Selected model fusion: an approach for improving the accuracy of monthly streamflow forecasting

Fereshteh Modaresi, Shahab Araghinejad and Kumars Ebrahimi

ABSTRACT

Monthly streamflow forecasting plays an important role in water resources management, especially for dam operation. In this paper, an approach of model fusion technique named selected model fusion (SMF) is applied and assessed under two strategies of model selection in order to improve the accuracy of streamflow forecasting. The two strategies of SMF are: fusion of the outputs of best individual forecasting models (IFMs) selected by dendrogram analysis (S1), and fusion of the best outputs of all IFMs resulting from an ordered selection algorithm (S2). In both strategies, five data-driven models including: artificial neural network, generalized regression neural network, least square-support vector regression, K-nearest neighbor regression, and multiple linear regression with optimized structure are performed as IFMs. The SMF strategies are applied for forecasting the monthly inflow to Karkheh reservoir, Iran, owning various patterns between predictor and predicted variables in different months. Results show that applying SMF approach based on both strategies results in more accurate forecasts in comparison with fusion of all IFMs outputs (S3), as the benchmark. However, comparison of the two SMF strategies reveals that the implementation of strategy (S2) considerably improves the accuracy of forecasts than strategy (S1) as well as the best IFM results (S4) in all months.

Key words | Karkheh River, K-nearest neighbor, monthly streamflow forecasting, neural networks, selected model fusion, support vector regression

Fereshteh Modaresi (corresponding author)
Shahab Araghinejad
Kumars Ebrahimi
University of Tehran,
Tehran,
Iran
E-mail: *fmodaresi@ut.ac.ir*

INTRODUCTION

One of the most important issues in the management of dams is awareness of the amount of inflows to reservoirs. As the rule curves of dams are usually in a monthly time scale, the accurate forecast of monthly inflow to dams is the main concern of hydrologists and water managers.

In order to forecast streamflow in long-lead time scales, like monthly, data-driven models such as neural networks (generalized regression (GRNN) and artificial (ANN)) (Cigizoglu 2005; Wu *et al.* 2010; Chen *et al.* 2014), K-nearest neighbor (KNN) regression (Araghinejad *et al.* 2006; Meidani & Araghinejad 2014), and support vector regression (SVR) (Wang *et al.* 2009; Su *et al.* 2014) are usually applied because they are able to recognize different relations between predictor and predicted variables for forecasting this process and produce the most appropriate forecasts. Nevertheless, each of these models contains estimation errors that are inevitable, and somehow lead to a decline in the accuracy of the forecasts.

In order to decrease the forecasting errors, model fusion method has been introduced and applied. This method can be performed in different forms including parallel, series, and mixed (Dasarathy 1994). In parallel form, the results of several individual models are combined by weighting or bootstrap methods (Srinivas & Srinivasan 2001), while in series form the results of only one model is post-processed and fed as input to another model (Bai *et al.* 2016; Chau 2017). The mixed form of model fusion is a combination of parallel and series forms, where the results of several

doi: 10.2166/hydro.2018.098

individual models are combined by another model (Shamseldin *et al.* 2007).

Since the performance of several models are applied in parallel and mixed forms of model fusion for increasing the accuracy of the results, these methods have been implemented in several studies for hydrological forecasting but mostly for short-term forecasting (e.g., See & Abrahart 2001; Xiong *et al.* 2001; Abrahart & See 2002; Shu & Burn 2004; Goswami & O'Connor 2007; Shamseldin *et al.* 2007; Srinivasulu & Jain 2009), where different data-driven and conceptual models have been used as individual models and various statistical methods including simple and weighted average, selection of the best initial results, fuzzy methods as well as different neural networks have been implemented for combining the results of individual models. In all of these studies, the results of all individual models were used in model fusion phase and their results showed that while model fusion methods produced more accurate forecasts than individual models, ANN model with the structure of multi layer perceptron (MLP) outperformed other methods for model fusion process. On this basis, Chau & Wu (2010), Chen *et al.* (2015), and Bai *et al.* (2016) developed hybrid or series models by the use of ANN (with MLP structure) for increasing the accuracy of daily and monthly rainfall and streamflow forecasting.

Furthermore, applying a mixed form of model fusion technique in the forecast of peak flow and seasonal streamflow has led to improving the accuracy of the results while it has been implemented by the use of neural networks and KNN methods (Azmi *et al.* 2010; Araghinejad *et al.* 2011).

However, due to the different abilities of individual models, Arsenault *et al.* (2015) applied an approach of model selection in order to combine the hydrographs obtained from four lumped hydrological models while nine optimization averaging methods based on optimizing probability distributions functions, minimizing root mean square error (RMSE) and maximizing Nash–Sutcliffe criteria were assessed to determine the number of individual models for combination. The results of this study revealed that the model fusion techniques based on minimized (RMSE) and maximized Nash–Sutcliffe produced more accurate results than other methods. Moreover, the best identified combinations of hydrographs did not include all the simulated hydrographs by all individual models.

Since the structure of data-driven models such as ANN, SVR, KNN, GRNN used as forecasting models are different, their abilities for forecasting process are different in various conditions like linearity and nonlinearity of relationships between predictor and predicted variables or the forecast of extreme values (Modaresi *et al.* 2018). Therefore, it seems that applying the approach of selected model fusion (SMF) in the forecasting process rather than the fusion of all individual models can promote the accuracy of the forecasts. By the implementation of this approach, the best models or results of the individual models are selected for model fusion so as to improve the forecasts. However, by choosing the outputs of the most appropriate individual models as in the research of Arsenault *et al.* (2015), some good results in the output of the eliminated models will be ignored, while applying them for model fusion process can improve the accuracy of the results.

Therefore, the aim of this study is to apply an approach of model fusion named SMF in order to improve the accuracy of monthly streamflow forecasting, and it is performed and assessed based on two strategies of combination consisting of combining the outputs of best individual forecasting models (IFMs) selected by dendrogram analysis (S1), and fusion of the best outputs of all IFMs resulting from an ordered selection algorithm (S2) presented in this study. The difference between strategies S1 and S2 is that in S1 all results of selected models are used for combination while in S2 the best result of all individual models is selected for each time step for model fusion process.

The performances of these strategies are evaluated based on the performance ratings of four assessment criteria and compared to the performance of two other strategies of model fusion as benchmarks which have been used in previous studies (See & Abrahart 2001; Abrahart & See 2002; Shu & Burn 2004; Goswami & O'Connor 2007; Shamseldin *et al.* 2007; Srinivasulu & Jain 2009), including combining all model outputs (S3) and selecting the outputs of the best individual model (S4). In the current study, the forecast of monthly inflow to Karkheh reservoir is implemented as a case study. Modaresi *et al.* (2018) revealed that there are different linear and nonlinear patterns between predictors and predicted variables in the process of forecasting the monthly inflow to this reservoir in various months, while the performance of IFMs are different in these conditions.

Therefore, using this case study, the proposed strategies for combining the IFMs' forecasts can be assessed in a better way in different months.

CASE STUDY AND DATA

In the current paper, the forecast of monthly inflow to the Karkheh reservoir is investigated. It is a strategic dam on the Karkheh River in the southwest of Iran, which supplies the water demands of the Khuzestan plain, which is an agricultural area of Iran. Karkheh River is the main river of the Karkheh basin, which drains an area of more than 50,000 km^2. The initial branches of this river are the Gamasiab and Gharesu rivers which join together and produce the Seimareh River. The confluence of the Seimareh and Kashkan rivers creates the Karkheh River. Figure 1 illustrates the location of the Karkheh Dam, Karkheh River and its branches in Iran.

Long-term data of inflow to Karkheh reservoir over 32 water years, from 1982 to 2013, demonstrate that the main period of water filling of this reservoir is from November to June when the streamflow is affected by precipitation and the snow melting process (Modaresi *et al.* 2018). Therefore, in the current study, the process of forecasting monthly streamflow has been done for this period, i.e., from November to June, while the best predictors for each month have been chosen from the variables, including monthly rainfall and snow area extent (SAE) of the upper sub-basin of the reservoir, and monthly inflow to the reservoir. SAE were obtained from images of the MODIS/TERRA satellite (MOD 10.A2).

Since the predictors of streamflow for each month are different from other months, the process of forecasting monthly streamflow has been performed for each month independently of the other months, while 22 data (1982–2003) and 10 data (2004–2013) have been used for calibration and validation phase of modeling for each month, respectively.

METHODOLOGY

In order to implement the SMF approach for monthly streamflow forecasting, a five-step algorithm, as shown in Figure 2, has been applied as follows.

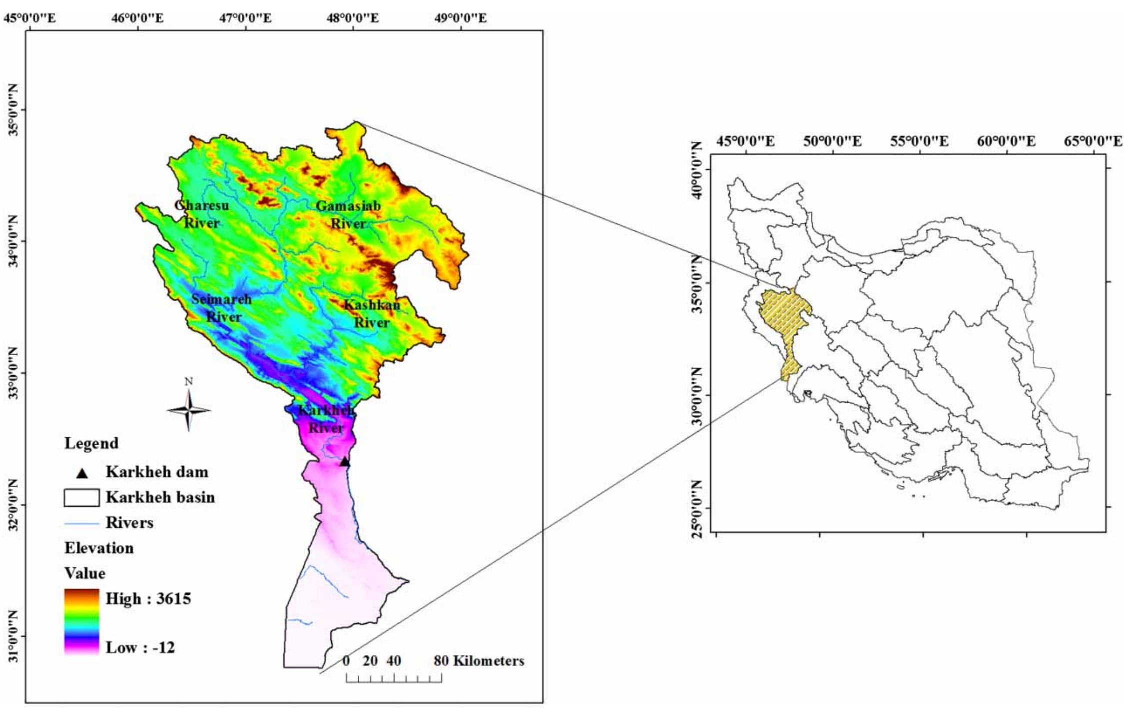

Figure 1 | Location of the Karkheh basin and its river and dam in Iran.

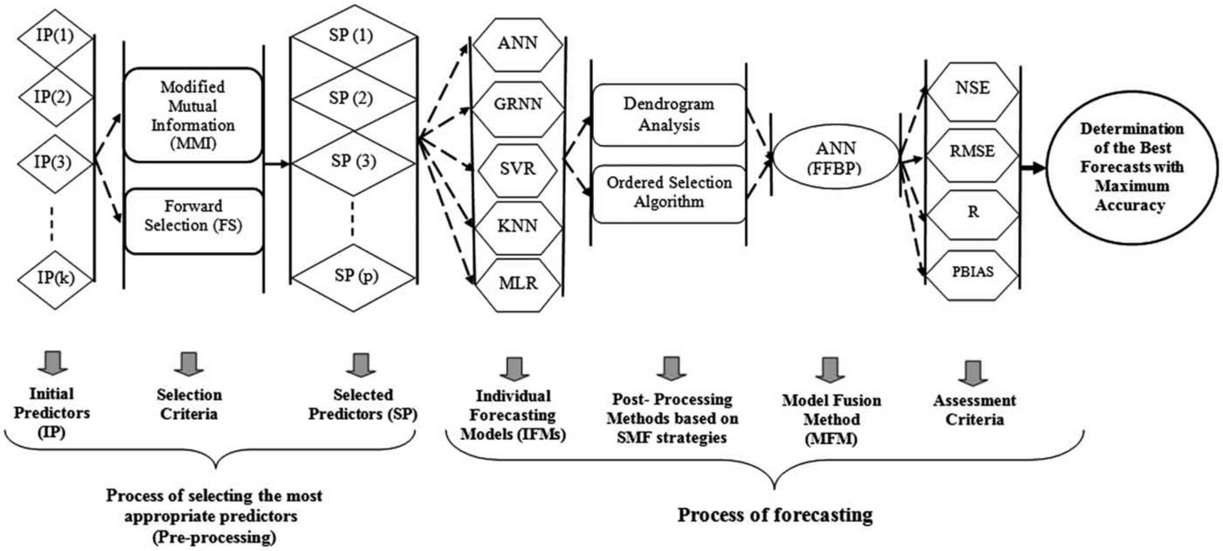

Figure 2 | Algorithm of the selected model fusion (SMF) approach applied in this study.

- Step 1: Determination of the most appropriate streamflow predictors for each month based on two criteria including modified mutual information (MMI), and forward selection method (FS).
- Step 2: Forecast of the monthly streamflow using five data-driven models with different structures, as IFMs, based on the most appropriate predictors specified in the previous step.
- Step 3: Post-process of the IFMs' outputs according to each of the two following strategies of model selection in order to choose the best models/results for combination:
  - Strategy 1: Selection of the best IFMs using dendrogram analysis (S1).
  - Strategy 2: Selection of the best outputs of all IFMs for each of the data using an ordered selection algorithm, presented in this study (S2).
- Step 4: Fusion of the IFMs' outputs, based on each of the model fusion strategies, presented in the previous step.
- Step 5: Evaluation of the performance of model fusion strategies according to the performance ratings of four assessment criteria consisting of Nash–Sutcliffe model efficiency coefficient (NSE), RMSE, correlation coefficient (R), and percent bias (PBIAS).

Furthermore, in order to evaluate the efficiency of the strategies of the SMF approach, the results of these strategies are compared to the results of two other strategies of model fusion as benchmarks which were used in previous studies (See & Abrahart 2001; Abrahart & See 2002; Shu & Burn 2004; Goswami & O'Connor 2007; Shamseldin et al. 2007; Srinivasulu & Jain 2009) including combination of all IFMs' outputs (S3), and selection of the outputs of the best individual models (S4).

The models and methods applied for each step are as follows.

Selection of the most appropriate predictors

The most appropriate predictors in this study have been selected based on two criteria as follows.

Modified mutual information

Modified mutual information (MMI) is an index based on mutual information index (MI) and entropy roles. While only linear relationships can be detected by correlation coefficient index, all of the dependencies between predictors and predicted variables including linear and nonlinear can be detected by MI (Nourani et al. 2015). If x and y are the monthly predictor and predicted variables, respectively, MMI is defined as follows (Modaresi et al. 2016):

$$MMI = \frac{MI(x, y)}{\min(H(x), H(y))} \tag{1}$$

where $MI\ (x,y)$ is the mutual information index and $H(x)$ and $H(y)$ are the '*simple entropy*' of the time series of x and y calculated as follows (Gray 2013):

$$MI(x, y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \tag{2}$$

$$\begin{aligned} H(x) &= -\sum_{x \in X} p(x) \log p(x) \geq 0 \\ H(y) &= -\sum_{y \in Y} p(y) \log p(y) \geq 0 \end{aligned} \tag{3}$$

where $p(x,y)$ is the joint probability density function of x and y, and $p(x)$ and $p(y)$ are the marginal probability density functions of x and y, respectively. If x does not give any information about y and vice versa which means that they are independent, their MI will be zero while a strong dependence between them leads to a high value of MI. Since the range of MI based on entropy rules is between '0' and 'Min (H(x) and H(y))' (Gray 2013), the range of the MMI index is between 0 and 1 for complete dependency and independency of the variables, respectively. While MI index does not have a constant range, the constant range of MMI allows the user to select simply the appropriate predictors based on a specified limit.

In the current study, an initial selection of the appropriate predictors has been done based on MMI index with the threshold of 0.5.

Forward selection method

After determination of the appropriate predictors for each month by MMI index, in order to specify the most appropriate of them, the FS method has been applied. Forward selection (FS) method is a three-step technique used for selection of the predictors (Chen *et al.* 1989; Eksioglu *et al.* 2005; Wang *et al.* 2006; Khan *et al.* 2007). Based on this method, at first the best predictor with the most value of MMI has been modeled with the predicted variable. Then, other predictors, one by one, have been added as inputs to the model according to their MMI in descent order. Finally, the best combination of the predictors that resulted in minimum error, i.e., maximum NSE or minimum RMSE, has been selected as the most appropriate predictors. As the patterns between predictors and predicted variables are linear and nonlinear in different months, the ANN model which is able to model different patterns has been applied in the FS method for modeling.

Individual forecasting models

Long-term forecasting process usually is performed using data-driven methods like neural networks, support vector machines, and statistical methods. Since several predictors have been identified as the most appropriate predictors for forecast of monthly inflow to Karkheh reservoir and the patterns between predictors and predicted variables are different in terms of linearity and nonlinearity, i.e., in some months the relationship between them is linear and in other months it is nonlinear (Modaresi *et al.* 2018), in this paper five models with different structures are considered as IFMs in order to ensure a wider spectrum of contributing models.

The optimum structure of all IFMs has been specified for each month by the leave-one-out cross-validation (LOOCV) method. LOOCV method is an n fold cross-validation method for n observed data where the learning algorithm is applied once for each data, using all other data as a training set and using the selected data as a single-item test set (Sammut & Webb 2010). In the current study, the LOOCV method has been performed based on all of the possible values for parameters of each model using programming (coding) in MATLAB, and the parameter's value resulting in the minimum average error has been selected as the optimum one.

The structure of considered IFMs and the names of optimized parameters of them by the LOOCV method are described briefly as follows.

IFM1: ANN

Artificial neural network (ANN) is a universal function approximator that is able to map any complicated linear and nonlinear functions. In this study, a three-layer, fully connected, feed forward back propagation (FFBP) neural network has been used, where the sigmoid and linear activation functions are applied in middle and output layer neurons, respectively.

In the considered neural network, the number of neurons in input and output layers is equal to the number of predictor and predicted variables, respectively, while the optimum number of middle layer neurons as well as the amounts of initial weights and biases of the neuron connections that affect the accuracy of the results have been determined via the LOOCV method for each month in order to forecast streamflow.

IFM2: GRNN

Generalized regression neural network (GRNN) is a probabilistic three-layer neural network involved in the radial basis function (RBF) network and used for solving regression problems (Araghinejad 2014). The probabilistic structure means this model does not face the problem of local minima which other neural networks encounter (Cigizoglu 2005).

In this network, the number of neurons in input and output layers is the same as ANN while the number of hidden layer neurons is specific and equal to the number of observed data (calibration data). However, the value of a parameter, namely, *spread*, which adjusts the RBF should be specified in the calibration phase. The optimized value of this parameter has been determined in this paper via the LOOCV method for each month in order to produce the most accurate forecasts.

IFM3: LS-SVR

Least square-support vector regression (LS-SVR) is a regression type of support vector machine (SVM) where the least square method is applied to find the optimum hyper plane (Suykens *et al.* 2002). In this method, the structural risk minimization principle, which is superior to the traditional empirical risk minimization principle used by conventional neural networks, is applied to recognize the pattern between predictors and predicted data (Modaresi & Araghinejad 2014).

The type of kernel function used in the structure of this model, resulting from a Lagrangian optimization for solving the objective function, affects the accuracy of the results. Therefore, in this paper, in order to produce the most accurate forecasts, all three types of kernel functions available for this model including linear, polynomial, and RBF have been assessed for each month, and the optimum amounts of their parameters have been calculated via the LOOCV method. Finally, the best kernel function that resulted in the most accurate results has been selected as the appropriate function for the intended month.

IFM4: KNN regression

K-nearest neighbor (KNN) regression is a nonparametric regression method in which, instead of defining a predetermined parametric relation (i.e., linear or nonlinear) between predictor and predicted variables, the information derived from the K number of observed data (nearest neighbors) which are the most similar to the real time data is applied.

In order to specify the amount of similarity of data, a Euclidean distance function is applied in this method where the amount of predictor weights affects the recognition of nearest neighbors. Furthermore, the results of this model are affected by the number of nearest neighbors (K). Therefore, in this paper, the optimum amounts of the two variables for each month have been determined via the LOOCV method.

IFM5: MLR

Multiple linear regression (MLR) is a parametric method, able to model a linear relationship between two or more variables. In a forecasting process, if there is a linear relationship between predictors and predicted variable, this method can model the relationship between them in a good manner (Araghinejad 2014). In the structure of this model for a forecast process, one coefficient is given to each of the predictors while its optimum amount is determined via the least square optimization by the use of calibration data.

Post-processing methods

In the current paper, two different post-processing methods have been performed on the results of IFMs based on each of the SMF strategies. The first method is to select the most appropriate IFMs by the use of dendrogram analysis, and the second one is to select the best output of IFMs for each data based on an ordered selection algorithm presented in this study.

The post-processing methods are as follows.

Dendrogram analysis

Dendrogram analysis is the post-processing method used for strategy 1 (S1). Dendrogram is a data mining procedure included in hierarchical clustering methods (Manning & Schütze 1999). Using a dendrogram as a post-processing method, the IFMs are classified as compared to the observed data based on the similarity of their results to the observed data. The IFMs placed in a class (or branch) with observed data have the most similarity to it, and they are the best IFMs for fusion. Therefore, only the forecasts of the best IFMs, identified by dendrogram, will be used for the model fusion process.

In this method, the similarity of the IFMs' results to the observed data is determined via calculating their distance from each other and from the observed data.

In this paper, seven distance functions have been used to determine the distances between the model forecasts (Y_t) and observed streamflow (T_t) as follows.

1. **Minkowski distance function** (Chino & Yaguchi 1994):

$$Dist_{Minkowski} = \sqrt[g]{\sum_{t=1}^{n} |T_t - Y_t|^g} \tag{4}$$

where n is number of observed data. The values of 1, 2, and ∞ for g produce **City block**, **Euclidean**, and **Chebyshev** distance functions, respectively.

2. **Correlation distance function** (Székely et al. 2007):

$$Dist_{Correlation} = 1 - \frac{\operatorname{cov}(T_t, Y_t)}{\sqrt{\operatorname{var}(T_t).\operatorname{var}(Y_t)}} \tag{5}$$

3. **Spearman distance function** (Singhal 2001):

$$Dist_{Spearman} = 1 - \frac{\operatorname{cov}(r_{T_t}, r_{Y_t})}{\sqrt{\operatorname{var}(r_{T_t}).\operatorname{var}(r_{Y_t})}} \tag{6}$$

where r_{Tt} and r_{Yt} are the rank vectors of T_t and Y_t, respectively. If the number of data is n, the average of each of r_{Tt} and r_{Yt} is equal to $(n+1)/2$.

4. **Cosine distance function (Cosine similarity)** (Singhal 2001):

$$Dist_{Co\sin e} = \frac{\sum_{t=1}^{n} T_t \times Y_t}{\sqrt{\sum_{t=1}^{n} (T_t)^2} \times \sqrt{\sum_{t=1}^{n} (Y_t)^2}} \tag{7}$$

It is worth noting that according to each distance function, a separate dendrogram is achieved.

In order to cluster the IFMs' results and observed data, there are several linkage methods consisting of single, complete, centroid, median, average, and ward; among them, only average method can be used for all distance functions (Jain et al. 1999). Therefore, this method has been applied in this paper, calculated as follows (Jain et al. 1999):

$$z(c, s) = \frac{1}{n_c n_s} \sum_{i=1}^{n_c} \sum_{j=1}^{n_s} Dist(Y_{ci}, Y_{sj}) \tag{8}$$

where n_c and n_s are the number of objects in the clusters c and s, respectively. Y_{ci} is the i^{th} object (IFM) in cluster c and Y_{sj} is the j^{th} object (IFM) in cluster s.

To specify the best and most confident dendrogram resulting from distance functions, the cophenetic correlation coefficient is calculated as follows (Sokal & Rohlf 1962):

$$Cophen = \frac{\sum_{i<j} (d_{ij} - \bar{d})(z_{ij} - \bar{z})}{\sqrt{\sum_{i<j} (d_{ij} - \bar{d})^2 \sum_{i<j} (z_{ij} - \bar{z})^2}} \tag{9}$$

where d_{ij} is the distance between i^{th} and j^{th} objects and $\bar{d}$ is the average of d. z_{ij} is the dendrogrammatic distance resulting from linkage method, and $\bar{z}$ is the average of z. The best value of this coefficient is equal to 1 which indicates the most faithful dendrogram.

Ordered selection algorithm

Ordered selection algorithm presented in this study is the post-processing method used for strategy 2 (S2). It is a three-step algorithm by which the best result of all IFMs is selected for each of the validation (real time) data to be used in the model fusion process. The selection process in this algorithm is based on the performance of IFMs for the observed data which is the most similar to the real-time data. Therefore, by using this algorithm, the beneficial abilities of all IFMs are handled to produce the most accurate forecasts in the model fusion process. The steps of this algorithm are as follows.

1. Creation of an error matrix ($D_{l,t}$) using the error of individual model outputs in calibration phase as follows:

$$D_{l,t} = |Y_{l,t} - T_t| \quad t = 1, 2, \ldots, n, \quad l = 1, 2, \ldots, p \tag{10}$$

   where, p is the number of individual models, n is the number of observed (calibration) data, $Y_{l,t}$ is the forecasted streamflow by l^{th} individual model for t^{th} observed data, and T_t is observed streamflow for t^{th} data.
2. Sorting of the matrix ($D_{l,t}$) in ascending order in each column and creation of two matrices based on it as follows:
   (I) An ordered matrix of the calibration outputs of IFMs ($OY_{l,t}$), where there are the forecasted streamflow of all IFMs in ascending order of errors in t^{th} column for t^{th} data.
   (II) An ordered matrix of the number or name of individual models ($OM_{l,t}$) corresponding to the matrix ($OY_{l,t}$). This matrix shows the name or number of the individual models in terms of efficiency for streamflow forecasting for each of the observed data.
3. Ordering of the validation (real-time) outputs of the IFMs resulting from validation (real time) data (X_r) in the matrix ($OV_{l,q}$) , $q = 1, 2, \ldots n_v$; n_{v} is the number of validation data) according to a column of matrix ($OM_{l,t}$) corresponding to the SX_t. SX_t is the most similar X_t to the X_r, having the minimum Euclidean distance from X_r, calculated as follows:

$$\begin{aligned} Dist_{r,t} &= \sqrt{|X_r - X_t|^2} \quad t = 1, 2, \ldots, n \\ SX_t &\approx Min\,(Dist_{r,t}) \end{aligned} \tag{11}$$

According to this algorithm, the best results of all IFMs placed in the first rows of ($OY_{l,t}$) and ($OV_{\text{l,t}}$) are selected and applied for model fusion process based on strategy 2 (S2) of the SMF approach. Figure 3 shows a schematic of the results of the selection processes performed in strategies (S1) and (S2).

Model fusion method (MFM)

In order to combine the results of IFMs, a mixed form of model fusion process has been used in this study where the simultaneous outputs of several models are post-processed and fed to another model as input. The benefit of this method is to apply the capabilities of several forecasting models for producing more accurate results. In order to perform the process of model fusion, an ANN with an optimum structure of three layers, fully connected, and feed forward (FF) has been applied where the number of neurons in the hidden layer as well as the initial weights, and biases of the connections of the neurons in the different layers have been optimized through the LOOCV method.

Although the ANN model (with MLP structure) is a common method, it has been applied in many previous studies and produced the most favorable results (e.g., See & Abrahart 2001; Shamseldin & O'Connor 2001; Xiong et al. 2001; Abrahart & See 2002; Shu & Burn 2004; Goswami & O'Connor 2007; Shamseldin et al. 2007; Chau & Wu 2010; Chen et al. 2015; Bai et al. 2016). Therefore, it has been used with an optimum structure in this study so that the efficiency of the strategies of the SMF approach can be assessed in a better way as compared to the strategies of the combination of all IFMs' outputs (S3), and the selection of the outputs of best individual models (S4), performed in most of the above studies.

It is worth noting that the post-processed outputs of calibration and validation phase of IFMs have been fed to the ANN model as inputs of calibration and validation phase, respectively.

Assessment criteria

The performance criteria used in this study in order to assess the forecasting models are as follows (Nash & Sutcliffe 1970; Moriasi et al. 2007; Araghinejad 2014):

- Nash-Sutcliffe: $NSE = 1 - \dfrac{\sum_{t=1}^{n} (T_t - Y_t)^2}{\sum_{t=1}^{n} (T_t - \bar{T})^2}$ (12)

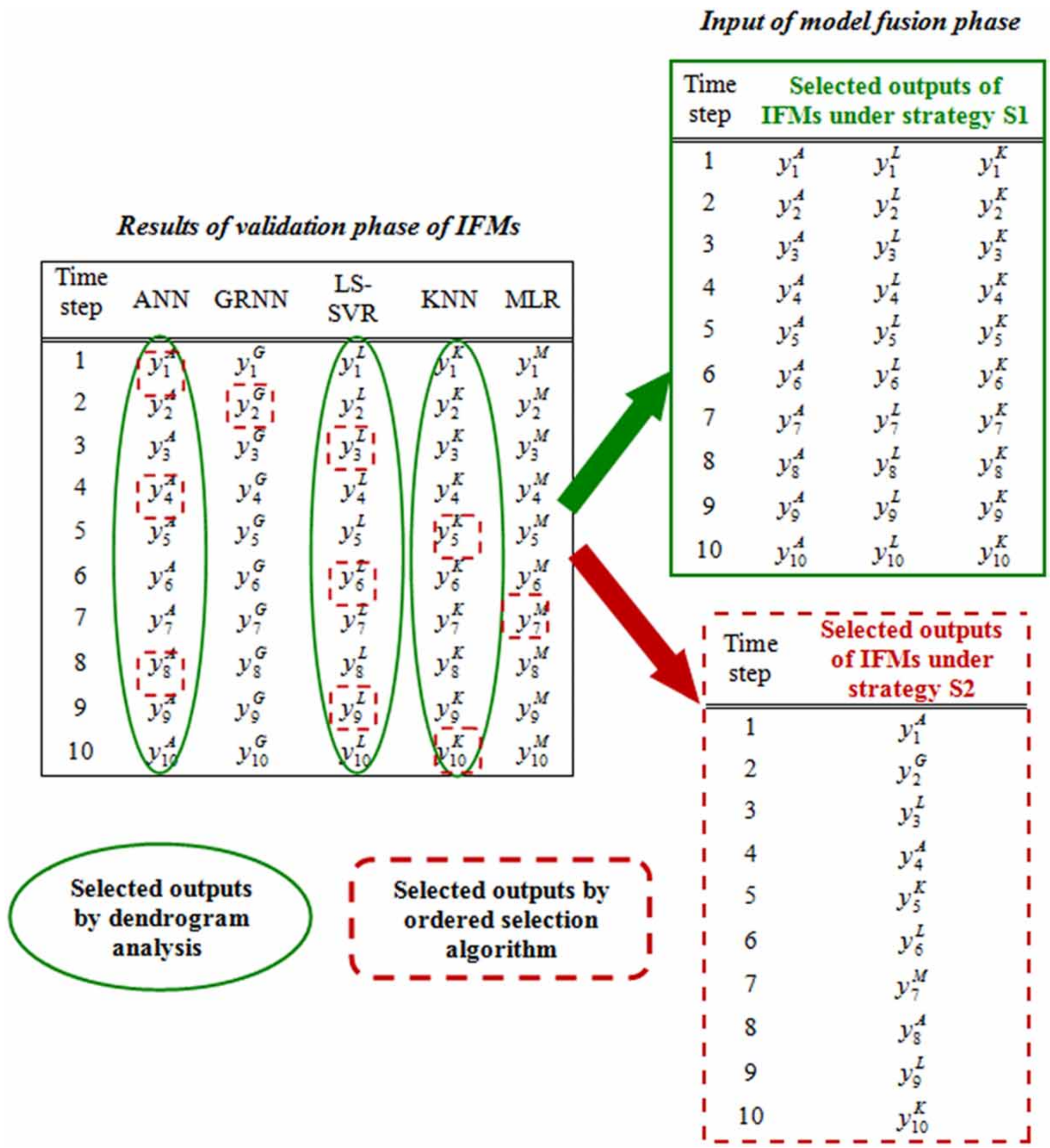

Figure 3 | A schematic of the results of the selection processes performed in strategies (S1) and (S2) of the SMF approach.

- Root mean square error: $RMSE = \sqrt{\frac{\sum_{t=1}^{n}(T_t - Y_t)^2}{n}}$ (13)

- Correlation coefficient:

$$R = \frac{\sum_{t=1}^{n}\left(T_t - \bar{T}\right)\left(Y_t - \bar{Y}\right)}{\sum_{t=1}^{n}\left(T_t - \bar{T}\right).\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)} \quad (14)$$

- Percent bias: $PBIAS(\%) = \frac{\sum_{t=1}^{n}(Y_t - T_t)}{\sum_{t=1}^{n} T_t} \times 100$ (15)

where Y_t and T_t are forecasted (estimated) and observed values of streamflow, respectively, for t^{th} data, $\bar{T}$ and $\bar{Y}$ are the average of observed and forecasted values of streamflow (forecasted variable), and n is the number of data. The general performance ratings of these criteria for monthly streamflow forecasting are presented in Table 1. It is worth noting that the positive and negative values of PBIAS index indicate the overestimation and underestimation of the models, respectively.

Table 1 | General performance ratings of assessment criteria for monthly time step (Moriasi et al. 2007; Diaz-Ramirez et al. 2011)

	Model efficiency limitation			
Performance rating	**NSE**	**RMSE**	**R**	**PBIAS**
Very good	$0.75 < NSE \leq 1$	$0 \leq RMSE \leq 0.5$ SD	$0.93 < R < 1.00$	$PBIAS < \pm 10$
Good	$0.65 < NSE \leq 0.75$	0.5 SD $< RMSE \leq 0.6$ SD	$0.88 < R < 0.92$	$\pm 10 \leq PBIAS < \pm 15$
Satisfactory (Fair)	$0.5 < NSE \leq 0.65$	0.6 SD $< RMSE \leq 0.7$ SD	$0.81 < R < 0.87$	$\pm 15 \leq PBIAS < \pm 25$
Unsatisfactory (Poor)	$NSE \leq 0.5$	$RMSE > 0.7$ SD	$R < 0.80$	$PBIAS > \pm 25$

RESULTS AND DISCUSSION

Result of predictor selection

The most appropriate predictors, achieved from the incorporation of MMI index and FS method, for each month, are shown in Table 2. This table presents the value of MMI and the results of FS method for the best predictors for each of the months.

Results of IFMs

In this study, five IFMs have been implemented in order to forecast monthly inflow to Karkheh reservoir, Iran, from November to June, based on the predictors shown in Table 2. Each of the IFMs has been trained and tested for each month based on the optimum structure determined by the LOOCV method while the optimum structures were different in various months. The optimum values of the IFMs' parameters for each month are shown in Table 3. Since all of the possible values of each parameter have been tested in the LOOCV method, the optimum structure achieved from this method is the absolute optimum one according to all data.

As the aim of this study is to increase the accuracy of the forecast results in the validation phase, in Figure 4, the monthly validation results of the IFMS are illustrated in terms of NSE (a), RMSE (b), R (c), and PBIAS (d).

Table 2 | MMI value and FS results of the most appropriate predictors for forecasting of monthly inflow to Karkheh reservoir from November to June

			Forecasted variable: Monthly streamflow (MCM)							
Predictors			**November**	**December**	**January**	**February**	**March**	**April**	**May**	**June**
Precipitation in October (mm)			0.647							
Precipitation in March (mm)								0.797		
Precipitation in May (mm)										0.617
Streamflow in November (MCM)				0.582						
Streamflow in December (MCM)					0.680					
Streamflow in January (MCM)						0.600				
Streamflow in March (MCM)									0.524	
Streamflow in April (MCM)									0.685	
Streamflow in May (MCM)										0.683
SAE[a] of February							0.515	0.670		
Result of FS method	**Calibration**	**NSE**	0.745	0.789	0.787	0.441	0.631	0.593	0.758	0.916
		RMSE	0.633	0.497	0.651	0.528	0.567	0.352	0.513	0.268
		R	0.922	0.901	0.894	0.676	0.791	0.759	0.821	0.957
	Validation	**NSE**	0.611	0.581	0.678	0.381	0.582	0.412	0.611	0.354
		RMSE	0.659	0.462	0.750	0.900	0.452	0.639	0.676	0.665
		R	0.846	0.833	0.886	0.674	0.756	0.563	0.792	0.832

[a]SAE, Snow Area Extent is the ratio of snow area to total area of the basin (dimensionless).

Table 3 | The optimized values of the IFMs' parameters in each month

	ANN model			LS-SVR model				KNN model			MLR model	
					Kernel parameters[a]				Weight of variables		Coefficient of the predictors	
Month	Initial Weights	Number of hidden layer neurons	GRNN Model Spread	Type of optimized kernel function	Param 1	Param 2	Gamma	K	w1	w2	β_1	β_1
Nov	0.6	6	0.4	Polynomial	1.1	1.0	10.3	5	1.0	–	0.60	–
Dec	0.9	4	1.2	Polynomial	0.1	5.0	5.3	10	1.0	–	0.76	–
Jan	0.1	10	0.3	Polynomial	2.0	3.0	0.1	2	1.0	–	0.60	–
Feb	0.8	5	0.1	Polynomial	0.1	5.0	5.0	2	1.0	–	0.63	–
Mar	0.4	2	0.1	Polynomial	0.7	5.0	0.1	4	1.0	–	0.48	–
Apr	0.1	10	0.2	RBF	0.1	–	53.9	10	0.9	0.1	4.03	2.58
May	0.3	5	0.1	Polynomial	0.9	5.0	10.1	2	0.1	0.9	0.34	4.32
Jun	0.6	7	0.2	Polynomial	2.0	2.0	3.7	4	0.6	0.4	0.75	0.11

[a]Param 1 and 2 for the polynomial kernel functions are constant value (τ) and power (d), respectively. In RBF kernel, param 1 is standard deviation (σ).

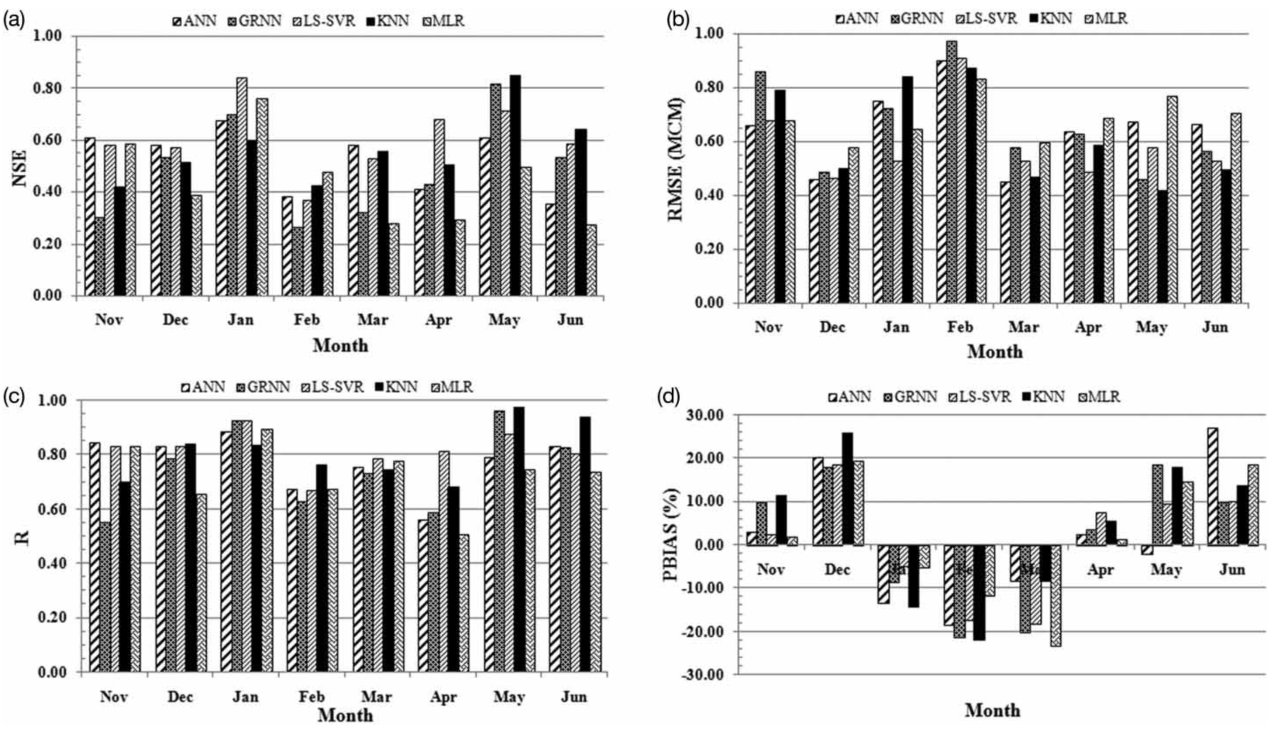

Figure 4 | Results of streamflow forecasting by individual forecasting models based on NSE (a), RMSE (b), R (c), and PBIAS (d) criteria.

It can be seen from Figure 4 that the accuracy of the individual models is different in various months and a specific model could not produce the best outputs for all of the months; for instance, according to NSE, RMSE, and R criteria, the best forecast for November, January, and May belongs to ANN, LS-SVR, and KNN, respectively. Furthermore, although several models with different and optimized structures have been used for monthly streamflow forecasting based on the best selected predictors, the accuracy of the best forecast results according to the performance ratings of NSE, RMSE, and R criteria (shown in Table 1), is in the category of 'Very good' only in January and March, while in other months the best performance is mostly 'Satisfactory' based on all three criteria.

Assessment of the results based on PBIAS index also confirms different performance of the models in different months; such that the performance of all IFMs is in the rating of 'Very good' and 'Good' in November, January, and April, while in other months it is 'Satisfactory' for most of them. Moreover, the results indicate overestimation of all models in November, December, April, May, and June except ANN for May, and underestimation of them for January, February, and March.

Results of post-processing the outputs of IFMs

Based on the two strategies of SMF approach, two post-processing methods including dendrogram analysis and the ordered selection algorithm have been performed on the outputs of IFMs.

In the dendrogram analysis, in order to identify and select the most appropriate IFMs, seven dendrograms according to seven distance functions have been drawn for each month, the most reliable of which with the highest cophenetic coefficient value is illustrated in Figure 5. It is worth noting that the considered threshold for classification in the dendrograms is equal to [$0.7 \times (Max\ z(c,s)$ in each dendrogram)], suggested by Jain et al. (1999).

According to the best dendrogram, the model selection for strategy 1 (S1) of the SMF approach has been performed according to the following two conditions:

1. If the observed data is in a group with one or more of the IFMs, only the results of the cohort IFMs with observed data will be used for model fusion.
2. But if the observed data is in a separate category than IFMs, the results of all IFMs will be used for model fusion because of having the same conditions as compared to the observed data.

Therefore, since the best dendrograms of November, February, and March illustrate that the average distance between the results of all IFMs and observed data is more than the threshold and they are in a separate group than the observed data, all of them are used in the model fusion process; however, it is not for other months. For example, for January and April, only the results of the LS-SVR model and for June, the results of GRNN, KNN, and LS-SVR models are applied for the model fusion process.

In Table 4, a summary of the results obtained from the most reliable dendrogram of each month including the type of best distance function, its cophenetic coefficient,

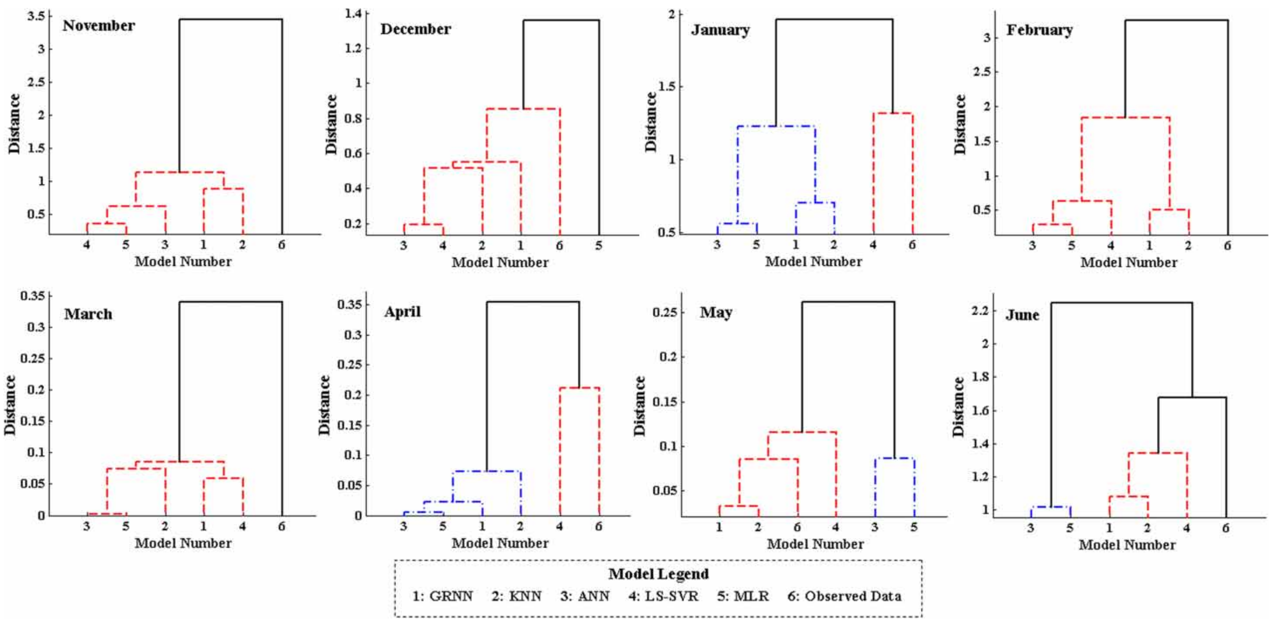

Figure 5 | The most reliable dendrogram of the IFMs' results for November to June.

Table 4 | The summary of the results obtained from the most reliable dendrogram of each month

			Selected IFMs for model fusion				
Month	**Type of best distance function**	**Cophenetic coefficient**	**ANN**	**GRNN**	**LS-SVR**	**KNN**	**MLR**
November	Euclidean	0.9773	✓	✓	✓	✓	✓
December	Chebysheve	0.9230	✓	✓	✓	✓	✗
January	Minkowski	0.8992	✗	✗	✓	✗	✗
February	Chebysheve	0.9861	✓	✓	✓	✓	✓
March	Correlation	0.9903	✓	✓	✓	✓	✓
April	Cosine	0.9169	✗	✗	✓	✗	✗
May	Cosine	0.9329	✗	✓	✗	✓	✗
June	Euclidean	0.8779	✗	✓	✓	✓	✗

and the selected IFMs for model fusion process for strategy 1 (S1) is presented.

According to Table 4, it can be suggested that because the cophenetic coefficient values of the best dendrograms are close to 1, the classification of IFMs resulting from the dendrograms is reliable. Furthermore, applying the various types of distance functions, it can be inferred that although City block, Euclidean, Chebyshev, and Minkowski distance functions have a similar base, Euclidean, and Chebyshev distance functions exceed the types of City block and Minkowski in terms of producing the most reliable dendrograms. Moreover, assessment of the results of correlation, Spearman, and cosine distance functions reveals that the ability of cosine distance function is better than the other two. While no reliable dendrogram has been produced by Spearman distance function, it can be said that applying rank vectors of observations rather than the vector of observations, used in correlation distance function, cannot lead to the most reliable dendrograms. However, because of the type of the most reliable dendrogram being different in various months, it is recommended that different types of dendrograms, especially those of the most reliable, are employed in classification by dendrogram. In addition, Table 4 shows that the performance of the LS-SVR model is better than other models because only for one month the results of this model were not applied for the model fusion process according to strategy 1.

Results of model fusion process

The post-processed outputs of individual models have been combined based on each of the strategies of the SMF approach by the use of an ANN model with an optimized structure achieved from the LOOCV method, which considered the model fusion method. The validation results of both SMF strategies (S1 and S2) are illustrated and evaluated in Figure 6 as compared to the two benchmark strategies, applied in the previous researches (S3 and S4) based on NSE (a), RMSE (b), R (c), and PBIAS (d) criteria. It is worth noting that in strategy 3 (S3) the results of all IFMs without any post-processing are fed as input to the model fusion method, and in strategy 4 (S4) the result of the best IFM is selected as the total result of model fusion; indeed, in this strategy, a weighted averaging is performed where the weight of 1 is given to the best model while other models receive the weight of 0.

As can clearly be seen from Figure 6, the strategy (S2) of the SMF approach has the best performance in all months based on all four criteria as compared to other strategies. Under this strategy (S2), the accuracy of the forecast results in all months except Feb is the rating of 'Very good' based on NSE, RMSE, and R criteria and for February it is 'Good' based on NSE and RMSE and 'Satisfactory' based on R index. However, the performance of this method based on PBIAS index is 'Very good' in all months.

Evaluation of the performance of strategy (S2) as compared to best IFMs (S4) demonstrates that implementation of strategy (S2) has led to improvement of the accuracy of the forecast results according to the performance ratings of NSE, RMSE, and R criteria from 'Satisfactory' to 'Very good' in November, December, March, and June while for February the results' accuracy has progressed from 'Unsatisfactory' to 'Good'. Moreover, based on PBIAS

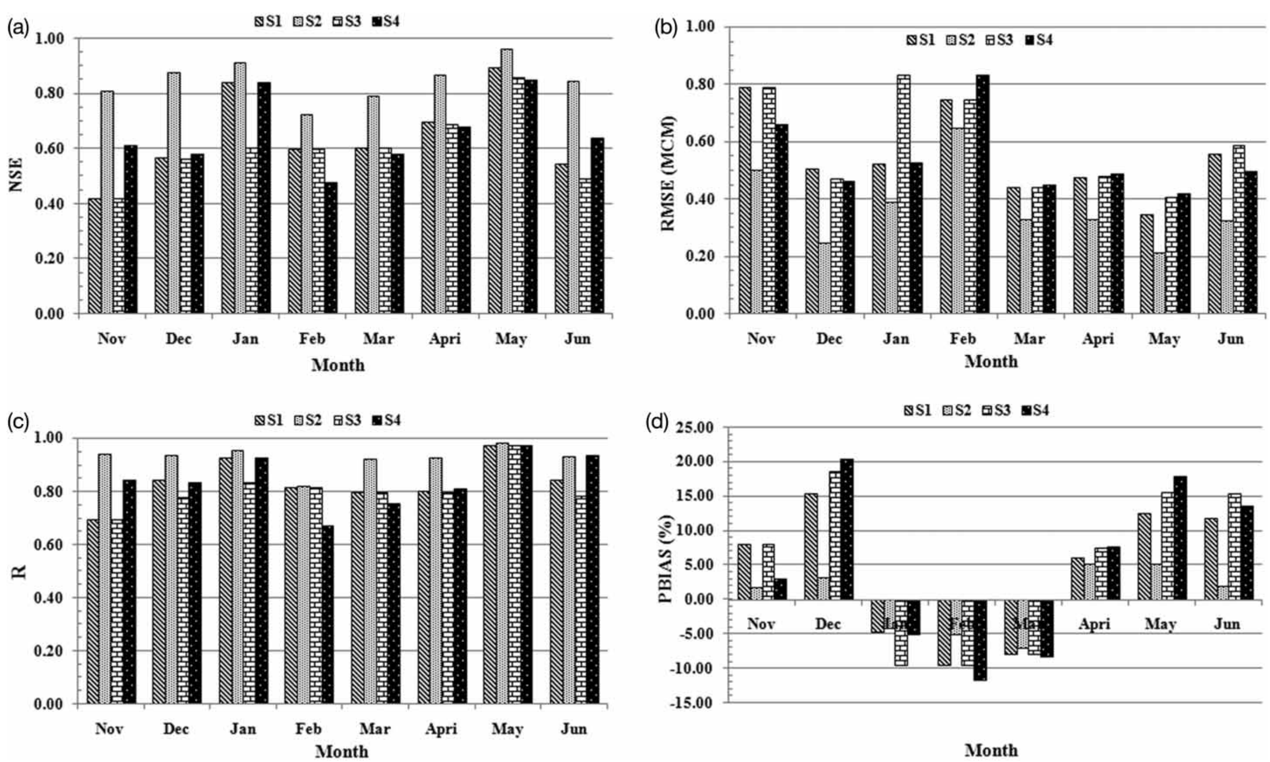

Figure 6 | The results of SMF strategies (S1 and S2) in terms of NSE (a), RMSE (b), R (c), and PBIAS (d) criteria as compared to the benchmark strategies (S3 and S4).

index, the accuracy of the results has improved from 'Satisfactory' to 'Very good' in December and May and from 'Good' to 'Very good' in February and June. Although in January the accuracy of the forecast results of strategy (S2) based on performance ratings of all criteria has not changed from strategy (S4), it has increased by 0.1 in terms of NSE index.

However, assessment of the efficiency of strategy (S1) of the SMF approach reveals that applying this strategy has improved the accuracy of the forecast results only from January to May as compared to strategy (S4) based on all criteria, while the amount of the increase is much less than strategy (S2), because only the accuracy of the forecast results of February has improved from the rating of 'Unsatisfactory' to 'Satisfactory', and in other months, in spite of improvement of the results' accuracy, the performance ratings have not been modified according to all criteria.

Assessing the efficiency of strategies (S1) and (S2) as compared to the strategy (S3) used in the previous studies, it can be suggested that the accuracy of the results of strategy (S3) is much less than strategy (S2) in all months based on all criteria while it is less or equal to strategy (S1) in all months. The accuracy of the results of strategy (S1) in November, February, and March is the same as that of strategy (S3) because all IFMs have been selected via dendrogram analysis for model fusion in these months; but in other months, the model selection strategy (S1) has led to more accurate results than strategy (S3).

The reason why the accuracy of the results of strategy (S2) is significantly more than strategy (S1) is that in strategy (S2) a dynamic selection process is performed because the selected model can be changed in each time step; as a result, all of the efficiency of the IFMs is applied for improvement of the accuracy of the forecast results. In other words, if each of the IFMs can forecast well the streamflow when the value of predictors is in a special range like min, max, or average, all of the models' abilities in order to produce the best forecasts will be employed with the implementation of the strategy (S2) for model fusion. Yet, in strategy (S1), a general selection of models is performed for all data by the dendrogram analysis and the results of the selected models are used for model fusion in all time steps. Therefore, although the output of the most appropriate models, having the least distance from the observed data, are selected and used for model fusion, leaving aside some models means some good results in those models are ignored. Moreover, applying all results of the selected IFMs for this strategy may produce inappropriate patterns for model fusion because some of them have not enough accuracy; as a result, the accuracy of the results of this strategy is less than strategy (S2) and even less than the best IFMs (S4) in November, December, and June.

Consequently, it can be inferred that strategy (S2) of the SMF approach is the best strategy for model fusion as compared to other strategies in order to promote the accuracy of the forecast results. Figure 7 shows the forecast results of strategy (S2) as compared to the results of the best IFMs and observed data in the validation phase for all months.

SUMMARY AND CONCLUSION

In this paper, an approach of model fusion technique named SMF was applied and assessed by two strategies of model selection in order to improve the accuracy of monthly streamflow forecasting. The performed SMF strategies were: fusion of the outputs of best IFMs selected by dendrogram analysis (S1), and fusion of the best outputs of all IFMs resulting from an ordered selection algorithm presented in this study (S2).

For this purpose, five data-driven models with different structures comprising ANN, GRNN, LS-SVR, KNN, and MLR were applied as IFMs in both strategies with optimized structures resulting from the LOOCV method. A mixed structure of model fusion method also was considered to combine the post-processed outputs of IFMs under each SMF strategy, where an ANN with an optimized structure of FFBP was used as the model fusion method.

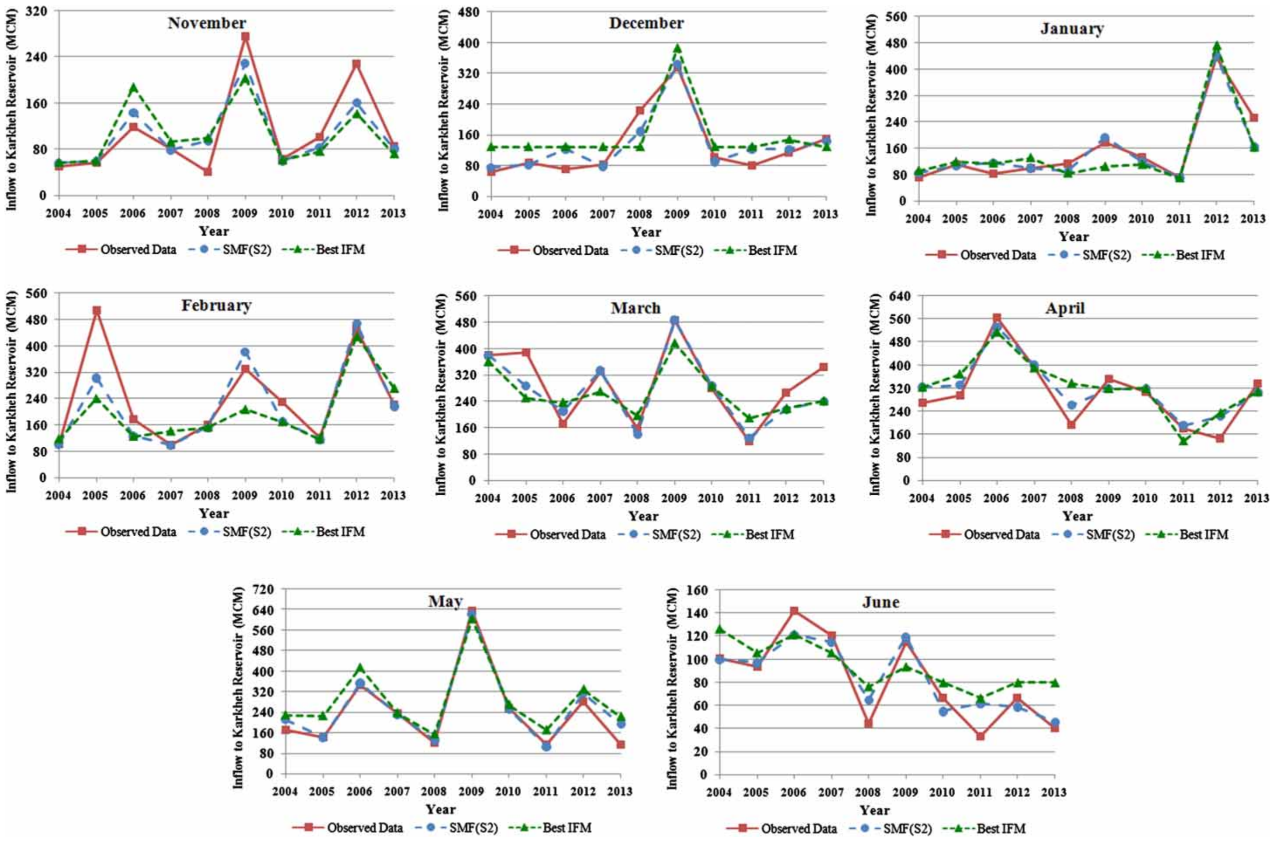

Figure 7 | The results of strategy (S2) of the SMF approach in validation phase in comparison to the results of best IFMs as well as observed data.

The SMF strategies were applied for forecasting the monthly inflow to Karkheh reservoir, Iran, possessing various patterns between predictor and predicted variables in different months. Results showed that while the best performance rating of IFMs for monthly streamflow forecasting was 'Satisfactory' to 'Very good' according to PBIAS index, it was in the rating of 'Satisfactory' in most months based on NSE, RMSE, and R criteria and no specific model could produce the best forecasts for all months. However, performing the SMF strategies, especially strategy (S2), considerably progressed the accuracy of the forecast results than the best results of IFMs from 'Satisfactory' to 'Very good' according to performance ratings of all criteria. Assessment of the SMF strategies as compared to the strategy of combination of the outputs of all IFMs used as a benchmark (S3) also revealed that both SMF strategies resulted in more accurate forecasts than strategy (S3); nevertheless, the forecast results of strategy (S2) were much more accurate than strategy (S3) as well as strategy (S1). This is due to the fact, that in strategy (S2) of the SMF approach, a dynamic selection process was performed and the selected model could be changed in each time step; therefore, the abilities of all IFMs were used for producing the most accurate forecasts. However, in strategy (S1), a general selection of models was performed for all data and the results of the selected models were used for model fusion in all time steps.

Consequently, it can be said that the limitation of this study was to apply only data-driven models for monthly streamflow forecasting which was due to the monthly time scale of predictors. However, the results showed that the presented strategy (S2) of the SMF approach is an efficient method in order to promote the accuracy of monthly forecasts obtained from data-driven models. With respect to the logic applied in this strategy, it can be suggested that the presented strategy of the SMF approach may be beneficial for improving the accuracy of the results in any forecasting process based on model fusion technique with any number and type of individual models.

ACKNOWLEDGEMENTS

The authors are grateful to the University of Tehran for providing the necessary facilities for this study.

REFERENCES

Abrahart, R. J. & See, L. 2002 Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. *Hydrol. Earth Syst. Sci.* **6**, 655–670.

Araghinejad, S. 2014 *Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering*, Vol. 67. Water Science and Technology Library, Springer, Switzerland.

Araghinejad, S., Burn, D. H. & Karamouz, M. 2006 Long-lead probabilistic forecasting of streamflow using ocean-atmospheric and hydrological predictors. *J. Water Resour. Res.* **42** (3), W03431. doi:10.1029/2004WR003853.

Araghinejad, S., Azmi, M. & Kholghi, M. 2011 Application of artificial neural network ensembles in probabilistic hydrological forecasting. *J. Hydrol.* **407**, 94–104.

Arsenault, R., Gatien, P., Renaud, B., Brissette, F. & Martel, J. L. 2015 A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *J. Hydrol.* **529** (3), 754–767.

Azmi, M., Araghinejad, S. & Kholghi, M. 2010 Multi model data fusion for hydrological forecasting using K-nearest neighbor method. *Iran J. Sci. Technol.* **34** (B1), 81–92.

Bai, Y., Xie, J., Wang, X. & Li, C. 2016 Model fusion approach for monthly reservoir inflow forecasting. *Water Sci. Technol.* **18** (4), 634–650.

Chau, K. W. 2017 Use of meta-heuristic techniques in rainfall-runoff modelling. *Water* **9** (3), 186. doi:10.3390/w9030186.

Chau, K. W. & Wu, C. L. 2010 A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. *J. Hydroinform.* **12** (4), 458–473. doi:10.2166/hydro.2010.032.

Chen, S., Billings, S. A. & Luo, W. 1989 Orthogonal least square methods and their application to nonlinear system identification. *Int. J. Contr.* **50**, 1873–1896.

Chen, L., Ye, L., Singh, V., Zhou, J. & Guo, S. 2014 Determination of input for artificial neural networks for flood forecasting using the copula entropy method. *J. Hydrol. Eng.* **19** (11), 04014021.

Chen, X. Y., Chau, K. W. & Wang, W. C. 2015 A novel hybrid neural network based on continuity equation and fuzzy pattern-recognition for downstream daily river discharge forecasting. *J. Hydroinform.* **17** (5), 733–744. doi:10.2166/hydro.2015.095.

Chino, M. & Yaguchi, H. 1994 Generalized Minkowski metrics for mixed feature-type data analysis. *IEEE Trans. Syst. Man Cybern.* **24**, 698–708.

Cigizoglu, H. K. 2005 Generalized regression neural network in monthly flow forecasting. *Civ. Eng. Environ. Syst.* **22** (2), 71–81.

Dasarathy, B. V. 1994 *Decision Fusion.* IEEE Computer Society Press, Los Alamitos, CA.

Diaz-Ramirez, J. N., McAnally, W. H. & Martin, J. L. 2011 Analysis of hydrological processes applying the HSPF model in selected watersheds in Alabama, Mississippi, and Puerto Rico. *Appl. Eng. Agric.* **27** (6), 937–954.

Eksioglu, B., Demirer, R. & Capar, I. 2005 Subset selection in multiple linear regression: a new mathematical programming approach. *Comput. Indust. Eng.* **49**, 155–167.
Goswami, M. & O'Connor, K. M. 2007 Real-time flow forecasting in the absence of quantitative precipitation forecasts: a multi-model approach. *J. Hydrol.* **334**, 125–140.
Gray, R. M. 2013 *Entropy and Information Theory.* Springer-Verlag, New York.
Jain, A. K., Murty, M. N. & Flynn, P. J. 1999 Data clustering: a review. *ACM. Comput. Surv.* **31** (3), 264–323.
Khan, J. A., Aelst, S. V. & Zamar, R. H. 2007 Building a robust linear model with forward selection and stepwise procedures. *Comput. Stat. Data Anal.* **52**, 239–248.
Manning, C. D. & Schütze, H. 1999 *Foundations of Statistical Natural Language Processing.* MIT Press, Cambridge, MA.
Meidani, E. & Araghinejad, S. 2014 Long-lead streamflow forecasting in the southwest of Iran by sea surface temperature of the Mediterranean Sea. *J. Hydrol. Eng.* **19** (8), 05014005.
Modaresi, F. & Araghinejad, S. 2014 A comparative assessment of support vector machines, probabilistic neural networks, and K-nearest neighbor algorithms for water quality classification. *Water Resour. Manage.* **28** (12), 4095–4111.
Modaresi, F., Araghinejad, S. & Ebrahimi, K. 2016 The combined effect of Persian Gulf and Mediterranean Sea surface temperature on operational forecast of spring streamflow for Karkheh basin, Iran. *Sustain. Water Resour. Manage.* **2**, 387–403.
Modaresi, F., Araghinejad, S. & Ebrahimi, K. 2018 A comparative assessment of artificial neural network, generalized regression neural network, least square support vector regression, and K-nearest neighbor regression for monthly streamflow forecasting in linear and nonlinear conditions. *Water Resour. Manage.* **32** (1), 243–258. doi:10.1007/s11269-017-1807-2.
Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systemic quantification of accuracy in watershed simulations. *Tran. ASABE* **50** (3), 885–900.
Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models. Part I – A discussion of principles. *J. Hydrol.* **10** (3), 282–290.
Nourani, V., Khanghah, T. R. & Baghanam, A. H. 2015 Application of entropy concept for input selection of wavelet-ANN based rainfall-runoff modeling. *J. Environ. Informat.* **26** (1), 52–70. doi:10.3808/jei.201500309.
Sammut, C. & Webb, G. I. 2010 *Encyclopedia of Machine Learning.* Springer. https://doi.org/10.1007/978-0-38730164-8.
See, L. & Abrahart, R. J. 2001 Multi-model data fusion for hydrological forecasting. *Comput. Geosci.* **27**, 987–994.
Shamseldin, A. Y. & O'Connor, K. M. 2001 A non-linear neural network technique for updating of river flow forecasts. *Hydrol. Earth Syst. Sci.* **5**, 577–597.
Shamseldin, A. Y., O'Connor, K. M. & Nasr, A. E. 2007 A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models. *Hydrolog. Sci. J.* **52** (5), 896–916.
Shu, C. & Burn, D. H. 2004 Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resour. Res.* **40** (9), W09301.
Singhal, A. 2001 Modern information retrieval: a brief overview. *Bull. IEEE Comput. Soc. Tech. Comm. Data Eng.* **24** (4), 35–43.
Sokal, R. R. & Rohlf, F. J. 1962 The comparison of dendrograms by objective methods. *Taxon* **11**, 33–40.
Srinivas, V. V. & Srinivasan, K. 2001 Post-blackening approach for modeling periodic streamflows. *J. Hydrol.* **241**, 221–269.
Srinivasulu, S. & Jain, A. 2009 River flow prediction using an integrated approach. *J. Hydrol. Eng.* **14** (1). doi:10.1061/(ASCE)1084-0699(2009)14:1(75).
Su, J., Wang, X., Liang, Y. & Chen, B. 2014 GA-based support vector machine model for the prediction of monthly reservoir storage. *J. Hydrol. Eng.* **19** (7), 1430–1437.
Suykens, J. A. K., De Brabanter, J., Lukas, L. & Vandewalle, J. 2002 Weighted least squares support vector machines: robustness and sparse approximation. *Neurocomputing (Special Issue on Fundamental and Information Processing Aspects of Neurocomputing)* **48** (1–4), 85–105.
Székely, G. J., Rizzo, M. L. & Bakirov, N. K. 2007 Measuring and testing dependence by correlation of distances. *Ann. Stat.* **35** (6), 2769–2794.
Wang, X. X., Chen, S., Lowe, D. & Harris, C. J. 2006 Sparse support vector regression based on orthogonal forward selection for the generalised kernel model. *Neurocomputing* **70**, 462–474.
Wang, W. C., Chau, K. W., Cheng, C. T. & Qui, L. 2009 A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol.* **374**, 294–306.
Wu, C. L., Chau, K. W. & Fan, C. 2010 Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *J. Hydrol.* **389** (1–2), 146–167.
Xiong, L., Shamseldin, A. Y. & O'Connor, K. M. 2001 A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. *J. Hydrol.* **245**, 196–217.

First received 30 July 2017; accepted in revised form 2 February 2018. Available online 21 February 2018