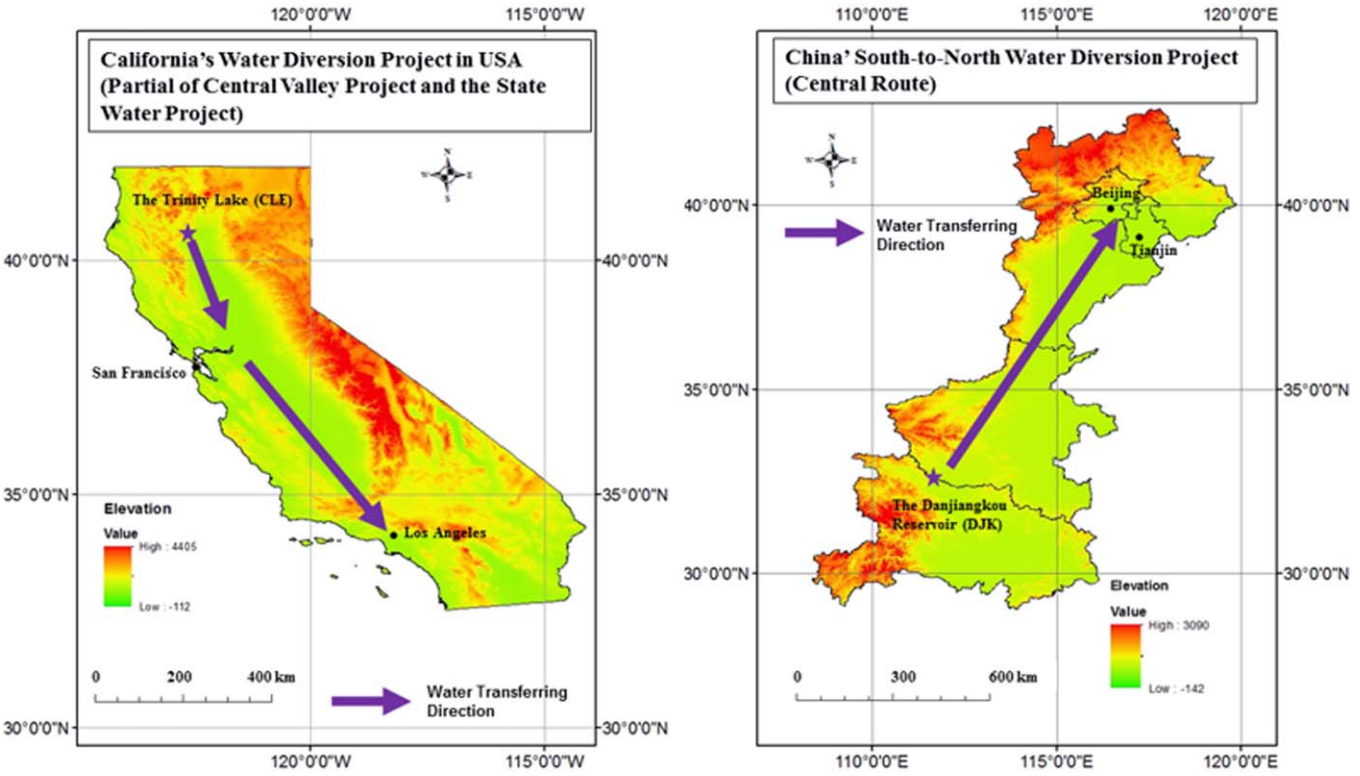

Figure 2. Location of the Trinity Lake (CLE) in USA and the Danjiangkou (DJK) Reservoir in China. The water diversion directions are shown as purple arrows.

Table 1. Information of Selected Climate Phenomenon Indices

Indices	Name	Source
PNA	Pacific North American Index	NOAA Climate Prediction Center (CPC)
WP	Western Pacific Index	NOAA Climate Prediction Center (CPC)
NAO	North Atlantic Oscillation	NOAA Climate Prediction Center (CPC)
SOI	Southern Oscillation Index	NOAA Climate Prediction Center (CPC)
WHWP	Western Hemisphere Warm Pool	*Wang and Enfield* [2001]
ONI	Oceanic Nino Index	NOAA Climate Prediction Center (CPC)
MEI	Multivariate ENSO Index	*Wolter and Timlin* [1998]
Nino1 + 2	Extreme Eastern Tropical Pacific SST (0-10S, 90W-80W)	CPC
Nino 3	Eastern Tropical Pacific SST (5N-5S)(150W-90W)	CPC
Nino 34	East Central Tropical Pacific SST (5N-5S)(170-120W)	CPC
Nino 4	Central Tropical Pacific SST (5N-5S) (160E-150W)	CPC
PDO	Pacific Decadal Oscillation	*Zhang et al.* [1997]
TNI	Indices of El Niño Evolution	*Trenberth and Stepaniak* [2001]
AO	First leading mode from the EOF analysis of monthly mean height anomalies	CPC
QBO	Quasi-Biennial Oscillation	NCEP/NCAR Reanalysis
CENSO	Bivariate ENSO Time series	Standardized SOI and standardized Nino3.4 SST Time Series
EPO	East Pacific/North Pacific Oscillation	NOAA Climate Prediction Center (CPC)
Seasonality	Month of a Year	Calendar

atmospheric activities, including teleconnections, atmosphere, ENSO, and the variations of surface temperatures of both the Pacific and Atlantic oceans.

3.3. Model Inputs

The regression for reservoir inflow is conducted using both current time step hydrological and climatic information and their lagging information up to two time steps. Specifically, the inputs for predicting current monthly inflow include (1) past inflow (1 month-lag and 2 month-lag); (2) current and previous (1 month-lag and 2 month-lag) accumulated precipitation; (3) current and previous (1 month-lag and 2 month-lag) lake evaporation; and (4) 17 monthly climate phenomenon indices at current and past (1 month-lag and 2 month-lag). Therefore, there are 62 model inputs included in the regression mode. Figure 3 shows the model inputs and output structure for the inflow regression. The inputs structure is set to be identical for all methods, including ANN, RF, and SVR experiments in order to produce a fair comparison.

The approach that treats the previous month's information as decision variables is similar to the one employed in *Hejazi and Cai* [2011], in which both the current and past month's inflow and release are identified as important state variables to enhance model performances. In addition, we also extend the investigation of climate indices up to 2 months-lag. The inflow is believed to be a result of climate and hydrological processes, and the previous months also have influences on the current month inflow due to mass continuity of snowmelt and seasonality.

3.4. Parameterization and Setting

In order to illustrate the strengths and weaknesses of different AI & DM methods, different parameterizations are used in ANN, RF, and SVR. In ANN, three sets of stopping criterion are tested, including 100, 1000, and 2000 iterations, in combination with different numbers of hidden nodes ranging from 5 to 20. In RF, based on some findings of using CART on the CLE reservoir presented in *Yang et al.* [2016], the maximum tree depth is set to 20, and the number of candidate trees is set to 100 for all simulations. One structural parameter of RF, namely the maximum features in growing candidate trees, is investigated in this study. The range of maximum features is set to vary from 5 to 60 for different RF simulation. Last, in SVR, four types of kernel functions are compared, including linear, RBF, Polynomial, and Sigmoid kernel functions in combination with penalty coefficients ranging from 0.001 to 100. Detailed parameter values are shown in the results and discussion sections.

In order to prevent the regression model from overfitting the data, in this study, the K-fold method is chosen for its simplicity and robustness. First, all the data are split equally into K subdata sets. Then, one subdata set is temporarily held out for testing, and the remaining subdata sets (total K-1 subdata sets) are used to train the predictive model (ANN, RF, or SVR). The performance of the trained model is evaluated on the

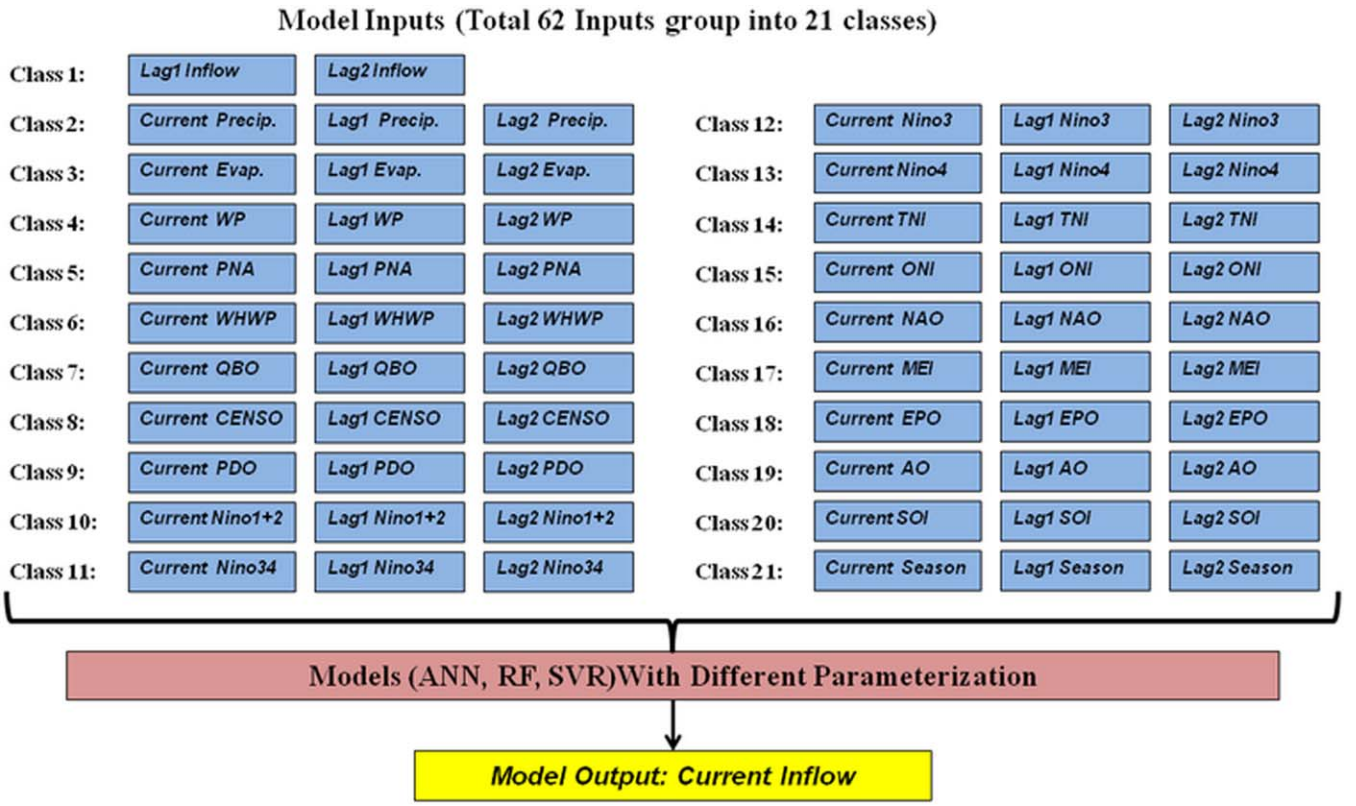

Figure 3. Structure of model inputs and output.

testing subdata set, and the results are achieved. The processes of training and testing are repeated using another subdata set and the remaining data until all subdata sets have been used as the testing data set. If the model performance on each fold of the testing data set is equally good, the regression model is identified as proper. In this study, identical K-fold cross-validation schemes are carried out in training the ANN, RF, and SVR models with K value equals to 51 and 49 for CLE and DJK reservoir, respectively. For example, for CLE reservoir, the reservoir inflow data in 1964 is temporarily held out for testing, and the predictive models are trained with the data from 1965 to 2015. After the prediction performances with ANN, RF, and SVR are tested on the data in 1964, another year (i.e., 1965) is selected to be the testing data set and the same model training procedure is carried out until all years of data have been used as testing data sets. Besides the K-fold cross-validation method used in this study, there are many other types of cross-validation schemes available in the literature, including the hold-out method, the leave-one-out method, the leave-p-out method, etc. [*Allen*, 1974; *Arlot and Celisse*, 2010; *Geisser*, 1993; *Stone*, 1977]. The general goal of all methods is to prevent overfitting, but the comparison of different cross-validation schemes are beyond the scope of this paper.

4. Results

4.1. Predictions

The reservoir inflow predictions using ANN for the CLE and DJK reservoirs are shown in Figures 4 and 5, respectively. The results obtained from RF for the CLE and DJK reservoir are shown in Figures 6 and 7, respectively. The predictions from SVR are presented in Figures 8 and 9, for CLE and DJK reservoir, respectively. It is worth mentioning the results for each single year as the testing year are obtained by an individual regression model calibrated without the testing year. Therefore, the demonstrated results in Figures 4–9 consist of 51 and 49 independent verification years for CLE and DJK reservoir, respectively. As shown in Figures 4–9, the general variation and magnitude of predicted inflows from ANN and RF have better agreement with the observations as compared to the results from SVR for both CLE and DJK reservoir. In SVR

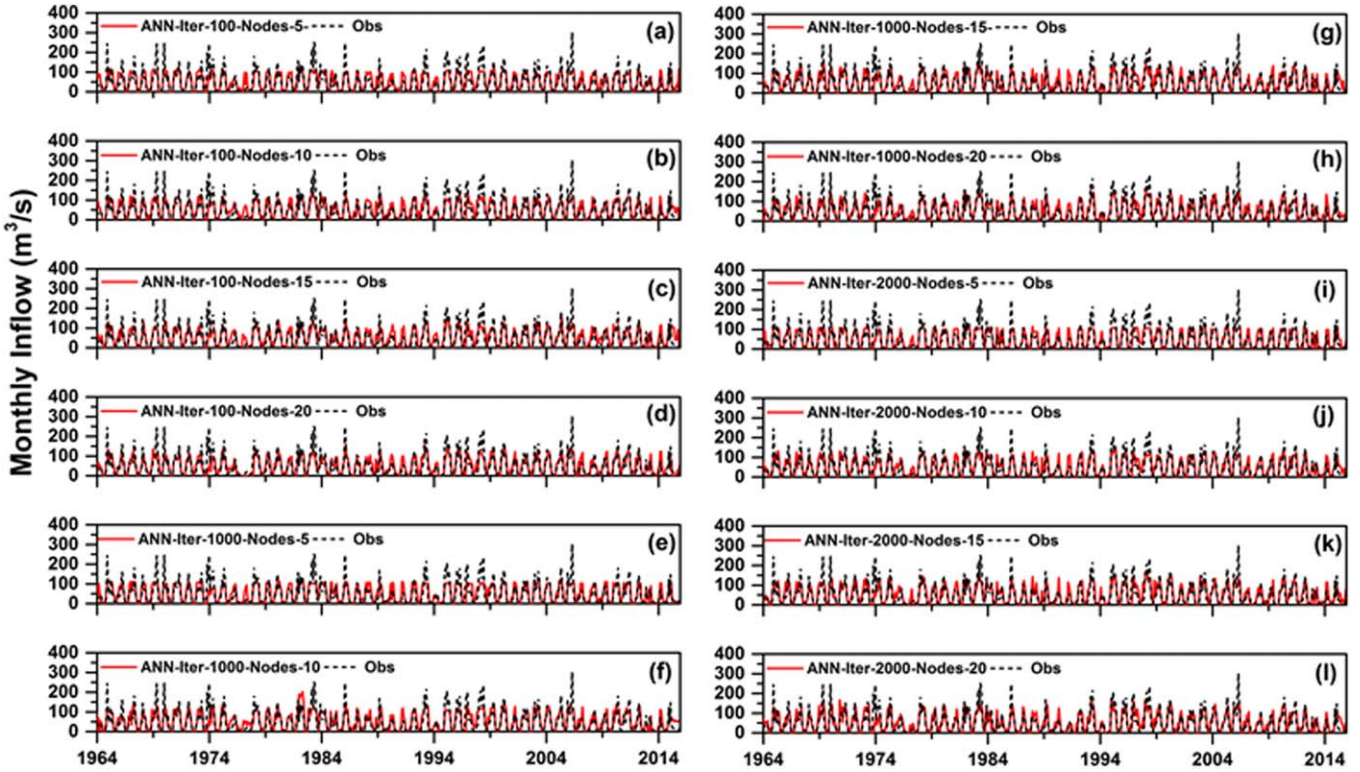

Figure 4. Comparison of predicted and observed inflow using ANN with different maximum iteration (Iter) and number of hidden nodes (H) for the CLE reservoir.

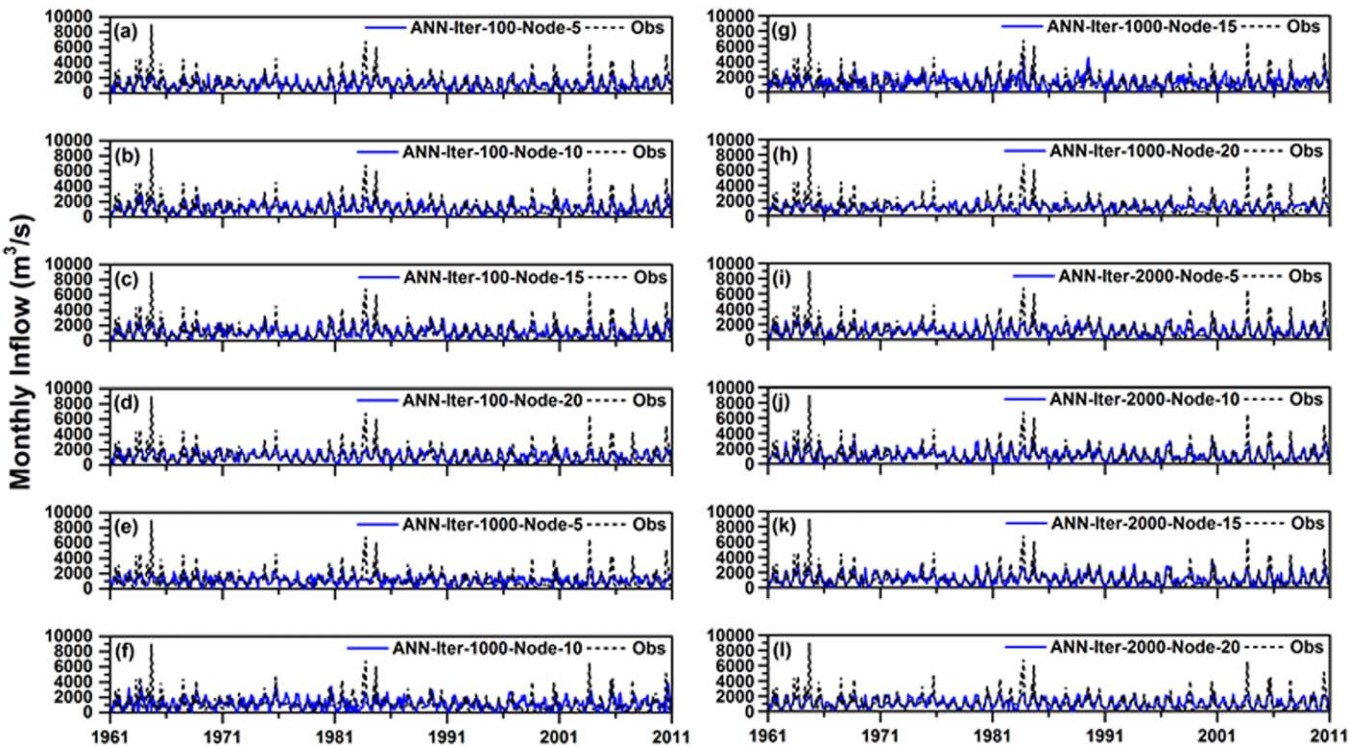

Figure 5. Comparison of predicted and observed inflow using ANN with different maximum iteration (Iter) and number of hidden nodes (H) for the DJK reservoir.

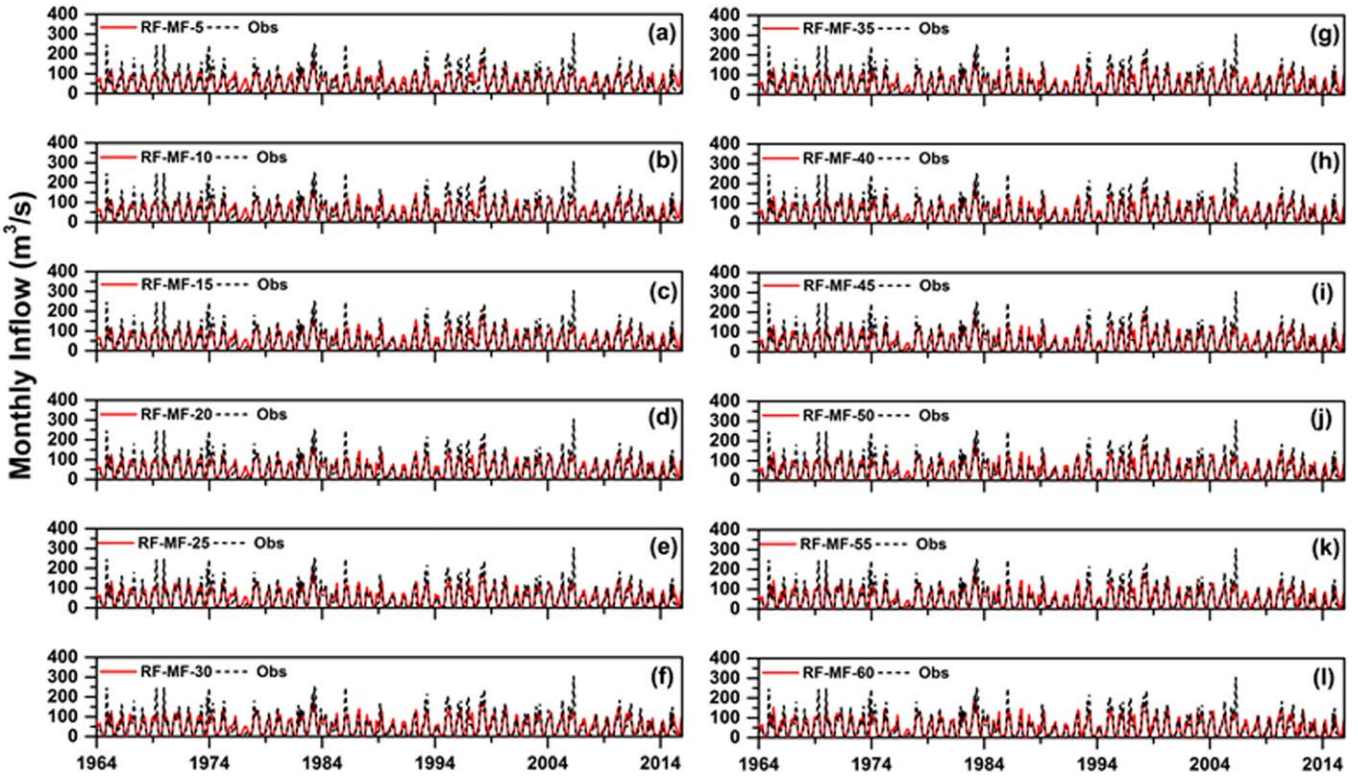

Figure 6. Comparison of predicted and observed inflow using RF with maximum features (MF) for the CLE reservoir Figure 7. Comparison of predicted and observed inflow using RF with maximum features (MF) for the DJK reservoir Figure 8. Comparison of predicted and observed inflow using SVR with RFB, Linear, Polynomial, and Sigmoid kernel functions, different penalty coefficient (c), and gamma (γ) for the CLE reservoir.

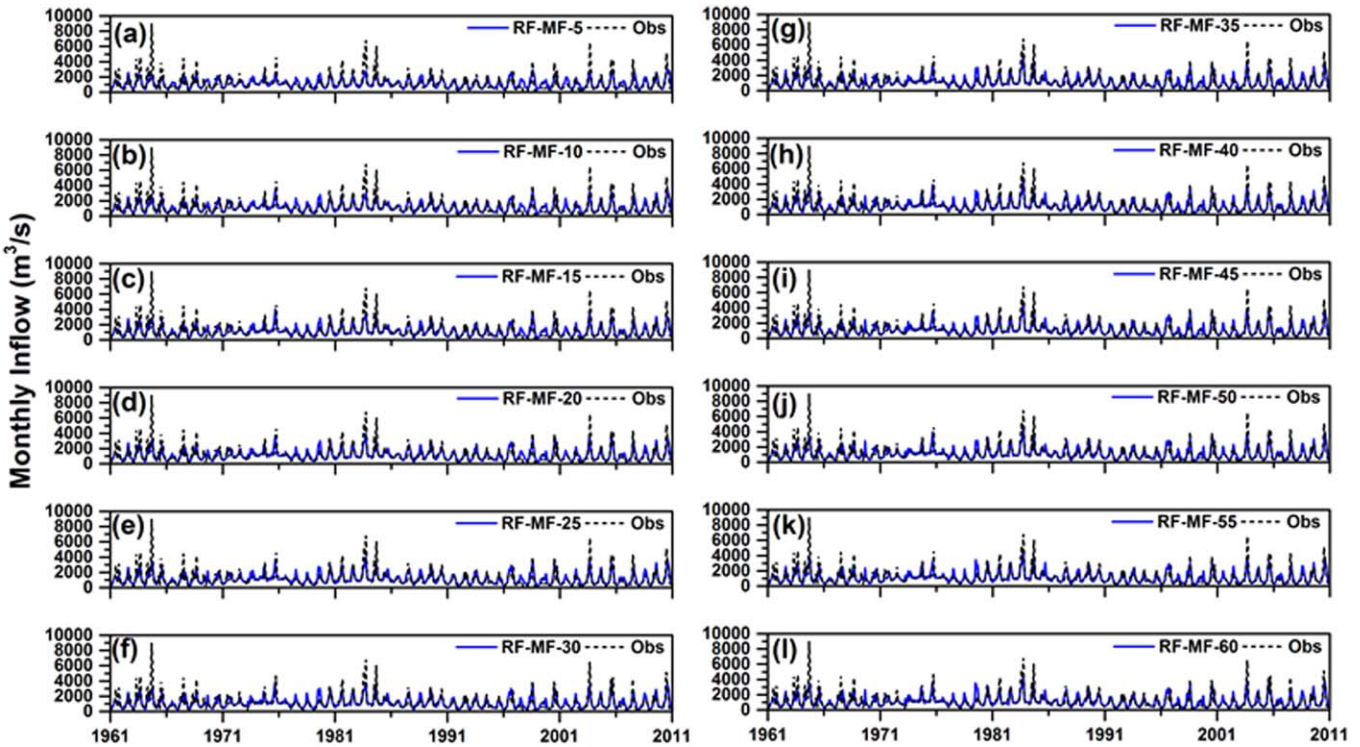

Figure 7. Comparison of predicted and observed inflow using RF with maximum features (MF) for the DJK reservoir.

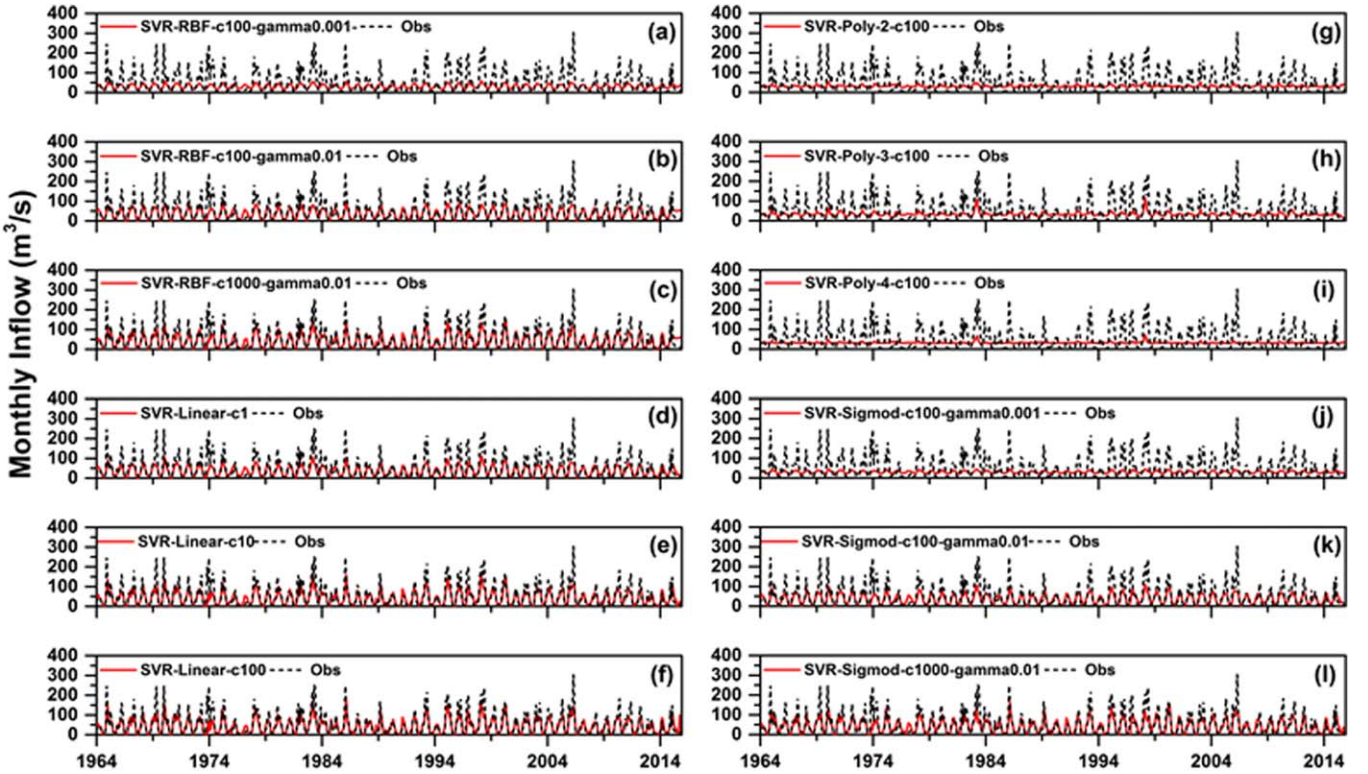

Figure 8. Comparison of predicted and observed inflow using SVR with RFB, Linear, Polynomial, and Sigmoid kernel functions, different penalty coefficient (c), and gamma (γ) for the CLE reservoir.

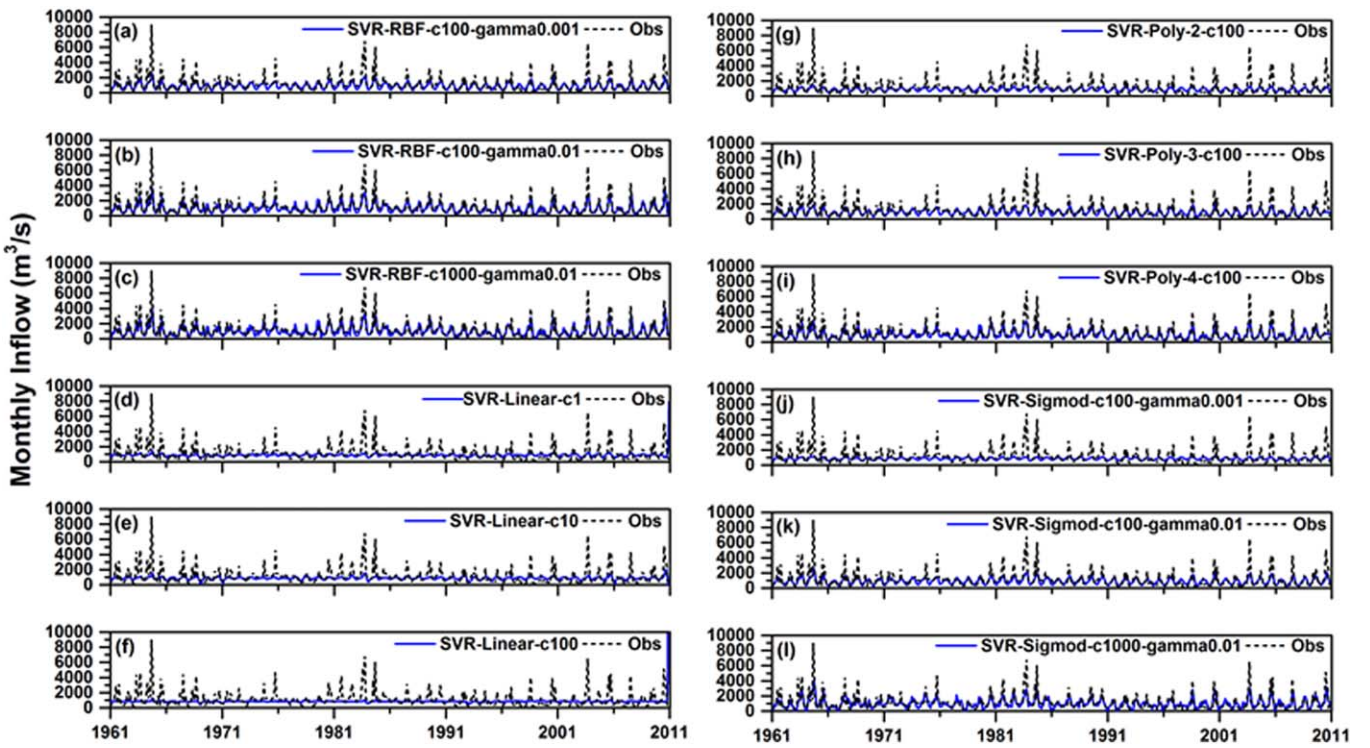

Figure 9. Comparison of predicted and observed inflow using SVR with RFB, Linear, Polynomial and Sigmoid kernel functions, different penalty coefficient (c), and gamma (γ) for the DJK reservoir.

(Figures 8 and 9), the results using RBF, Polynomial, and Sigmoid kernels with a high penalty (c = 1000) are significantly closer to observations when compared to those with other settings. Furthermore, the RBF kernels tend to generate more accurate simulations than other kernels (Figures 8 and 9) for both the CLE and DJK reservoirs. In ANN and RF (Figures 4–7), the differences are less significant among the results using various parameters. For the CLE reservoir, the peak inflows during 2006–2007 are underestimated by all models (Figures 4, 6 and 8). For the DJK reservoir, the high flows around 1963, 1982–1984, and 2003 are consistently underestimated by all models (Figure 5, 7 and 9). According to Figure 6, the second largest peak inflows happened on the CLE reservoir during 1983–1984 are well captured only by the RF algorithm, whereas the prediction results from the other two methods still remain underestimated (Figures 4 and 8). Similarly, the peak flows during 1982–1984 for the DJK reservoir can be better simulated by the RF method (Figure 7) than that from ANN (Figure 5) and SVR (Figure 9).

In order to mathematically quantify the predictive performances of ANN, RF, and SVR on both CLE and DJK reservoirs, four statistical measures are calculated: the Root Mean Square Error (RMSE), Correlation Coefficient (CORR), the Nash-Sutcliffe Model Efficiency Coefficient (NSE) [*Nash and Sutcliffe*, 1970], and the Kling-Gupta efficiency (KGE) [*Gupta et al*., 2009]. The KGE is an enhanced statistical measure of the NSE. As pointed by a recent study by *Revilla-Romero et al*. [2015], the KGE is able to avoid some limitations of using the NSE, because it decomposes the NSE into three independent components to present linear correlation (r), variability (α), and bias ratio (β) between the simulations and observations. The equations of the selected statistics are shown in equation (13)–(16):

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=0}^{n}\left(Q_{sim,i}-Q_{obs,i}\right)^2}{n}} \tag{13}$$

$$\mathrm{CORR}=\frac{\sum_{i=0}^{n}\left(\left(Q_{sim,i}-\bar{Q}_{sim,i}\right)\left(Q_{obs,i}-\bar{Q}_{obs,i}\right)\right)}{\sqrt{\sum_{i=0}^{n}\left(Q_{sim,i}-\bar{Q}_{sim,i}\right)^2\sum_{i=0}^{n}\left(Q_{obs,i}-\bar{Q}_{obs,i}\right)^2}} \tag{14}$$

$$NSE=1-\frac{\sum_{i=0}^{n}\left(Q_{obs,i}-Q_{sim,i}\right)^2}{\sum_{i=0}^{n}\left(Q_{obs,i}-\bar{Q}_{obs,i}\right)^2} \tag{15}$$

$$KGE=1-\sqrt{(1-r)^2+(1-\alpha)^2+(1-\beta)^2} \tag{16}$$

where $Q_{obs,i}$ and $Q_{sim,i}$ are the observed and simulated values for inflows during the *i*-th month; $\bar{Q}_{obs,i}$ and $\bar{Q}_{sim,i}$ are the mean of the observed and simulated values; and *n* is the total number of observations. In equation (16), $\alpha=\sigma_s/\sigma_o$, and $\beta=\mu_s/\mu_o$, where μ_s and σ_s are the mean and standard deviation of the simulations, respectively, and μ_o and σ_o are the mean and standard deviation of the observations, respectively. The KGE values are bounded to negative infinity to 1, in which a KGE = 1 indicates a perfect match between simulated and observed values [*Gupta et al*., 2009].

The calculated RMSE, CORR, NSE, and KGE between observed and simulated inflows obtained from ANN, RF, and SVR for both CLE and DJK reservoir are shown in Tables 2–4, respectively. In Tables 2–4, the bolded and underlined values indicate the best statistics with regards to different parameterizations for each AI & DM method. For CLE reservoir (Tables 2–4), the best NSE values obtained by ANN, RF, and SVR during the entire water years are 0.613, 0.691, and 0.627, respectively, and the best KGE values for ANN, RF, and SVR are 0.732, 0.737, and 0.583, respectively. For DJK reservoir (Tables 2–4), the best NSE values achieved by ANN, RF, and SVR during the entire water years are 0.440, 0.620, and 0.603, respectively, and the best KGE values for ANN, RF, and SVR are 0.578, 0.645, and 0.583, respectively. As suggested by *Moriasi et al*. [2007], a KGE value or a NSE value that is larger than 0.50 can be considered as satisfactory model performance for monthly time series. Therefore, we can conclude that except the ANN simulation on the DJK reservoir, other AI & DM methods are satisfactory in simulating the monthly reservoir inflow. Furthermore, the employed AI & DM methods in the CLE reservoir tend to produce better statistics during runoff season than the flood control season, while for the DJK reservoir the simulation during flood control season is slightly better than the runoff season. The results from RF (Table 3) on both CLE and DJK reservoir suggest that the RF has a consistently good performance for inflow prediction during both flood control and runoff seasons. According to Table 4, SVR only provides satisfactory performance (both KGE and NSE values are larger than 0.50)

Table 2. Statistical Performances of ANN With Different Maximum Iteration (MI) of 100, 1000, and 2000, and Different Numbers of Hidden Nodes (H) Ranging From 5 to 20 for the Trinity Lake (CLE) and the Danjiangkou Reservoir (DJK)

	RMSE-WY[a] (m^3/s)	RMSE-Runoff[a] (m^3/s)	RMSE-FC[a] (m^3/s)	CORR-WY	CORR-Runoff	CORR-FC	NSE-WY	NSE-Runoff	NSE-FC	KGE-WY	KGE-Runoff	KGE-FC
					Reservoir CLE							
MI = 100 H = 5	38.503	33.805	44.472	0.720	0.780	0.579	0.516	0.605	0.321	0.624	0.646	0.478
MI = 100 H = 10	36.777	33.109	41.315	0.748	0.792	0.647	0.559	0.621	0.414	0.642	0.650	0.529
MI = 100 H = 15	37.201	34.722	40.919	0.742	0.764	0.654	0.548	0.583	0.426	0.638	0.660	0.504
MI = 100 H = 20	36.633	31.799	43.086	0.750	0.809	0.614	0.562	0.651	0.363	0.660	0.681	0.530
MI = 1000 H = 5	37.379	33.186	43.608	0.739	0.788	0.602	0.544	0.619	0.348	0.656	0.677	0.512
MI = 1000 H = 10	37.271	32.135	43.867	0.745	0.803	0.610	0.547	0.643	0.340	0.685	0.711	0.549
MI = 1000 H = 15	34.955	32.230	39.448	0.777	0.801	0.687	0.601	0.641	0.466	0.705	0.717	0.597
MI = 1000 H = 20	**34.437**[b]	**30.198**[b]	41.064	**0.784**[b]	**0.830**[b]	0.659	**0.613**[b]	**0.685**[b]	0.421	0.704	0.723	0.578
MI = 2000 H = 5	36.634	32.394	42.685	0.750	0.800	0.623	0.562	0.637	0.375	0.664	0.677	0.540
MI = 2000 H = 10	35.422	31.925	40.917	0.770	0.805	0.661	0.591	0.648	0.426	0.700	0.723	0.576
MI = 2000 H = 15	35.165	30.345	41.002	0.776	0.826	0.664	0.596	0.682	0.423	0.716	0.745	0.593
MI = 2000 H=20	34.794	30.367	**40.200**[b]	0.782	0.826	**0.682**[b]	0.605	0.681	**0.446**[b]	**0.732**[b]	**0.758**[b]	**0.626**[b]
					Reservoir DJK							
MI = 100 H = 5	906.477	1125.785	610.729	0.481	0.393	0.264	0.191	0.117	−1.619	0.390	0.233	0.083
MI = 100 H = 10	910.486	1135.844	622.974	0.481	0.373	0.272	0.184	0.102	−1.725	0.397	0.225	0.062
MI = 100 H = 15	902.451	1131.967	582.449	0.497	0.404	0.348	0.198	0.108	−1.382	0.424	0.275	0.142
MI = 100 H = 20	922.125	1156.159	604.325	0.447	0.326	0.265	0.163	0.069	−1.564	0.340	0.141	0.095
MI = 1000 H = 5	992.412	1255.939	627.834	0.314	0.148	0.202	0.030	0.198	−1.767	0.193	0.135	0.040
MI = 1000 H = 10	1058.636	1302.863	747.315	0.295	0.176	0.118	−0.104	0.182	−2.921	0.243	0.085	−0.185
MI = 1000 H = 15	1002.892	1220.518	744.763	0.412	0.281	0.182	0.010	0.037	−2.894	0.370	0.187	−0.185
MI = 1000 H = 20	923.154	1192.981	524.940	0.409	0.226	0.239	0.161	0.009	0.135	0.219	0.171	0.017
MI = 2000 H = 5	783.746	1015.085	456.029	0.633	0.532	0.415	0.395	0.282	0.160	0.530	0.378	0.328
MI = 2000 H = 10	803.574	1052.657	429.510	0.605	0.480	0.412	0.364	0.228	−0.295	0.474	0.235	0.404
MI = 2000 H = 15	754.230	978.675	425.234	**0.667**[b]	**0.578**[b]	**0.497**[b]	**0.440**[b]	**0.333**[b]	0.270	**0.578**[b]	**0.416**[b]	0.452
MI = 2000 H=20	**762.717**[b]	**995.741**[b]	**420.338**[b]	0.657	0.558	0.497	0.427	0.310	**0.240**[b]	0.560	0.391	**0.460**[b]

[a]WY stands for all the water years of the entire study period. Runoff Season includes April to September and FC stands for Flood Control Season which starts from October to next year March.

[b]The bold and underlined values indicate the best statistics for each column.

with a certain combination of kernel functions and penalty. Comparing different methods (Tables 2–4), in general, the results obtained from RF have the best and the most stable statistical performances compared with the other two methods. The results from ANN in certain experimental settings could reach similar KGE values as that from RF, but the RMSE, CORR, and NSE are slightly worse. For the CLE reservoir, the general ranking among the three methods is RF > ANN > SVR when comparing the best performing parameterizations. For the DJK reservoir, the best performing models rank as RF > SVR > ANN. More discussion on the parameterization of each method is provided in the discussion section.

In both the CLE and the DJK reservoir areas, the runoff and flood control seasons are from April to September, and October to March the following year, respectively. The inflow volumes, variability, and the associated reservoir scheduling rules are different between these two periods. According to Tables 2–4, the model performance during runoff and flood control seasons are also different. During runoff seasons, based on NSE and KGE values, the best performances of models rank as RF (NSE = 0.763 and KGE = 0.769) > ANN (NSE = 0.685 and KGE = 0.758) > SVR (NSE = 0.627 and KGE = 0.517) for the CLE reservoir, and RF (NSE = 0.504 and KGE = 0.514) > SVR (NSE = 0.503 and KGE = 0.477) > ANN (NSE = 0.333 and KGE = 0.416) for the DJK reservoir. Based on the KGE values during the flood control season, model ranks as DT (KGE = 0.629) > ANN (KGE = 0.626) > SVR (KGE = 0.542) for the CLE reservoir, and DT (KGE = 0.673) > SVR (KGE = 0.571) > ANN (KGE = 0.460) for the DJK reservoir. However, if only considering the NSE values during the flood control season, the ranking of best model performances become RF (NSE = 0.568) > SVR (NSE = 0.489) > ANN (NSE = 0.446) for the CLE reservoir, and RF (NSE = 0.618) > SVR (NSE = 0.437) > ANN (NSE = 0.240) for the DJK reservoir. Similarly, the simulated inflows to the CLE reservoir from RF give the best CORR values during flood control season (CORR = 0.754), followed by that from SVR (CORR = 0.721) and ANN (CORR = 0.682), and the best CORR values for the DJK reservoir follow as RF (CORR = 0.788) > SVR (CORR = 0.669) > and ANN (CORR = 0.497). In a summary, the uses of calibrated RF for predicting 1 month-ahead reservoir inflow are suggested, because the results given by RF are consistently better than ANN or SVR according to our experiment results in Tables 2–4.

Table 3. Statistical Performances of Random Forests With Different Maximum Features Ranging From 5 to 60 for the Trinity Lake (CLE) and the Danjiangkou Reservoir (DJK)

	RMSE-WY[a] (m/s)	RMSE-Runoff[a] (m/s)	RMSE-FC[a] (m/s)	CORR-WY	CORR-Runoff	CORR-FC	NSE-WY	NSE-Runoff	NSE-FC	KGE-WY	KGE-Runoff	KGE-FC
						Reservoir CLE						
MF = 5	35.230	27.974	40.924	0.791	0.844	0.658	0.618	0.730	0.425	0.641	0.704	0.453
MF = 10	36.270	30.696	42.946	0.763	0.834	0.613	0.571	0.675	0.367	0.588	0.649	0.387
MF = 15	33.811	26.920	37.377	0.792	0.870	0.726	0.610	0.750	0.521	0.696	0.746	0.556
MF = 20	32.233	26.395	37.068	0.818	0.875	0.738	0.652	0.759	0.541	0.706	0.761	0.587
MF = 25	32.993	28.103	38.740	0.807	0.859	0.704	0.645	0.727	0.485	0.669	0.720	0.506
MF = 30	**30.763**[b]	**26.219**[b]	36.165	**0.833**[b]	0.876	0.745	**0.691**[b]	**0.763**[b]	0.551	0.736	0.767	0.598
MF = 35	31.299	27.257	36.246	0.826	0.868	0.712	0.660	0.743	0.498	0.681	0.732	0.525
MF = 40	30.990	26.747	36.170	0.830	0.870	0.743	0.687	0.753	0.551	0.732	0.768	0.614
MF = 45	31.423	26.272	37.026	0.826	**0.877**[b]	0.731	0.678	0.762	0.530	0.709	0.755	0.571
MF = 50	30.983	27.120	**35.495**[b]	0.829	0.865	**0.754**[b]	0.687	0.746	**0.568**[b]	**0.737**[b]	**0.769**[b]	**0.629**[b]
MF = 55	30.977	26.447	36.116	0.830	0.874	0.746	0.687	0.758	0.552	0.723	0.764	0.597
MF = 60	31.051	27.086	36.654	0.829	0.866	0.752	0.685	0.747	0.554	0.730	0.767	0.607
						Reservoir DJK						
MF = 5	733.226	993.018	300.095	0.696	0.592	0.626	0.471	0.313	0.368	0.482	0.274	0.433
MF = 10	691.090	937.113	280.800	0.736	0.651	0.678	0.530	0.388	0.446	0.544	0.360	0.513
MF = 15	662.338	900.035	262.935	0.761	0.683	0.723	0.568	0.436	0.515	0.581	0.415	0.577
MF = 20	650.806	885.549	254.666	0.769	0.690	0.742	0.583	0.454	0.545	0.604	0.447	0.602
MF = 25	639.888	871.376	248.265	0.777	0.700	0.756	0.597	0.471	0.567	0.621	0.469	0.625
MF = 30	648.081	884.838	243.416	0.770	0.686	0.766	0.586	0.455	0.584	0.616	0.461	0.644
MF = 35	633.465	863.883	242.160	0.781	0.703	0.768	0.605	0.480	0.588	0.638	0.491	0.648
MF = 40	633.092	865.584	**233.295**[b]	0.780	0.699	**0.788**[b]	0.605	0.478	**0.618**[b]	**0.645**[b]	0.499	0.660
MF = 45	**621.274**[b]	**846.668**[b]	239.669	**0.790**[b]	**0.716**[b]	0.773	**0.620**[b]	**0.501**[b]	0.597	0.652	**0.514**[b]	0.661
MF = 50	633.100	862.669	244.364	0.780	0.699	0.762	0.605	0.482	0.581	0.651	0.511	**0.673**[b]
MF = 55	622.864	850.214	234.742	0.782	0.712	0.784	0.618	0.497	0.613	0.650	0.508	0.657
MF = 60	631.966	861.057	245.222	0.780	0.700	0.761	0.607	0.484	0.578	0.651	0.489	0.672

[a]WY stands for all the water years of the entire study period. Runoff Season includes April to September and FC stands for Flood Control Season which starts from October to next year March.

[b]The bold and underlined values indicate the best statistics for each column.

4.2. Variable Sensitivity Analysis

In order to investigate the sensitivity of decision variables, we employ a normalized Gini diversity index [*Breiman et al*., 1984] to measure the importance of model inputs in explaining the variations of model output. The normalized Gini diversity index is based on the standard Gini diversity index. The standard Gini diversity index [*Breiman et al*., 1984] is one of the most commonly and broadly used criterion in the standard decision tree algorithm [*Yang et al*., 2016; *Chandra et al*., 2010; *Guyon and Elisseeff*, 2003; *Hapfelmeier and Ulm*, 2014]. There are two steps to calculate the normalized Gini diversity index. First, the standard Gini diversity index for each decision variable is summed up for all splits of the "Parent Nodes" in the decision tree. Then, the summation of the Gini diversity index for each decision variable is further normalized to 0–1 in a manner that a normalized Gini diversity index that is equal to zero means the decision variable is not used in growing the regression tree. A normalized Gini diversity index equal to 1 indicates this decision variable is extensively used for all of the splits in the decision tree, which means all data points can be successfully and accurately classified with this decision variable. Another merit of using the normalized Gini diversity index is the contributions of each model inputs in explaining the variability target values are able to be quantified in the form of percentages. The pros and cons of other metrics, such as information gains, Twoing index, etc., were summarized in *Mingers* [1989] and *Buntine and Niblett* [1992] for interested readers. A detailed calculation of the Gini diversity index can be found in *Breiman et al*. [1984], *Yang et al*. [2016], and in support information of this paper. The normalized Gini diversity index is individually calculated for the inflow for the CLE and DJK reservoir as shown in Figure 10. The red and blue lines represent the normalized Gini index for the CLE and DJK reservoir, respectively. The subplots in Figures 10a–10l are corresponding to the results with different maximum feature shown in Figures 4 and 5.

In Figure 10, there are 21 total classes of model inputs. Each class indicates the accumulated importance percentage from particular types of inputs in explaining the variability of the model output (current inflow) shown in Figure 3. For example, class one indicates the percentage of a combined contribution from two raw model inputs, i.e., the 1 month-lag and 2 month-lag inflows, in explaining the current month inflow. Class two indicates the accumulated importance of three raw model inputs, i.e., the current, 1 month-lag, and 2 month-lag precipitation, in explaining the variability of current inflow. The accumulation of an

Table 4. Statistical Performances of SVR Using RBF, Linear, Polynomial, and Sigmoid Kernel Functions With Different Gamma (γ) and Penalty Coefficients (C), for the Trinity Lake (CLE) and the Danjiangkou Reservoir (DJK)

		RMSE-WY[a] (m^3/s)	RMSE-Runoff[a] (m^3/s)	RMSE-FC[a] (m^3/s)	CORR-WY	CORR-Runoff	CORR-FC	NSE-WY	NSE-Runoff	NSE-FC	KGE-WY	KGE-Runoff	KGE-FC
						Reservoir CLE							
RBF	$\gamma = 0.0001$ C = 100	53.047	48.505	58.654	0.701	0.789	0.576	0.082	0.187	−0.181	0.045	0.086	−0.049
	$\gamma = 0.01$ C = 100	43.101	39.766	48.041	0.739	0.800	0.622	0.394	0.454	0.208	0.329	0.322	0.231
	$\gamma = 0.01$ C = 1000	35.626	**32.876**[b]	40.157	**0.803**[b]	**0.844**[b]	**0.721**[b]	0.586	**0.627**[b]	0.447	0.538	**0.517**[b]	0.474
Linear	C = 1	42.610	40.067	46.810	0.753	0.823	0.657	0.408	0.445	0.248	0.338	0.296	0.265
	C = 10	36.238	34.380	**40.026**[b]	0.795	0.837	0.713	0.571	0.592	0.450	0.529	0.469	0.489
	C = 100	**34.960**[b]	33.353	38.573	0.798	0.835	0.713	**0.601**[b]	0.616	**0.489**[b]	**0.583**[b]	0.512	**0.542**[b]
Polynomial	Deg = 2 C = 100	60.423	52.561	68.291	0.099	0.541	0.011	−0.191	0.046	−0.600	−0.165	−0.062	−0.195
	Deg = 3 C = 100	181.044	53.663	250.982	−0.015	0.280	−0.043	−9.696	0.005	−20.611	−1.336	−0.218	−2.648
	Deg = 4 C = 100	53.793	50.551	58.180	0.577	0.684	0.514	0.056	0.117	−0.162	0.007	0.033	−0.046
Sigmoid	$\gamma = 0.0001$ C = 100	56.041	51.065	61.788	0.688	0.785	0.560	−0.025	0.099	−0.310	−0.036	0.028	−0.129
	$\gamma = 0.01$ C = 100	42.714	40.145	47.006	0.754	0.822	0.657	0.405	0.443	0.242	0.334	0.294	0.259
	$\gamma = 0.01$ C = 1000	37.383	35.538	41.009	0.770	0.807	0.688	0.544	0.564	0.423	0.520	0.468	0.472
						Reservoir DJK							
RBF	$\gamma = 0.0001$ C = 100	807.453	1094.749	327.105	0.700	0.665	0.516	0.358	0.165	0.249	0.304	0.136	0.346
	$\gamma = 0.01$ C = 100	668.081	893.546	311.010	0.788	0.745	0.601	0.561	0.444	0.321	0.518	0.399	0.522
	$\gamma = 0.01$ C = 1000	**634.797**[b]	**845.210**[b]	304.587	**0.799**[b]	**0.752**[b]	0.624	**0.603**[b]	**0.503**[b]	0.349	**0.583**[b]	**0.477**[b]	**0.571**[b]
Linear	C = 1	1031.122	1333.177	588.912	0.171	0.324	0.031	−0.047	−0.238	−1.435	−0.093	−0.206	−0.012
	C = 10	974.236	1297.568	461.486	0.370	0.419	0.171	0.065	−0.172	−0.495	0.014	−0.130	0.148
	C = 100	3017.498	1392.018	4040.286	−0.021	0.149	−0.032	−7.966	−0.349	−9.607	−1.072	−0.363	−8.649
Polynomial	Deg = 2 C = 100	929.421	1266.451	350.740	0.599	0.509	0.408	0.149	−0.117	0.136	0.087	−0.086	0.087
	Deg = 3 C = 100	820.601	1115.855	319.964	0.682	0.604	0.542	0.337	0.133	0.281	0.287	0.110	0.323
	Deg = 4 C = 100	687.896	931.367	**283.184**[b]	0.778	0.716	**0.669**[b]	0.534	0.396	**0.437**[b]	0.485	0.362	0.545
Sigmoid	$\gamma = 0.0001$ C = 100	969.903	1320.259	369.115	0.563	0.432	0.327	0.074	−0.214	0.043	0.010	−0.161	−0.031
	$\gamma = 0.01$ C = 100	810.738	1099.237	328.810	0.696	0.659	0.508	0.353	0.159	0.241	0.299	0.131	0.344
	$\gamma = 0.01$ C = 1000	697.678	905.421	383.432	0.739	0.709	0.493	0.521	0.429	−0.032	0.528	0.413	0.493

[a]WY stands for all the water years of the entire study period. Runoff Season includes April to September and FC stands for Flood Control Season which starts from October to next year March.
[b]The bold and underlined values indicate the best statistics for each column.

identical type of input is only for better visualization of results. The breakdown of each input class importance percentage is shown in Figures 11 and 12 for CLE and DJK reservoir, respectively. Figures 11 and 12 depict the normalized Gini diversity index for 62 raw model inputs with lag information.

Because the decision variables for the studied AI & DM models (ANN, RF, and SVR) include hydrological and climatic information, the current and previous time steps variables, and seasonality, the following discussion section is organized specifically to investigate (1) the parameterization of different models, (2) auto and cross correlations of variables, and (3) the physics of climatic phenomenon and their implications on reservoir inflow forecasts.

5. Discussion

5.1. Parameterization of Models

In this study, we investigate the maximum feature parameter used in the RF model. The sensitivity analysis on other structural parameters, i.e., the tree depth can be found in *Yang et al.* [2016]. According to Table 3,

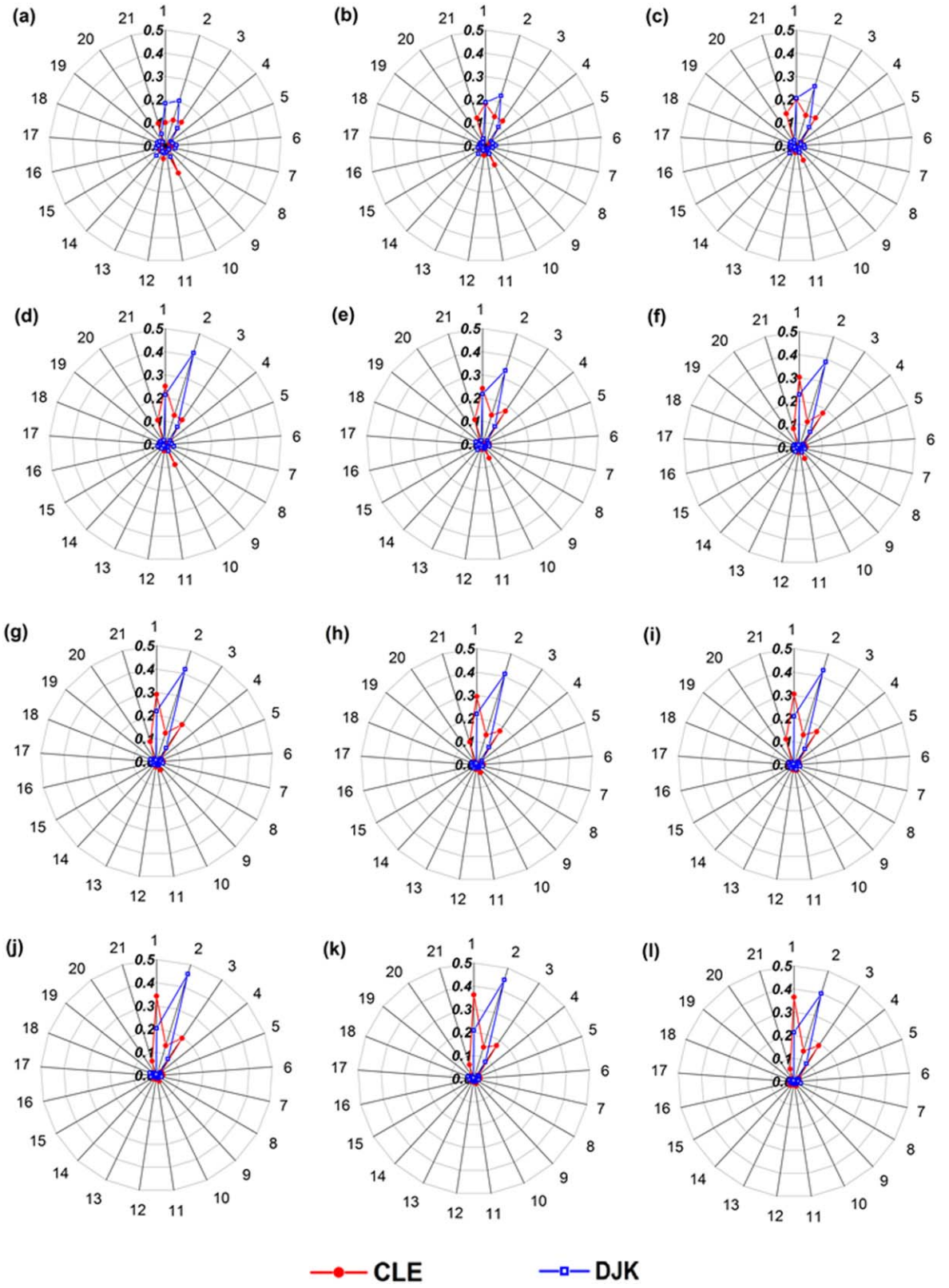

Figure 10. The normalized Gini diversity index for 21 classes of inputs with different maximum features in RF for the CLE and DJK reservoir.

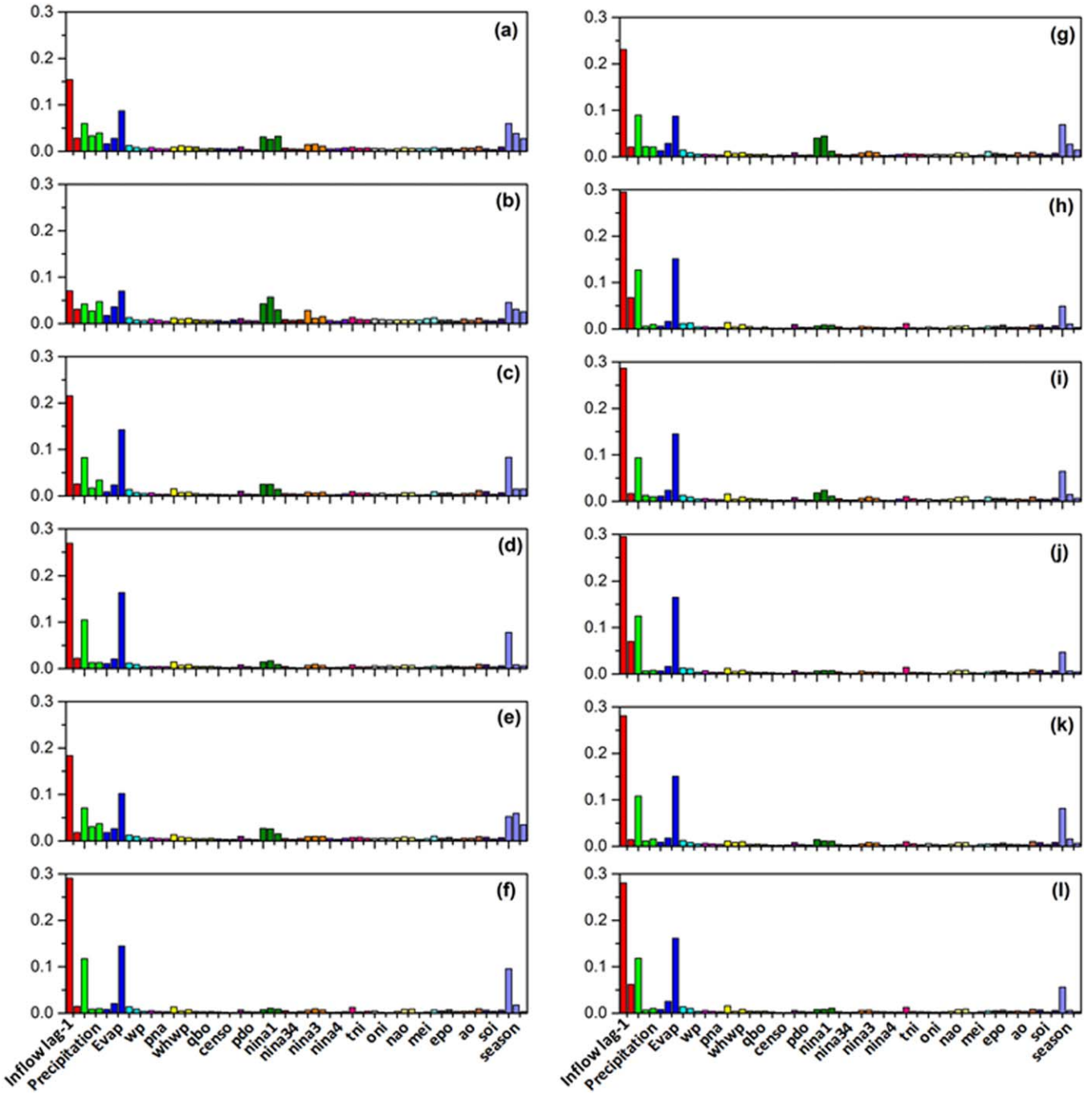

Figure 11. The breakdown of the Normalized Gini diversity for 62 raw inputs for the CLE reservoir.

there is an interesting pattern about the maximum feature, in which the statistics tend to improve as the maximum feature increases from 5 to 30, i.e., equals approximately half of the number of decision variables. After the maximum feature reaches 30, the performances of RF stays stable and no significant improvement is observed when increasing the parameter values. After the maximum feature goes beyond 50, the statistical performances of RF even becomes worse. The models with best performance are associated with values of the maximum feature larger than 50% and smaller than 85% of the total number of decision variables. The existence of an optimal range of the maximum feature parameter can be explained with respect to the RF mechanism. In the tree-growing process of RF, a single candidate tree grows using only a random selection of decision variables instead of the full numbers of decision variables [*Liaw and Wiener*, 2002]. The

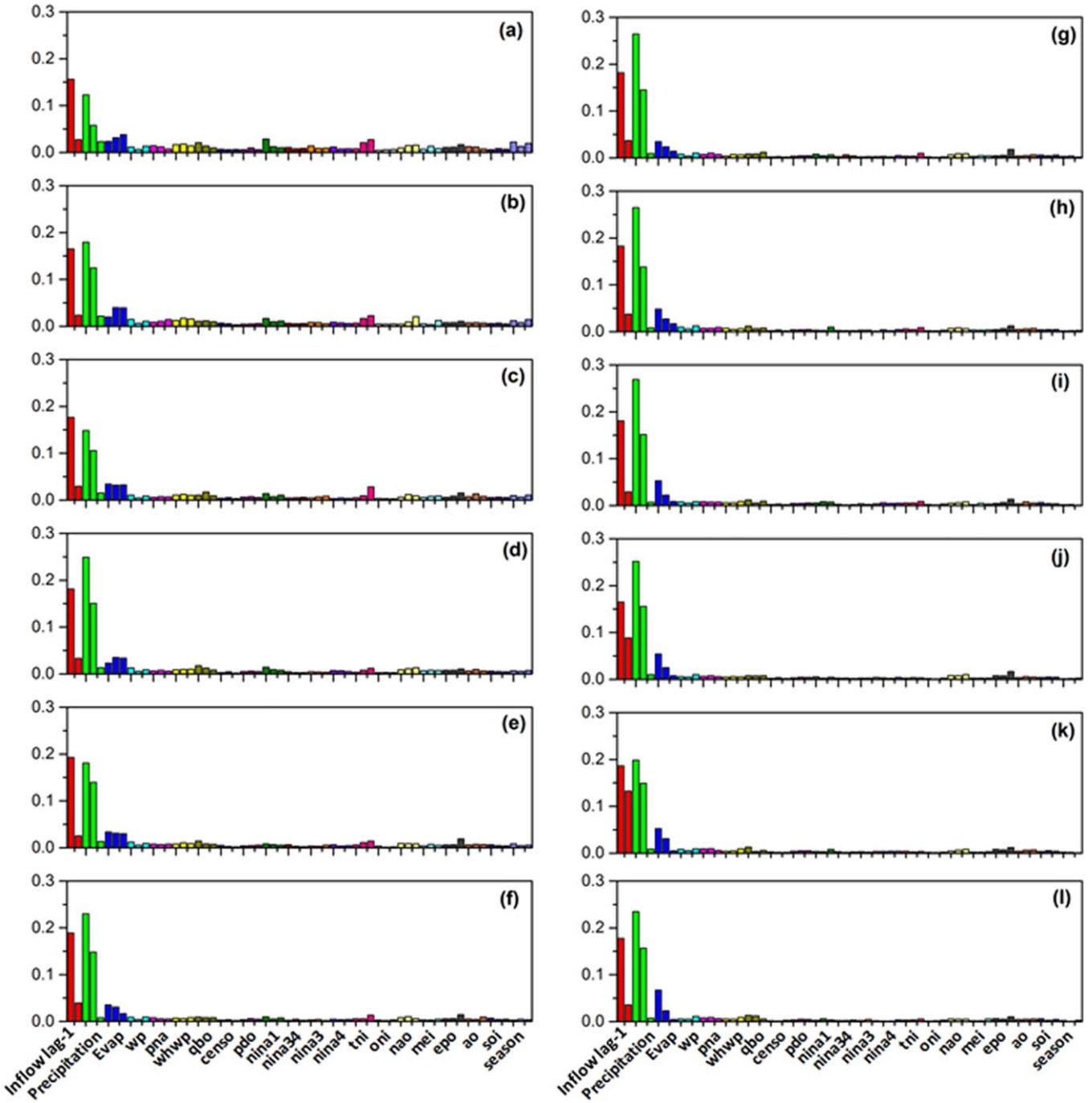

Figure 12. The breakdown of the Normalized Gini diversity for 62 raw inputs for the DJK reservoir.

number variables selected in growing candidate trees is termed the maximum feature. When the maximum feature value is relatively low, such as less than 20% of the total number of decision variables, the random selection process is not able to guarantee a sufficient number of candidate trees can be built using decision variables with possibly high predictive capability. For example, under an extreme scenario, if the maximum feature is set to 1, each candidate tree will grow using only one randomly selected decision variable, which will result in important hydrological and climatic information not being selected in the final candidates group, which will contain 100 single variable tree members. On another extreme scenario, if the maximum feature value is too large, i.e., larger than 85% of the total number of decision variables, the variability

among candidate trees will be relatively small. After all, the advantage of ensemble trees is to purposely create sufficient numbers of both "good" learners and "bad" learners [*Breiman*, 1996, 2001; *Liaw and Wiener*, 2002] in the candidate pool. Therefore, in the reservoir inflow forecast problem shown in this paper, it is found that 50–85% of decision variables used as predictors for RF is a suitable range. As the RF input structure changes (input data), the acceptable range of the maximum feature could change correspondingly, nevertheless, the existence of an optimal range of such parameter applies to other practical uses of the RF algorithm.

According to the statistics shown in Table 2 for ANN, it is found (1) in general increases of the maximum iterations and the number of hidden nodes are able to improve the statistics; and (2) the increases of hidden nodes are capable of further increasing model performances after the iterations are relaxed. For example, the best ANN statistics for CLE and DJK reservoir (Table 2) are consistently associated with the scenario of 15 or 20 hidden nodes for a maximum number of iteration equals to 1000 or 2000. In the use of ANN in our study, a backpropagation and a gradient-based optimization scheme are employed. As the number of maximum iterations increases, the gradient optimization scheme continuously explores the response surface of the objective function (equation (5)), and the evolution is not terminated until the algorithm reaching the allowable maximum iterations number. However, as the search evolves with more iterations, the gradient of objective function will decrease, resulting in any further increase in the maximum iteration number having less improvement on the objective function value, as it is shown in the 1000 and 2000 maximum iterations scenarios in Table 2. The number of hidden nodes, as an ANN structural parameter, shows positive effects on improving model performances, and the tested ANN models with the best performance are all with relatively high hidden nodes (H = 15 or 20). As mentioned in *Yao* [1999], the size of the ANN, such as the number of hidden nodes, is crucial for model performance, and should be jointly designed and optimized with a proper training algorithm. In our applied ANN, the increase of hidden nodes will create more computational challenges for the backpropagation and the gradient-based optimization scheme. This is also the reason the best performing models in the scenarios tested are all with a high maximum iteration (MI = 1000 or 2000) and a high number of hidden nodes (H = 15 or 20). As demonstrated in many other studies [*Ding et al*., 2013; *Ilonen et al*., 2003; *Khan and Sahai*, 2012; *Slowik and Bialko*, 2008], the use of other optimization schemes, such as the evolutionary algorithm or goal programming, could be suitable approaches to improve ANN model performances, however, such work falls into the category of optimal design of ANN and optimization, which should be addressed in another study other than this paper.

With respect to the performances of SVR (Figures 8 and 9 and Table 4), the variability of the model performance varies dramatically from one to another. Using some kernel functions and parameter settings, the SVR models are able to produce satisfactory results, such as RBF, linear, and Sigmoid with high penalty and low gamma cases in CLE reservoir (Table 4, and Figures 8c, 8f, and 8i). Except these three cases, the NSE and KGE values for other scenarios of CLE reservoir are all below 0.50, which is not acceptable for monthly streamflow simulation [*Gupta et al*., 2009; *Moriasi et al*., 2007]. In the DJK reservoir, only RBF kernel function with high penalty and low gamma, and the high-order polynomial kernel function could generate satisfactory statistics for the reservoir inflow simulation (Table 4, and Figures 9c and 9i). Therefore, we suggest the use of SVR for reservoir inflow forecasts should be carefully examined before practical uses. According to our experiments shown in this study, the SVR method can provide satisfactory statistics with proper kernel functions and parameters. In the study presented by *Asefa et al*. [2006], the authors also concluded the selection of kernel functions and model coefficients could result in large uncertainties in predictions. The penalty coefficient terms appear very sensitive in our study, which is in agreement with the results *Lin et al*. [2006] obtained in their discharge prediction case study. Another interesting thing shown in the SVR experiment is the predictions with polynomial kernel functions are consistently worse than those with linear kernels for the CLE reservoir inflow simulation, and as the degree of polynomial increases the statistics become worse. This is seen because the polynomial kernel functions used in our study are not allowed to have zero coefficients on each term as shown in equation (12). As the degree of the polynomial function increases, the roughness and nonlinearity of the hyperplane becomes higher. On the contrary, the high-order of polynomial function could produce relatively better results than linear kernels when simulating the reservoir inflows for the DJK reservoir (Table 4). It is highly possible the raw inputs for the CLE reservoir could be classified fairly well with a linearized hyperplane instead of a hyperplane with higher dimensions that created by the RBF, Sigmoid, and Polynomial kernel functions. However, the model inputs for the DJK reservoir are

more complexed and linearized hyperplane is not able to represent the relationship between model inputs and outputs well. Investigations on the complexity of the response surface and data structure using SVR are out of the scope of this paper, but are encouraged for an investigation in another study. As suggested by *Lin et al*. [2006], perhaps automatic calibration on SVR model parameters could be beneficial in determining the proper kernel function and coefficients, both of which in our study are found to be the keys to achieve acceptable predictions when using SVR models.

5.2. Cross and Autocorrelations of Variables

Large-scale climatic phenomenon and local hydrology are closely related and the former always influence on latter. Furthermore, streamflow, precipitation, evaporation, and climate phenomenon indices have autocorrelation effects due to the intrusion of lagging information into model inputs. In the following Figure 13, the cross-correlation coefficients among 62 raw model inputs, and 21 model inputs classes for both CLE and DJK reservoirs are shown using the data from 1964 to 2010. In details, Figures 13a and 13b present the cross correlation among model inputs of the CLE reservoir, and Figures 13c and 13d indicate the calculated coefficients among the model inputs of the DJK reservoir. Note that in Figures 13a and 13c, the 62 ticks on x and y axis correspond to the 62 raw model inputs shown in Figure 3. In Figures 13b and 13d, the first x and y ticks indicate the 1 month-lag inflow as the first model input, second and third x and y ticks indicate the current precipitation and evaporation without lagging information, and etc. The 45° lines in Figure 13 suggest the self-correlation coefficients among decision variables, which always equal to 1.

According to the correlation coefficients shown in Figure 13a, most climate phenomenon indices have strong autocorrelation within one time step, and the autocorrelation effects become less significant when lagging the variable for two time steps, though a number of climate indices are found to have strong autocorrelation effects up to 2 months, including CENSO, PDO, TNI, WHWP, ONI, MEI, and Nino indices. Given the facts most of these autocorrelated climate indices, such as PDO, are calculated from the Rotated Principal Component Analysis (RPCA) or Empirical Orthogonal Function (EOF) analysis, the interdecadal or interanual variabilities always dominate the subseasonal variation as the leading principal components. In addition, the interdecadal variability of PDO is obvious and well documented in many studies [*Hamlet and Lettenmaier*, 1999; *Stenseth et al*., 2003; *Redmond and Koch*, 1991; *Mantua et al*., 1997; *Hamlet et al*., 2002]. Therefore, it is not surprising the indices, such as PDO, have strong autocorrelation effects up to 2 months. Similarly, the autocorrelation effects in enso-related climate indices, i.e., CENSO, TNI, ONI, and MEI, are also significant. The CENSO index represents the bivariate ENSO timeseries calculated from both standardized SOI and Nino3.4 surface temperature timeseries [*Smith and Sardeshmukh*, 2000]. The TNI index is calculated using standardized Nino1 + 2 and 4, and the calculation of the NOI index is slightly different from the calculation for SOI index [*Schwing et al*., 2002]. The different ENSO phases typically last for several months to 1 or a few years [*Holton et al*., 1989]. Therefore, the correlation coefficients among current, 1 month-lag and, 2 month-lag CENSO, TNI, ONI, and MEI indices are very high due to the facts that ENSO phases have not shifted. Last, according to the study by *Wang and Enfield* [2001], the WHWP, as the second largest tropical warm pool on Earth [*Weisberg*, 1996], has a dominating annual cycle and significant interannual departures with regard to location and intensity. This is the reason that the autocorrelation of WHWP within 2 months is also strong.

As shown in Figure 13, the direct correlation between hydrological information, i.e., precipitation and evaporation, and climate phenomenon indices are not very strong. Nevertheless, the weak correlation does not mean there is no interaction between large-scale climatic conditions and local hydrology. In California, winter precipitations, especially for Atmospheric River (AR) events are highly correlated [*Ralph and Dettinger*, 2012; *Leung and Qian*, 2009; *Lavers et al*., 2016; *Dettinger et al*., 2011; *Dettinger*, 2011]. In general, during winter or reservoir flood control seasons, a massive amount of water vapor is transferred in a short time from the tropical regions of the Pacific Ocean to the western U.S., and eventually results in precipitation landfalls. The driving forces for such quick movement of water vapor in the atmosphere are due to the interactions of sea surface temperature, air pressure, wind, and many other climatic forcings. Another finding in Figures 13b and 13d is that after removing lagging information, strong cross correlations among many climate phenomenon indices are observed. This result is seen because those climate phenomenon indices essentially can be grouped into similar categories based on their physical descriptions of (1) teleconnections, (2) atmosphere, (3) precipitation, (4) ENSO, (5) Pacific Ocean sea surface temperature, and (6) Atlantic Ocean sea surface temperature, only with different calculation and standardization philosophies.

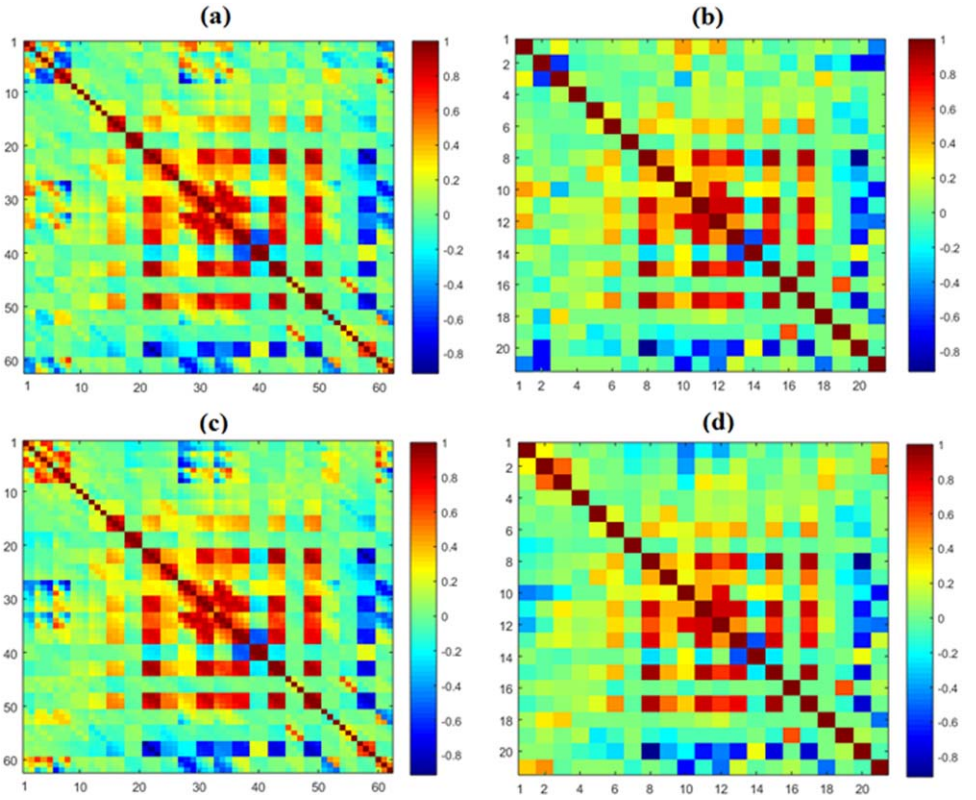

Figure 13. Cross-correlation coefficients among (a) 62 raw model inputs of the CLE reservoir, (b) reduced model inputs including lag-1 month inflow and other variables without lagging information of the CLE reservoir, (c) 62 raw model inputs of the DJK reservoir, and (d) reduced model inputs including lag-1 month inflow and other variables without lagging information of the DJK reservoir.

One minor limitation of this study is the investigation on the optimal latency of climate phenomenon indices, of which perhaps a longer period of lagging information may provide higher predictability to streamflow. The current time lag for model inputs is set up to 2 months. As shown in Figures 13a and 13c, some climate indices have high autocorrelation effects up to 2 months, which could be purposely lagged a few more steps as additional model inputs. However, the investigation is beyond the scope of this paper and deserves a separate comprehensive study. After all, better predictions will be achievable if the model inputs contain sufficient predictability and representation of all useful information that not included in our current study. Furthermore, the nonstationarity of climate phenomenon and streamflow are not considered in our model settings. As suggested by *Turner and Galelli* [2016], the regime state variable shift could be a major cause of suboptimal reservoir operation. Future studies, when using the historical climate phenomenon indices and streamflow observation at decadal scale to support reservoir operation, are suggested to take the nonstationarity of regime state variable into consideration.

5.3. Signals of Climate Phenomenon Indices

Figure 10 shows the accumulated contributions of model inputs in explaining the predicted inflows. According to Figure 10, the contributions from 1 month-lag and 2 month-lag inflow, current and past precipitation, and current and past lake evaporation are able to explain approximately 20–30%, 10–15%, and 15–20% of the total variation of monthly inflow, respectively. By a robust estimate, the hydrological information (past inflows, and both current and past precipitation and evaporation) together contributes about 45–65% of the total predictability to current inflows. About 35–55% input predictability comes from the climatic conditions. According to Figure 10, precipitation has larger impacts on reservoir inflow for the DJK reservoir than the CLE reservoir. This is because the high correlation between the precipitation in DJK area and the seasonality (Figures 13b and 13d). The impacts of seasonality on CLE reservoir are much higher than that on the DJK reservoir (Figure 10). It is possible that RF model identifies that the precipitation in the DJK area is highly correlated with seasonality. Therefore, precipitation inputs are

used more frequently in the tree-growing process of DJK inflow regression than that in the simulating the inflow to CLE reservoir.

As discussed in the previous subsection, the climate phenomenon indices can be strongly correlated to each other due to their similarity of physical meanings. The contribution percentages are distributed among all kinds of climate indices as shown in the bar plots of Figures 11 and 12. The climate index with the highest contribution is Nino 1 for both CLE and DJK reservoir, followed by Nino 3, Nino 4, TNI, QBO, WP, and WHWP. This is a result of the sea-surface temperature variability in the tropical regions of the Pacific Ocean being the key driving force for fast horizontal water vapor transport, which bring large and heavy precipitation via AR events and directly becomes streamflow or snowpack, especially for the CLE reservoir. Nevertheless, the authors also suspect the contribution of Nino indices on streamflow may also come from other indices, such as PDO, CENSO, ONI, MEI, and other ENSO-related indices. As shown in Figures 13b and 13d, the Nino indices are highly related to many other climate phenomenon indices with high correlation coefficients, such as PDO, MEI, and ONI. It is possible the model overestimates the contribution of Nino indices, whereas underestimates the contributions from other climate indices which are highly correlated to Nino indices. *Gutiérrez and Dracup* [2001] found there were significant correlations between climate indicators, including MEI, SOI, and Nino indices, and reservoir discharges in the Columbia River Basin. Based on the correlation analysis, *Gutiérrez and Dracup* [2001] concluded efficient reservoir operation policies can be established according to the changes in these climate indicators. In another study conducted in the Yellow River Basin in China, which also belongs to the same type of semiarid region as California, *Lü et al*. [2011] identified that Nino indices, and SOI have strong correlations with the river discharge in the Yellow River basin. A similar conclusion was drawn by *Kalra et al*. [2013], who demonstrated that the use of NAO is able to increase the streamflow forecast length in the upper Colorado River Basin.

In our case study, the predictive capability of climate indices on streamflow is also identified, as they explain approximately 35–55% of the inflow variability. Even though multiple satisfactory reservoir inflow forecasts are achieved in our study cases, some questions still remain with regards to (1) what are the relationships among climate phenomenon and (2) how the large-scale ocean-climate variability influences local hydrology. Future studies are suggested to investigate the interactions of certain correlated climate phenomenon and their impacts on local weather and water supplies. Based on the experiments and findings in this study, the authors reach a similar conclusion as *Burley et al*. [2012], who identified that climate phenomenon indices as useful information sources for supporting reservoir decision-making, and the shift of many climate phenomenon indices should be emphasized in future studies on streamflow forecasts and afterwards reservoir reoperation. Similar suggestions of using climate phenomenon indices in supporting water resources management and planning have also been made by *Pagano and Garen* [2003] and *Garen* [1993]. *Garen* [1993] pointed out many climate phenomenon indices can be used as a useful predictor in estimating the water-supply conditions in many parts of the western United State. *Pagano and Garen* [2003] summarized that although it was still a slow migration process for the actual use of climate indicators in water-supply forecasts, the predictability of climate information was already acknowledged by many agencies, especially in the western regions of United States.

The AI & DM methods are powerful and flexible tools for incorporating climate indices into regression modes. One advantage of RF over ANN and SVR is the transparency of regression. In ANN, the neurons in hidden layers represent a combination of the weighted sum of input information, which does not explain how one output value is obtained. In addition, the Hyperbolic Tangent function normalizes the input values to a bounded value between -1 and 1, solely for faster convergence. The information and integrity of the magnitudes of input values are compromised. *Li et al*. [2010] reviewed many pros and cons of several AI and DM methods, including ANN and DT, and concluded neural networks cannot illustrate how prediction is achieved, while DT methods have the advantage of transparency as we demonstrated with respect to climate indices contribution to streamflow prediction. However, *Li et al*. [2010] also criticized DT methods because they display a weakness in treating nonaxis-parallel class boundaries of training data. A joint use of AI & DM methods may better address this weakness of DT methods, and provide auxiliary information for some operational streamflow regression approaches, such as multivariable and analogue regression.

In SVR, as demonstrated in Figures 8 and 9, and Table 4, the model performances depend on the selection of kernel functions and proper parameters. The parameter calibration in SVR and the selection of multiple kernel functions will be a tedious task. The transparency and flexibility of single decision tree and ensemble

tree methods, such as RF algorithm, are preferred by many researchers, who have successfully applied these methods to many water resources management and planning problems [*Galelli and Castelletti*, 2013; *Bessler et al*., 2003]. In addition, according to *Hobbs et al*. [1997], the recommendation is also made to use DT methods to evaluate the signals and impacts of climate variability on water resources supply and management decisions. In summary, our study presents a novel way of using ANN, RF, and SVR to incorporate climate information into reservoir inflow prediction and exhibits a systematic comparison with regard to these three popular AI & DM methods for a practical use on the CLE and DJK reservoir.

6. Conclusions

In this study, three of the most popular AI & DM methods, namely ANN, RF, and SVR, are employed and compared in predicting monthly reservoir inflow of CLE reservoir in USA and the DJK Reservoir in China using about 50 years of historical reservoir-operation records, multiple climate phenomenon indices, as well as the lagged information. Results show that RF is able to provide reasonably good predictions, and ANN and SVR can only produce satisfactory prediction in some cases using proper structural settings. The use of RF method is recommended for reservoir inflow prediction, not only because of its robust prediction capability over ANN and SVR, but also because of its transparency in model structures and that allows dam operators to monitor and customize different inputs based on their regional hydrology and favored climate phenomenon indices. The authors believe this unique feature of the AI & DM methods, especially the RF method, is able to further the help decision makers understand the implications of climate phenomenon indices on streamflow, and provide more flexibility of adding desired decision variables by operators. According to the experiments conducted in this study, the following conclusions can be drawn with regard to the methodologies and applications to CLE reservoir in United States and the DJK reservoir in China.

1. The AI & DM methods are robust and useful tools for simulating monthly reservoir inflows, thus, assisting reservoir operations. The statistics obtained by the best performing models are satisfactory as shown in the experiments on the CLE and DJK reservoirs in United States and China, respectively. In details, for the CLE reservoir, the highest NSE values are 0.613, 0.691, and 0.627, respectively, and the highest KGE values for ANN, RF, and SVR are 0.732, 0.737, and 0.583, respectively. For the DJK reservoir, the highest NSE values are 0.440, 0.620, and 0.603, for ANN, RF, and SVR respectively, and the highest KGE values derived by ANN, RF, and SVR are 0.578, 0.645, and 0.583, respectively.
2. Different parameterizations were compared with regard to ANN, RF, and SVR to enhance the practical uses of these three methods. The Random Forest (RF) methods turn out to be better models than the ANN and SVR methods. The reservoir inflows generated by RF models have consistently better and stable values as compared to any other methods. Furthermore, an optimal range of maximum feature is suggested when using the RF algorithm. In ANN, the increase of maximum iteration and number of hidden nodes will improve model performances. However, with a larger number of hidden nodes, the maximum iteration is suggested to increase correspondingly in order to produce accurate predictions. Last, the prediction from SVR has a large variability so that the proper selection of kernel functions and parameterization are required before practical uses.
3. The general model performances of ANN, RF, and SVR are better during runoff seasons (April to September) than the flood control season (October to next year March) for the CLE reservoir. The reason is that in California reservoir inflow during runoff season is largely contributed by snowmelt instead of heavy precipitation during the winter time.
4. The variability of reservoir inflow is influenced by multiple climate phenomenon indices, including Nino 1 for both CLE and DJK reservoirs, followed by Nino 3, Nino 4, TNI, QBO, WP, and WHWP. In general, the local hydrological information can explain about 45–65% of the total variability of current reservoir inflow, while the rest 35–55% are due to climate conditions for the two reservoirs selected in this paper.
5. Most AI & DM models cannot successfully predict the peak flows. The reason is the employed models are not particularly trained on high flows. Therefore, it is not surprising that AI & DM methods may not outperform hydrological models on high flows. However, AI & DM methods have the flexibility to incorporate auxiliary information and soft data, i.e., expert knowledge, into the modeling framework, which could assist decision makers or dam operators in managing their facilities.

Based on this study, which provides a novel way of incorporating climate information into streamflow regression, and investigates the physical interactions between large-scale climatic phenomenon and local hydrological processes, future studies can be carried out. For example, the Atmospheric Rivers (AR) events, which quickly bring large amounts of precipitable water to the western United State during the winter season, could be set as prediction variables using AI & DM methods. In addition, the uses of climatic information are suggested in order to consider nonstationarity and the shift of regime variables in streamflow. With regard to the methodology, the ANN training schemes, the backpropagation scheme and gradient-based optimization scheme used in ANN cannot guarantee the training weights are a global optimum for all types of problems, and the searching is sometimes associated with risks of being trapped in a local optimum. However, the methodologies presented in our study are universally adaptable for other case studies, and are flexible for a customized design for any particular water resources planning problem.

Acknowledgments
The financial supports for this study are from the DOE (Prime Award # DE-IA0000018), CEC (Award # 300-15-005), the CDWR Seasonal Forecasting via Database Enhancement Program (DWR agreement 4600010378), the NSF CyberSEES project (Award CCF-1331915), the NOAA/NESDIS/NCDC (Prime award NA09NES4400006 and NCSU CICS and subaward 2009-1380-01), and the Army Research Office (award W911NF-11-1-0422). We also thank many anonymous reviewers for their suggestive comments. The data for this paper are available at the California Data Exchange Center (http://cdec.water.ca.gov/index.html) and the NOAA Earth System Research Laboratory (http://www.esrl.noaa.gov/psd/data/climateindices/list/).

References:

Araújo, M. B., and M. New (2007), Ensemble forecasting of species distributions, *Trends Ecol. Evol.*, *22*(1), 42–47.
Asefa, T., M. Kemblowski, M. McKee, and A. Khalil (2006), Multi-time scale stream flow predictions: The support vector machines approach, *J. Hydrol.*, *318*(1), 7–16.
Ashaary, N. A., W. H. W. Ishak, and K. R. Ku-Mahamud (2015), Forecasting model for the change of reservoir water level stage based on temporal pattern of reservoir water level, in *Proceedings of the 5th International Conference on Computing & Informatics*, pp. 692–697, ICOCI, Istanbul, Turkey.
Bessler, F. T., D. A. Savic, and G. A. Walters (2003), Water reservoir control with data mining, *J. Water Resour. Plann. Manage.*, *129*(1), 26–34.
Breiman, L. (1996), Bagging predictors, *Mach. Learn.*, *24*(2), 123–140.
Breiman, L. (2001), Random forests, *Mach. Learn.*, *45*(1), 5–32.
Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984), *Classification and Regression Trees*, CRC Press, Boca Raton, Fla.
Buntine, W., and T. Niblett (1992), A further comparison of splitting rules for decision-tree induction, *Mach. Learn.*, *8*(1), 75–85.
Burley, N., K. Maher, and J. Lund (2012), *Re-Operation of Major Reservoirs for Flood and Environmental Management*, pp. 2388–2394, Am. Soc. of Civ. Eng., Albuquerque, N. M.
CDWR (2014a), *System Reoperation Study Forecast-Based Operations Analysis System Reoperation Report: Phase 2 Report Attachment A*, Sacramento, Calif.
CDWR (2014b), *System Reoperation Study Strategy Formulation and Refinement. System Reoperation Report: Phase 2 Report*, Sacramento, Calif.
Chandra, B., R. Kothari, and P. Paul (2010), A new node splitting measure for decision tree construction, *Pattern Recognit.*, *43*(8), 2725–2731.
Chebrolu, S., A. Abraham, and J. P. Thomas (2005) Feature deduction and ensemble design of intrusion detection systems, *Comput. Secur.*, *24*(4), 295–307.
Cheng, C. C., N. S. Hsu, and C. C. Wei (2008), Decision-tree analysis on optimal release of reservoir storage under typhoon warnings, *Nat. Hazards*, *44*(1), 65–84.
Cheng, C. T., Z. K. Feng, W. J. Niu, and S. L. Liao (2015), Heuristic methods for reservoir monthly inflow forecasting: A case study of Xinfengjiang Reservoir in Pearl River, China, *Water*, *7*(8), 4477–4495.
Cichocki, A., and S.-I. Amari (2002), *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley, Chichester, West Sussex, England.
Coulibaly, P., F. Anctil, and B. Bobee (2001), Multivariate reservoir inflow forecasting using temporal neural networks, *J. Hydrol. Eng.*, *6*(5), 367–376.
Dettinger, M. (2011), Climate change, atmospheric rivers, and floods in California: A multimodel analysis of storm frequency and magnitude changes, *J. Am. Water Resour. Assoc.*, *47*(3), 514–523.
Dettinger, M. D., F. M. Ralph, T. Das, P. J. Neiman, and D. R. Cayan (2011), Atmospheric rivers, floods and the water resources of California, *Water*, *3*(2), 445–478.
Ding, S., H. Li, C. Su, J. Yu, and F. Jin (2013), Evolutionary artificial neural networks: A review, *Artif. Intell. Rev.*, *39*(3), 251–260.
Egmont-Petersen, M., D. de Ridder, and H. Handels (2002), Image processing with neural networks: A review, *Pattern Recognit.*, *35*(10), 2279–2301.
Elith, J., and J. R. Leathwick (2009), Species distribution models: Ecological explanation and prediction across space and time, *Annu. Rev. Ecol. Evol. Syst.*, *40*(1), 677–697.
Erdal, H. I., and O. Karakurt (2013), Advancing monthly streamflow prediction accuracy of CART models using ensemble learning paradigms, *J. Hydrol.*, *477*, 119–128.
Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth (1996), From data mining to knowledge discovery in databases, *AI Magaz.*, *17*(3), 37.
Galelli, S., and A. Castelletti (2013), Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling, *Hydrol. Earth Syst. Sci.*, *17*(7), 2669–2684.
Garen, D. C. (1993), Revised surface-water supply index for western United States, *J. Water Resour. Plann. Manage.*, *119*(4), 437–454.
Goyal, M. K., C. Ojha, R. Singh, and P. Swamee (2013a), Application of artificial neural network, fuzzy logic and decision tree algorithms for modelling of streamflow at Kasol in India, *Water Sci. Technol.*, *68*(12), 2521–2526.
Goyal, M. K., C. Ojha, R. Singh, P. Swamee, and R. Nema (2013b), Application of ANN, fuzzy logic and decision tree algorithms for the development of reservoir operating rules, *Water Resour. Manage.*, *27*(3), 911–925.
Guo, J., J. Zhou, H. Qin, Q. Zou, and Q. Li (2011), Monthly streamflow forecasting based on improved support vector machine model, *Exp. Syst. Appl.*, *38*(10), 13,073–13,081.
Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, *377*(1), 80–91.
Gutiérrez, F., and J. A. Dracup (2001), An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño–Southern Oscillation indicators, *J. Hydrol.*, *246*(1–4), 181–196.
Guyon, I., and A. Elisseeff (2003), An introduction to variable and feature selection, *J. Mach. Learn. Res.*, *3*, 1157–1182.

Hamlet, A. F., and D. P. Lettenmaier (1999), Columbia River streamflow forecasting based on ENSO and PDO climate signals, *J. Water Resour. Plann. Manage.*, *125*(6), 333–341.
Hamlet, A. F., D. Huppert, and D. P. Lettenmaier (2002), Economic value of long-lead streamflow forecasts for Columbia River hydropower, *J. Water Resour. Plann. Manage.*, *128*(2), 91–101.
Hancock, T., R. Put, D. Coomans, Y. Vander Heyden, and Y. Everingham (2005), A performance comparison of modern statistical techniques for molecular descriptor selection and retention prediction in chromatographic QSRR studies, *Chemometr. Intell. Lab. Syst.*, *76*(2), 185–196.
Hapfelmeier, A., and K. Ulm (2014), Variable selection by Random Forests using data with missing values, *Comput. Stat. Data Anal.*, *80*, 129–139.
Hejazi, M. I., and X. M. Cai (2009), Input variable selection for water resources systems using a modified minimum redundancy maximum relevance (mMRMR) algorithm, *Adv. Water Resour.*, *32*(4), 582–593, doi:10.1016/j.advwatres.2009.01.009.
Hejazi, M. I., and X. M. Cai (2011), Building more realistic reservoir optimization models using data mining: A case study of Shelbyville Reservoir, *Adv. Water Resour.*, *34*(6), 701–717.
Hobbs, B. F., P. T. Chao, and B. N. Venkatesh (1997), Using decision analysis to include climate change in water resources decision making, *Clim. Change*, *37*(1), 177–202.
Holton, J. R., R. Dmowska, and S. G. Philander (1989), *El Niño, La Niña, and the Southern Oscillation*, Academic press, San Diego, Calif.
Hopfield, J. J. (1988), Artificial neural networks, *IEEE Trans. Circuits Dev. Magaz.*, *4*(5), 3–10.
Hsu, K.-l., X. Gao, S. Sorooshian, and H. V. Gupta (1997), Precipitation estimation from remotely sensed information using artificial neural networks, *J. Appl. Meteorol.*, *36*(9), 1176–1190.
Ilonen, J., J.-K. Kamarainen, and J. Lampinen (2003), Differential evolution training algorithm for feed-forward neural networks, *Neural Process. Lett.*, *17*(1), 93–105.
Jain, A. K., J. Mao, and K. Mohiuddin (1996), Artificial neural networks: A tutorial, *Computer*, *29*(3), 31–44.
Jain, A. K., R. P. Duin, and J. Mao (2000), Statistical pattern recognition: A review, *IEEE Trans. Pattern Anal. Mach. Intell.*, *22*(1), 4–37.
Jain, S., A. Das, and D. Srivastava (1999), Application of ANN for reservoir inflow prediction and operation, *J. Water Resour. Plann. Manage.*, *125*(5), 263–271.
Kalra, A., W. P. Miller, K. W. Lamb, S. Ahmad, and T. Piechota (2013), Using large-scale climatic patterns for improving long lead time streamflow forecasts for Gunnison and San Juan River Basins, *Hydrol. Processes*, *27*(11), 1543–1559.
Khan, K., and A. Sahai (2012), A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context, *Int. J. Intell. Syst. Appl.*, *4*(7), 23–29.
Kişi, Ö. (2007), Streamflow forecasting using different artificial neural network algorithms, *J. Hydrol. Eng.*, *12*(5), 532–539.
Kohonen, T. (2012), *Self-Organization and Associative Memory*, Springer, Heidelberg, Fed. Rep. of Germany.
Kumar, A. R. S., M. K. Goyal, C. S. P. Ojha, R. D. Singh, and P. K. Swamee (2013), Application of artificial neural network, fuzzy logic and decision tree algorithms for modelling of streamflow at Kasol in India, *Water Sci. Technol.*, *68*(12), 2521–2526.
Lavers, D. A., D. E. Waliser, F. M. Ralph, and M. D. Dettinger (2016), Predictability of horizontal water vapor transport relative to precipitation: Enhancing situational awareness for forecasting western US extreme precipitation and flooding, *Geophys. Res. Lett.*, *43*, 2275–2282, doi:10.1002/2016GL067765.
Lek, S., and J.-F. Guégan (1999), Artificial neural networks as a tool in ecological modelling, an introduction, *Ecol. Modell.*, *120*(2), 65–73.
Leung, L. R., and Y. Qian (2009), Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model, *Geophys. Res. Lett.*, *36*, L03820, doi:10.1029/2008GL036445.
Li, X.-Z., L.-Z. Xu, and Y.-G. Chen (2010), *Implicit Stochastic Optimization With Data Mining for Reservoir System Operation*, pp. 2410–2415, IEEE, Qingdao, China.
Liaw, A., and M. Wiener (2002), Classification and regression by random, *Forest. R News*, *2*(3), 18–22.
Lima, A. R., A. J. Cannon, and W. W. Hsieh (2016), Forecasting daily streamflow using online sequential extreme learning machines, *J. Hydrol.*, *537*, 431–443.
Lin, J.-Y., C.-T. Cheng, and K.-W. Chau (2006), Using support vector machines for long-term discharge prediction, *Hydrol. Sci. J.*, *51*(4), 599–612.
Linares-Rodriguez, A., V. Lara-Fanego, D. Pozo-Vazquez, and J. Tovar-Pescador (2015), One-day-ahead streamflow forecasting using artificial neural networks and a meteorological mesoscale model, *J. Hydrol. Eng.*, *20*(9), 05015001.
Lü, A., S. Jia, W. Zhu, H. Yan, S. Duan, and Z. Yao (2011), El Niño-Southern Oscillation and water resources in the headwaters region of the Yellow River: Links and potential for forecasting, *Hydrol. Earth Syst. Sci.*, *15*(4), 1273–1281.
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*(7), 917–923.
Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis (1997), A Pacific interdecadal climate oscillation with impacts on salmon production, *Bull. Am. Meteorol. Soc.*, *78*(6), 1069–1079.
Mingers, J. (1989) An empirical comparison of selection measures for decision-tree induction, *Mach. Learn.*, *3*(4), 319–342.
Montoya, E., J. Dozier, and W. Meiring (2014), Biases of April 1 snow water equivalent records in the Sierra Nevada and their associations with large-scale climate indices, *Geophys. Res. Lett.*, *41*, 5912–5918, doi:10.1002/2014GL060588.
Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith (2007), Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *Trans. ASABE*, *50*(3), 885–900.
Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I: A discussion of principles, *J. Hydrol.*, *10*(3), 282–290.
Pagano, T., and D. Garen (2003), *World Water and Environmental Resources Congress 2003*, pp. 1–9, American Society of Civil Engineers, Philadelphia, Pa.
Quinlan, J. R. (1986), Induction of decision trees, *Mach. Learn.*, *1*(1), 81–106.
Quinlan, J. R. (1990), Decision trees and decision-making, *IEEE Trans. Syst. Man Cybern.*, *20*(2), 339–346.
Ralph, F., and M. Dettinger (2012), Historical and national perspectives on extreme West Coast precipitation associated with atmospheric rivers during December 2010, *Bull. Am. Meteorol. Soc.*, *93*(6), 783–790.
Redmond, K. T., and R. W. Koch (1991), Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices, *Water Resour. Res.*, *27*(9), 2381–2399.
Revilla-Romero, B., H. E. Beck, P. Burek, P. Salamon, A. de Roo, and J. Thielen (2015), Filling the gaps: Calibrating a rainfall-runoff model using satellite-derived surface water extent, *Remote Sens. Environ.*, *171*, 118–131.

Schnier, S., and X. M. Cai (2014), Prediction of regional streamflow frequency using model tree ensembles, *J. Hydrol.*, *517*, 298–309, doi: 10.1016/j.jhydrol.2014.05.029.
Schwing, F., T. Murphree, and P. Green (2002), The Northern Oscillation Index (NOI): A new climate index for the northeast Pacific, *Progr. Oceanogr.*, *53*(2), 115–139.
Shamim, M. A., M. Hassan, S. Ahmad, and M. Zeeshan (2016), A comparison of artificial neural networks (ANN) and local linear regression (llr) techniques for predicting monthly reservoir levels, *KSCE J. Civ. Eng.*, *20*(2), 971–977.
Slowik, A., and M. Bialko (2008), *Training of Artificial Neural Networks Using Differential Evolution Algorithm*, pp. 60–65, IEEE, Krakow, Poland.
Smith, C. A., and P. D. Sardeshmukh (2000), The effect of ENSO on the intraseasonal variance of surface temperatures in winter, *Int. J. Climatol.*, *20*(13), 1543–1557.
Smola, A., and V. Vapnik (1997), Support vector regression machines, *Adv. Neural Inform. Process. Syst.*, *9*, 155–161.
Smola, A. J., and B. Schölkopf (2004), A tutorial on support vector regression, *Stat. Comput.*, *14*(3), 199–222.
Stenseth, N. C., G. Ottersen, J. W. Hurrell, A. Mysterud, M. Lima, K. S. Chan, N. G. Yoccoz, and B. Ådlandsvik (2003), Studying climate effects on ecology through the use of climate indices: The North Atlantic Oscillation, El Nino Southern Oscillation and beyond, *Proc. R. Soc. London Ser. B*, *270*(1529), 2087–2096.
Thirumalaiah, K., and M. C. Deo (1998), River stage forecasting using artificial neural networks, *J. Hydrol. Eng.*, *3*(1), 26–32.
Trenberth, K. E., and D. P. Stepaniak (2001), Indices of El Niño evolution, *J. Clim.*, *14*(8), 1697–1701.
Turner, S. W. D., and S. Galelli (2016), Regime-shifting streamflow processes: Implications for water supply reservoir operations, *Water Resour. Res.*, *52*, 3984–4002, doi:10.1002/2015WR017913.
Vapnik, V. N. (1999), An overview of statistical learning theory, *IEEE Trans. Neural Netw.*, *10*(5), 988–999.
Wang, C., and D. B. Enfield (2001), The tropical Western Hemisphere warm pool, *Geophys. Res. Lett.*, *28*(8), 1635–1638.
Wang, W., P. H. Van Gelder, J. Vrijling, and J. Ma (2006), Forecasting daily streamflow using hybrid ANN models, *J. Hydrol.*, *324*(1), 383–399.
Wei, C. C. (2012), Discretized and continuous target fields for the reservoir release rules during floods, *Water Resour. Manage.*, *26*(12), 3457–3477.
Wei, C. C., and N. S. Hsu (2009), Optimal tree-based release rules for real-time flood control operations on a multipurpose multireservoir system, *J. Hydrol.*, *365*(3-4), 213–224.
Weisberg, R. (1996), *On the Evolution of SST Over the PACS Region*, p. 378, American Meteorological Society, Atlanta, Ga.
Werbos, P. (1974), *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Harvard University, Cambridge, Mass.
Wolter, K., and M. S. Timlin (1998), Measuring the strength of ENSO events: How does 1997/98 rank?, *Weather*, *53*(9), 315–324.
Wu, C., K. Chau, and Y. Li (2009), Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques, *Water Resour. Res.*, *45*, W08432, doi:10.1029/2007WR006737.
Yang, T., X. Gao, S. Sorooshian, and X. Li (2016), Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme, *Water Resour. Res.*, *52*, 1626–1651, doi:10.1002/2015WR017394.
Yao, X. (1999), Evolving artificial neural networks, *Proc. IEEE*, *87*(9), 1423–1447.
Yaseen, Z. M., A. El-Shafie, O. Jaafar, H. A. Afan, and M. N. Sayl (2015), Artificial intelligence based models for stream-flow forecasting: 2000-2015, *J. Hydrol.*, *530*, 829–844.
Zealand, C. M., D. H. Burn, and S. P. Simonovic (1999), Short term streamflow forecasting using artificial neural networks, *J. Hydrol.*, *214*(1), 32–48.
Zhang, Y., J. M. Wallace, and D. S. Battisti (1997), ENSO-like interdecadal variability: 1900-93, *J. Clim.*, *10*(5), 1004–1020.

Erratum

One of the funders of this study was omitted from the Acknowledgments section of the originally published version of record. The article has been updated, and this may be considered the official version of record.