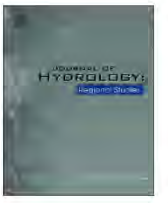
Journal of Hydrology: Regional Studies 44 (2022) 101228

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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



Daily reservoir inflow forecasting using weather forecast downscaling and rainfall-runoff modeling: Application to Urmia Lake basin, Iran

Amirreza Meydani [a,b], Amirhossein Dehghanipour [c], Gerrit Schoups [c,*], Massoud Tajrishy [a]

[a] Department of Civil Engineering, Sharif University of Technology, Tehran, Iran
[b] Department of Geography and Spatial Sciences, University of Delaware, Newark, DE, USA
[c] Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands

ARTICLE INFO

Keywords:
Downscaling
ECMWF
NCEP
Runoff Forecast System
Urmia Lake
Bukan reservoir

ABSTRACT

Study region: This study develops the first daily runoff forecast system for Bukan reservoir in Urmia Lake basin (ULB), Iran, a region suffering from water shortages and competing water demands.
Study focus: A weather forecast downscaling model is developed for downscaling large-scale raw weather forecasts of ECMWF and NCEP to small-scale spatial resolutions. Various downscaling methods are compared, including deterministic Artificial Intelligence (AI) techniques and a Bayesian Belief Network (BBN). Downscaled precipitation and temperature forecasts are then fed into a rainfall-runoff model that accounts for daily snow and soil moisture dynamics in the sub-basins upstream of Bukan reservoir. The multi-objective Particle Swarm Optimization (MOPSO) method is used to estimate hydrological model parameters by maximizing the simulation accuracy of observed river flow (NSE_Q) and the logarithm of river flow (NSE_{LogQ}) in each sub-basin.
New hydrological insights for the region: Results of the weather forecast downscaling model show that the accuracy of the BBN is greater than the various deterministic AI methods tested. Calibration results of the rainfall-runoff model indicate no significant trade-off between fitting daily high and low flows, with an average NSE_Q and NSE_{LogQ} of 0.43 and 0.63 for the calibration period, and 0.54 and 0.57 for the validation period. The entire forecasting system was evaluated using inflow observations for years 2020 and 2021, resulting in an NSE of 0.66 for forecasting daily inflow into Bukan reservoir. The inflow forecasts can be used by policymakers and operators of the reservoir to optimize water allocation between agricultural and environmental demands in the ULB.

1. Introduction

Increased agricultural water consumption along with repercussions of climate change, drought, and mismanagement of water resources leads to competition for water supply between agriculture and environmental sectors and has brought adverse

* Correspondence to: Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands.
E-mail addresses: meydani@udel.edu (A. Meydani), a.dehghanipour@tudelft.nl (A. Dehghanipour), g.h.w.schoups@tudelft.nl (G. Schoups), tajrishy@sharif.edu (M. Tajrishy).

https://doi.org/10.1016/j.ejrh.2022.101228
Received 17 February 2022; Received in revised form 12 August 2022; Accepted 3 October 2022
Available online 8 October 2022

environmental effects such as the destruction of ecosystems and natural bodies of water, especially in arid and semi-arid regions (Dehghanipour et al., 2020; Dunn et al., 2003; Khorasani et al., 2018; Mancosu et al., 2015; Sisto, 2009; Valipour et al., 2015; Xue et al., 2017). Seasonal weather forecast models can be an invaluable tool to manage competing water demands from agriculture and the environment (Rayner et al., 2005). Seasonal weather forecast models enable water resources managers to adapt their actions to pre-emptively resolve any shortages or decide optimal planning. Applying weather forecasts in water resources management leads to better foresight of water-related risks (Grillakis et al., 2018) and increases sustainability in agriculture and natural ecosystems. Several studies accounted for using weather forecasts in flood management (Voisin et al., 2011; Wang et al., 2012), multi-reservoir operation (Ficchì et al., 2016), and optimal agricultural water allocation (Kaune et al., 2020).

The European Center for Medium Range Weather Forecast (ECMWF) and the National Center for Environmental Prediction (NCEP), provide real-time and seasonal weather forecasts. Various studies have evaluated ECMWF seasonal forecast capabilities, including its temperature forecast performance over South America (Gubler et al., 2020), precipitation performance over central Himalaya (Chen et al., 2021), temperature and precipitation skills in the Danube basin (Voisin et al., 2011), and prediction of Caspian sea level (Arpe et al., 2014). The NCEP forecasts showed good performance in various studies for precipitation and temperature forecasting (Phan-Van et al., 2018; Sangelantoni et al., 2019; Yuan et al., 2013) and drought prediction (Lang et al., 2020).

These forecasts usually have a coarse spatial resolution, limiting their direct use in regional-scale applications (Siegmund et al., 2015). Therefore, a downscaling method should be applied to convert the forecasts for regional-scale applications. In this regard, several studies have applied AI in developing precipitation (Kisi and Cimen, 2012), temperature (Chevalier et al., 2011), evaporation (Dehghanipour et al., 2021), and runoff (Adamowski, 2013; Munoz et al., 2021) forecast models using observed historical data (Liu et al., 2021). However, using observational data to forecast future precipitation increases uncertainty because precipitation is a random phenomenon requiring the use of complex dynamics and numerical procedures to be forecast (ECMWF, 2013). Hybrid AI models combine several methods that can reduce uncertainty (Ahmadi et al., 2019). However, few studies have applied hybrid AI for downscaling seasonal weather forecasts (Sun and Lan, 2021).

In addition to using deterministic AI methods, probabilistic models such as BBNs have been applied for estimating weather variables such as precipitation (Hruschka and Nicoletti, 2013). While BBNs are an appropriate approach in situations with high uncertainty and complexity (Molina et al., 2013), no study has used this method to downscale seasonal weather forecasts.

In this study, we focus on weather and runoff forecasting in the ULB. Bukan reservoir is the largest reservoir in ULB, and is located upstream of the Zarrineh Rood River (ZRR). Recent research investigated the long-term climate change effects on water resources in the ZRR basin and indicated a decrease in precipitation upstream of ZRR, leading to a decline in runoff to Bukan reservoir in the future (Dariane and Pouryafar, 2021). In the face of reducing precipitation and increasing agricultural water consumption in ULB, the development of a runoff forecast system for this reservoir can help policymakers optimize water allocation between agricultural and environmental sectors in ULB.

While there have been numerous modeling studies in ULB, runoff forecasting has received little attention, with past studies focusing on either simulating historical hydrological dynamics (Dehghanipour et al., 2020, 2019) or developing long-term scenarios and projections for the distant future (Emami and Koch, 2019; Mahmudi et al., 2021). Moreover, the simulation of snow accumulation and river runoff in sub-basins upstream of Bukan reservoir has only been done on a monthly scale so far (Ahmadaali et al., 2017; Mostafazade and Alizadeh, 2020). Finally, a model that can forecast future water balance components using real-time weather forecasts data has not been proposed for the region because most studies used observed historical data instead of forecast runoff (Gavahi et al., 2019).

In this study, a weather forecast downscaling model is coupled to a daily hydrological rainfall-runoff model to forecast inflow into Bukan reservoir. The resulting runoff and inflow forecasts support water resources management plans to increase the inflow to Urmia lake (UL) and reduce the impacts of both water scarcity and flooding. This research has the following innovations compared to previous studies. First, a weather forecast downscaling model is developed, comparing hybrid deterministic AI and probabilistic BBN methods for downscaling the raw ECMWF and NCEP weather forecasts. Second, an existing rainfall-runoff model is implemented for the first time in this region for modeling daily snow and soil moisture dynamics with a two-storage method that accounts for snow accumulation and melt, actual evapotranspiration, irrigation, surface and subsurface runoff, and groundwater recharge. In contrast, previous studies have only considered monthly dynamics resulting in a less accurate simulation of water balance components. Third, although some research has developed a rainfall-runoff model for this region, these models (e.g., SWAT) are complex and require many input data and calibration parameters. Since the required data of these models are not available in the study area, simplifying assumptions were used to estimate the required data and parameters of the model, leading to errors and uncertainties in river runoff simulation. However, our method uses fewer input data collected in the region leading to no simplifying assumptions, and at the same time, the complexity of the model improves its performance and maintains the capabilities of previous models.

The paper is divided into five sections. Section 2 introduces the study area. Section 3 describes the methodology, including developing the weather forecast model, the downscaling methods, the hydrological rainfall-runoff model, and the calibration approach. Results and discussion are presented in Section 4, and the conclusions are presented in Section 5.

2. Study area

ULB, with an approximate area of 51,760 km^2, is one of the largest endorheic basins in Iran, located in the mountainous region of northwestern Iran. UL is located in the center of ULB, and it is a drainage place of the whole basin's surface water. The surface water supply of UL originates from six important rivers: ZRR, Simineh Rood, Ajichay, Ghadarchay, Shahrchay, and Zolachay (Fig. 1a).

ZRR is the largest and most important river in UL. The average annual runoff of this river is about 2000 MCM, and it supplies more

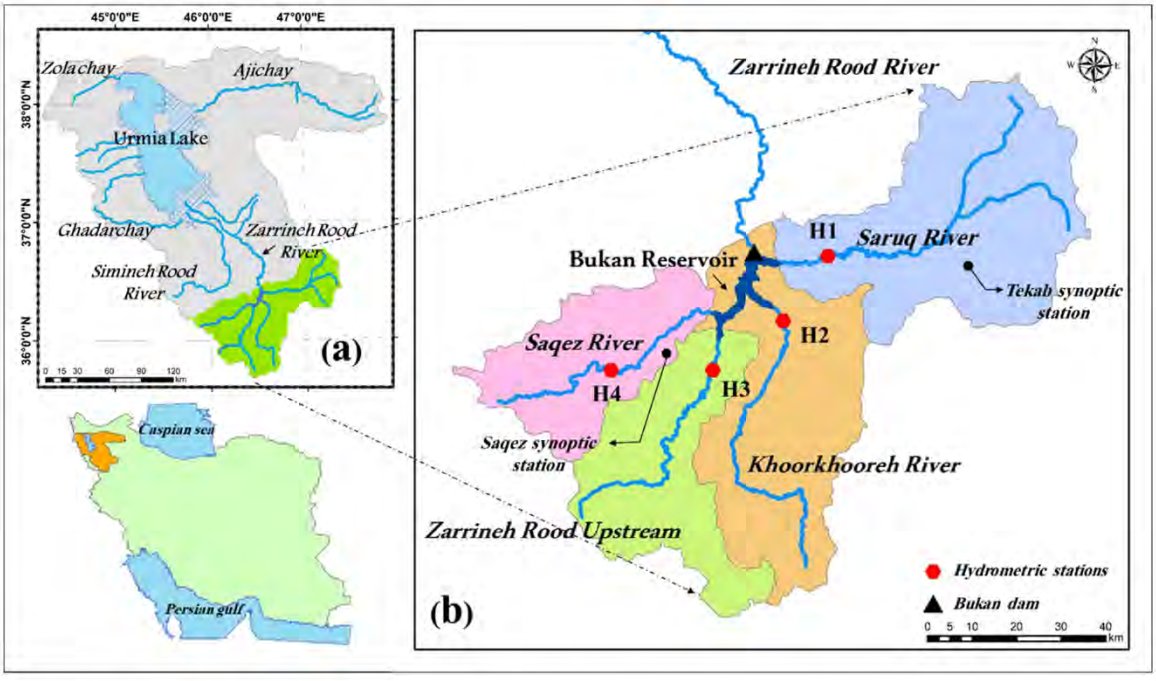


Fig. 1. Location of the study area (a) Bukan reservoir and its main upstream rivers in ULB, Iran; and (b) Sub-basins in upstream Bukan reservoir (See the names of hydrometric stations (H1, H2, H3, and H4) in Table 5).

than 40 % of the total annual environmental flow to UL (Ghaheri et al., 1999; Meydani et al., 2021). Moreover, Bukan reservoir, as the biggest reservoir in the ULB, is located on the ZRR (Dehghanipour et al., 2020, 2019).

In this study, four sub-basins of the ZRR basin located upstream of Bukan reservoir are selected as a case study (Fig. 1). The sub-basins include Saqez River, ZRR upstream, Khoorkhooreh River, and Saruq River. The mean annual precipitation and temperature are about 650 mm and 10 °C, respectively, for these sub-basins (Yekom Consulting Engineers, 2016). Most of the precipitation from mid-December to early March is snow (Fig. 2), and the average snow area is 53 % of the total area (Yekom Consulting Engineers, 2016). Therefore, snow is one of this region's primary water balance components, and the flow regime in these sub-basins is snow-rainy (Yekom Consulting Engineers, 2016). Snowmelt accounts for about 67 % of annual runoff. There are no significant aquifers in the region.

Bukan reservoir was built in 1971 with a storage capacity of 650 MCM, which was increased to 808 MCM in 2005. Saruq, Khoorkhooreh, and Saqez rivers are tributaries of the ZRR and supply additional inflow into Bukan reservoir (Fig. 1b). Bukan reservoir provides drinking water for major cities such as Tabriz, Bukan, and Miyandoab, and it is also a strategic surface water source for irrigated agriculture in Sain Qaleh and Miyandoab plains downstream of the reservoir. Sain Qaleh and Miyandoab plains, with an approximate area of 17,000 ha and 100,000 ha, are the most modern and most extensive agricultural irrigated plains of this basin, and about 64% of the rural employment is in the agricultural sector (Dehghanipour et al., 2020). The net annual releases from Bukan reservoir amount to 500 and 80 MCM of water to meet agricultural and urban water demands in Miyandoab and Sain Qala plains, respectively (Yekom Consulting Engineers, 2016).

3. Methodology

In this study, a runoff forecast model was developed to forecast inflow to Bukan reservoir, an essential and strategic reservoir in ULB. The runoff forecast model helps policymakers optimize allocation to agricultural demands, drinking water, and environmental requirements downstream of Bukan reservoir. As outlined in Fig. 3, this model consists of (1) a weather forecast downscaling model and (2) a hydrological rainfall-runoff model. Meteorological input data for the models were based on observed data from Tekab synoptic station for the Saruq river sub-basin, and from Saqez synoptic station for the other sub-basins (Fig. 1). These two stations were chosen because they have high-quality data and little missing data during the modeling period.

The weather forecast downscaling model was developed to downscale large-scale raw weather forecast data of ECMWF and NCEP to small-scale resolutions. The downscaling procedure helps to reduce errors between observed data and the nearest grid point. The forecasted weather variables include monthly precipitation and daily temperature. Since the rainfall-runoff model has a daily time scale, the monthly forecasted precipitation was disaggregated to daily forecasted precipitation based on the daily distribution of raw forecast data in each month. The spatially downscaled and temporally disaggregated precipitation and temperature data are subsequently used as inputs to the hydrological rainfall-runoff model. The hydrological rainfall-runoff model simulates water balance

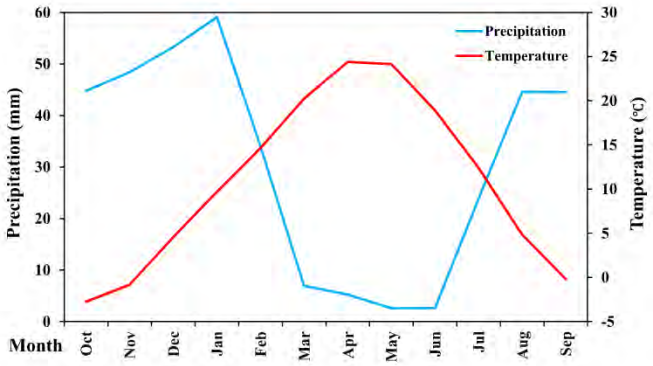


Fig. 2. Ombrothermic diagram for sub-basins in upstream Bukan reservoir using monthly time-averaged observed precipitation and temperature data for the period 1995–2020.
Source: Saqez and Tekab stations, Iran Meteorological Organization, see https://data.irimo.ir/.

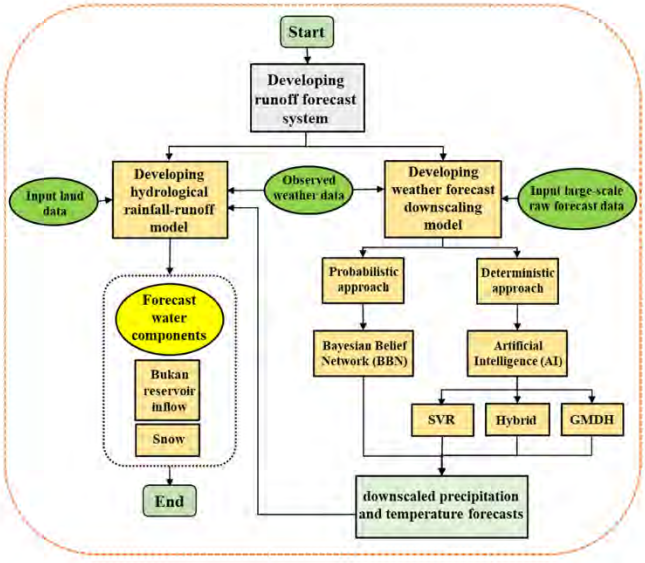


Fig. 3. Outline of the developed runoff forecast system including integrated weather forecast downscaling model and hydrological rainfall-runoff model.

components, including snow storage, snowmelt, actual evapotranspiration, surface runoff, and subsurface flow for sub-basins upstream of Bukan reservoir. The output of the hydrological rainfall-runoff model consists of daily forecasted inflow to Bukan reservoir at the beginning of each month and for a 1-month lead time. In the following sections, we discuss the various parts of both models in more detail. Furthermore, more details are provided in Table 1 about the periods and input data used in each model.

Table 1
Summary of the calibration and validation process for weather forecast downscaling and hydrological rainfall-runoff models.

Model		Data	Time scale	Calibration period (Train)	Validation period (Test)	Forecast period
Weather forecast downscaling	Temperature	ECMWF and NCEP's temperature and observed temperature	Daily	2010–2017	2017–2019	2019–2021
	Precipitation	ECMWF and NCEP's precipitation, Month number, and observed precipitation	Monthly			
Hydrological rainfall-runoff		Observed weather data and observed river discharge	Daily	2003–2014	1999–2003	

3.1. Weather forecast downscaling model

3.1.1. ECMWF forecasts

ECMWF is run by an independent intergovernmental entity supported by 34 European nations. These forecasts are based on mathematical models of the atmosphere and oceans to forecast weather variables based on current and initial weather conditions. Therefore, different meteorological organizations operate different models for forecasting atmospheric conditions, and the results vary for the same area. The ECMWF system has been upgraded every five years, with SEAS5 the fifth and latest seasonal forecast system (Johnson et al., 2019). Seasonal forecasts are possible due to the slow evolution of some of the earth's components and their effect on the atmosphere. The seasonal forecasts provide information on the upcoming season's weather conditions and are useful for various sectors, such as water resources management, agriculture, health, and energy (ECMWF, 2013). Here, we use SEAS5. ECMWF provides seasonal forecast data from 1993 to 2016 as hindcast data calibrated and verified using a set of retrospective seasonal forecasts for past dates and is more accurate compared to the historical data (Johnson et al., 2019). From 2017 onward, forecast instead of hindcast data is available. In this study, the ensemble mean of seasonal monthly and daily weather forecast data, i.e., respectively "Seasonal forecast monthly statistics on the single level" and "Seasonal forecast daily and sub-daily data on the single level", were downloaded (ECMWF, 2021a,b). These global products have a horizontal resolution of 1 degree. The lowest monthly precipitation lead time, i.e. 1-month lead time, was selected for downscaling to have more accurate forecasts. It should be noted that initially, the daily precipitation data were downscaled, while the results were not satisfactory because precipitation is a highly uncertain phenomenon, so monthly precipitation downscaling was considered as the main objective of this study. Daily temperature data were downscaled with a 6-month lead time because temperature forecasts are much less uncertain than precipitation.

3.1.2. NCEP forecasts

The US Government runs NCEP under the leadership of the National Oceanic and Atmospheric Administration and its subsidiary agencies. The second version of the Climate Forecast System (CFSv2) became operational in 2011. CFSv2 has higher skill than version 1, due to improvements in the initial state of its underlying model (Saha et al., 2014). Data used in this study are from the "CFSv2 operational forecasts time series" product and monthly precipitation and temperature, and daily temperature. In this study, to prevent confusion and easy understanding of results, GFSv2's products are called NCEP. For NCEP, precipitation and temperature forecasts were downloaded for 1-month and 6-month lead times, respectively (NCEP, 2011).

3.1.3. AI downscaling methods

The meteorological models used in this study are global and large-scale. The outputs of meteorological models need to be downscaled to small-scale resolutions in the study region to reduce forecast errors. Recently, AI has been used for downscaling weather forecasts due to its ability to identify complex patterns and unforeseen behavior (Benderskaya, 2013), and its capability of extracting non-linear relations from data without prior knowledge of the parametric form of these relations (Valverde Ramírez et al., 2006). Here, we downscale the forecasts to the point-scale of synoptic weather stations, whose data are used to force the hydrological rainfall-runoff model. In this study, three AI methods are used for downscaling raw weather forecast data, as detailed below. Each method contains parameters that are estimated by minimizing the discrepancy between the downscaled precipitation and temperature estimates and their weather-station observed counterparts, with discrepancy quantified by Eq. 1. In addition to NSE, the RMSE criterion is used to quantify performance of the downscaled forecasts:

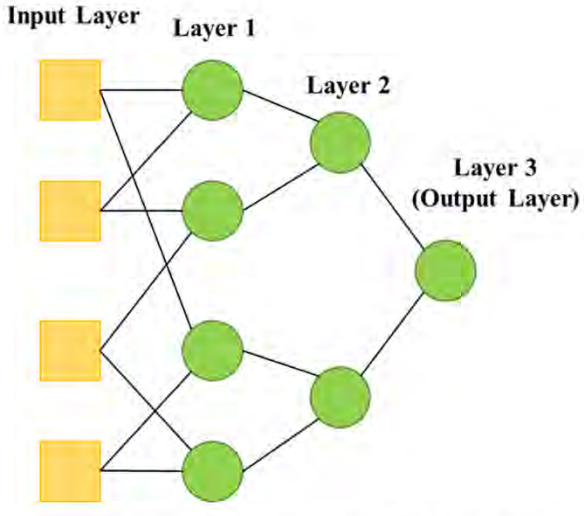


Fig. 4. Structure of GMDH neural network with four inputs.

$$NSE = 1 - \frac{\sum_{i=1}^{m}\left(V_{obs}^{i} - V_{down}^{i}\right)^2}{\sum_{i=1}^{m}\left(V_{obs}^{i} - \bar{V}_{obs}\right)^2} \tag{1}$$

$$RMSE = \sqrt{\sum_{i=1}^{m}\frac{\left(V_{down}^{i} - V_{obs}^{i}\right)^2}{m}} \tag{2}$$

where, V_{obs}^{i} and V_{down}^{i} are observed data and downscaled forecasts, respectively, at the synoptic station. Furthermore, the ratio of the train data to the test data was 70–30. More information on three methods is presented in the following sections.

3.1.3.1. Group method of data handling (GMDH). The first method is based on a multilayer network, where each layer acts as a non-linear function of its inputs (Ivakhnenko, 1970). Input data define the number of neurons. For instance, if the number of input variables is k, the number of neurons in each layer equals $m = \binom{k}{2}$, which will enter the neuron as a binary permutation. Fig. 4 shows the multilayer structure of the GMDH neural network model.

The main objective of this method is to define a function that minimizes the difference between downscaled forecast data and observed data of synoptic stations. At first, as one of the best polynomial equations (Eq. 3), Volterra equation was applied to construct a high order polynomial by (Ivakhnenko, 1970). Generally, Eq. 3 is used to calculate up to the square terms as Eq. 4 to determine the correlation between raw forecast data and observed gauge data.

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \ldots \tag{3}$$

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j \tag{4}$$

where, a_0, a_i, a_{ij} are coefficients of variables in the polynomial that can be calculated by the least squares error method, x_i, x_j, x_k are input variables which are combined pairwise in each neuron based on Eq. 4, and n is the number of variables in each combination ($n = 2$). This method considers all pairwise combinations in a time series, and each combination is fed to each neuron in Fig. 4. In other words, the coefficients of Eq. 4 are estimated in each neuron, and the output is calculated using input variables. Then, the NSE values between model output and data of each neuron are calculated, and neurons are sorted based on NSE values in order to eliminate deficient neurons step by step (Nariman-zadeh et al., 2002). Inputs to the GMDH model for downscaling precipitation are a combination of raw forecast data of precipitation at time step t and t-1, the month number and precipitation of observed gauge data in time step t. Raw forecast and observed gauge data of temperature are input variables to downscaling temperature. Selected neurons are the inputs to the next layer. The process is continued until the last layer in which the obtained neuron is the estimated a_0, a_i, a_{ij} for the time series.

3.1.3.2. Support vector regression (SVR). The idea of SVR is to approximate the best value within the margin called ε-sensitive zone by minimizing the error inside a specific threshold like Fig. 5 for a linear function. Due to the non-linear nature of the problem, with the help of SVR, regression function in the form of Eq. 5 is obtained.

$$f(x) = w\phi(x) + b \tag{5}$$

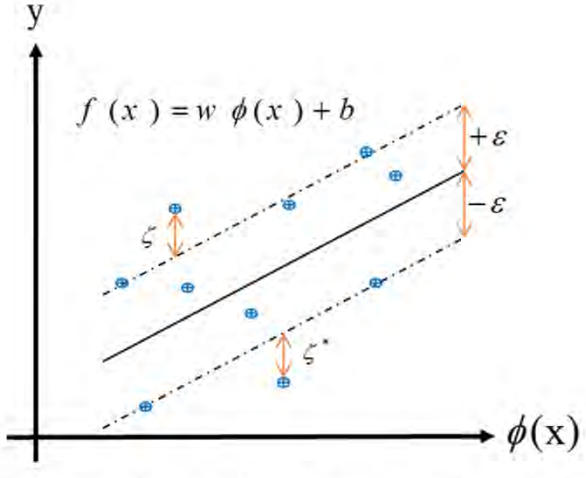


Fig. 5. The best approximation value ε to minimize the error in a specific threshold corresponds to SVR.

where $\phi(x)$ is a non-linear transfer function to map raw forecast data to higher dimensional space. w is a weight factor, and b is a deviation constant.

In Fig. 5, the straight line is the hyperplane that separates the line between two classes in a higher dimension than the actual dimension. The dashed line in Fig. 5 shows the boundaries at a distance ε to create a margin between, for example, raw precipitation forecast data. Indeed, ε represents the quality of optimization and states deviation between observed gauge data and downscaled forecast data. With the help of risk minimization (Eq. 6), Eq. 5 is solved at a higher dimension for raw forecast data to make linear separation possible.

$$R\left(w,\zeta,\zeta^{*}\right)=\frac{1}{2}\|w\|^{2}+C\sum_{i=1}^{m}\left(\zeta_i+\zeta_i^{*}\right)$$

$$\text{subject to}: \begin{cases} y_i-(w\phi(x_i)+b_i)\le\varepsilon+\zeta_i \\ (w\phi(x_i)+b_i)-y_i\le\varepsilon+\zeta_i^{*} \\ \zeta_i^{*},\zeta_i\ge 0 \end{cases} \tag{6}$$

where, x and y are raw and downscaled forecast data, respectively. In the above equations, given the training sample set (x_i,y_i), x_i is the n-dimensional input vector for different variables mentioned in GMDH method and y_i is the desired 1-dimensional output at the point i, $x_i\in R^n$, $y_i\in R^n$. ζ,ζ^{*} are slack variables to measure errors for any variable that falls outside of ε in upside and downside, respectively. $b\in R^n$ and $w\in R^n$ are constants for showing deviation and weight factor, respectively, and C is a regularization constant for balancing error minimization and maximization. m is the size of raw forecast data, $f(x)$ is the modeled target that calculates downscaled forecast data based on the observed gauge data (Dodangeh et al., 2020). Using the Lagrange function method, Non-linear regression becomes Eq. 7, where Kernel function is used to transform data into higher dimension space (for more detail about Lagrange and Kernel transformations, see (Geng et al., 2020; Yaghoubi et al., 2019).

$$f(x)=\sum_{i=1}^{m}\left(\alpha_i+\alpha_i^{*}\right)k\left(x,x_i\right)+b \tag{7}$$

where $k(x,x_i)=(\phi(x),\phi(x_i))$ is Kernel function, a Gaussian type of Kernel equation is used in this study, and α_i,α_i^{*} are Lagrange coefficients. NSE function (Eq. 1) is used between downscaled forecast and observed gauge data to optimize model parameters, including C,ε, and Kernel function parameters.

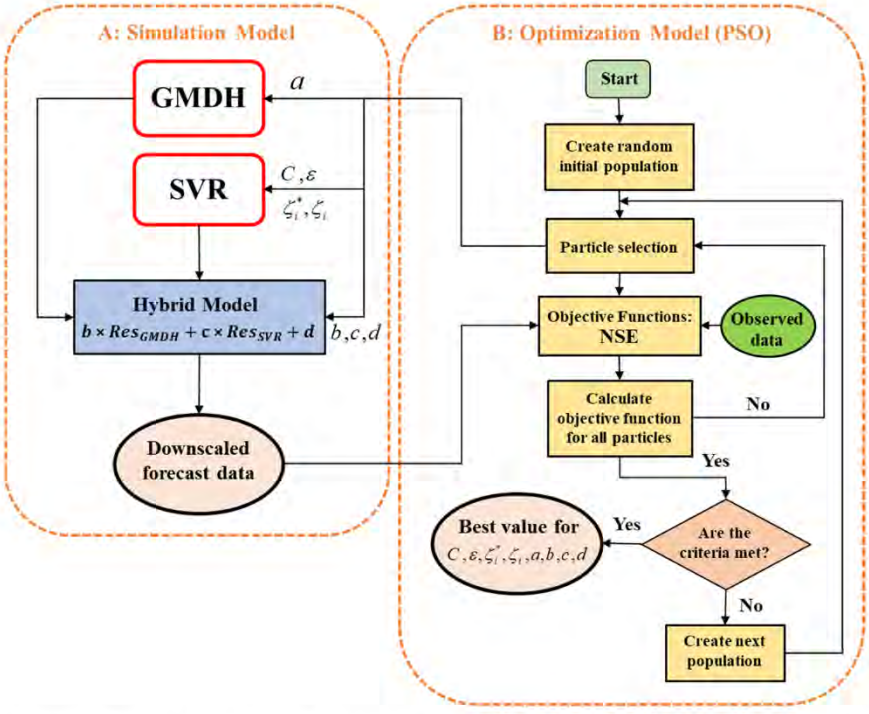


Fig. 6. A hybrid SVR-GMDH model for downscaling precipitation data. The parameters of hybrid model are optimized by PSO.

3.1.3.3. Hybrid AI downscaling method (SVR-GMDH). We also test a hybrid approach that combines GMDH and SVR (Fig. 6). Applying the GMDH and SVR forms an efficient downscaling model because of GMDH's capability in modeling highly non-linear relations, and SVR's independence on the dimensionality of the input space due to its model computational complexity (Dodangeh et al., 2020). In this case, raw precipitation forecast data are used to produce downscaled estimates with both GMDH and SVR, and the final hybrid downscaled values are then estimated as a linear combination of the GMDH and SVR outputs. Coefficients *b*, *c*, and *d* of the linear combination and coefficients of the individual methods are jointly optimized by Particle Swarm Optimization (PSO) with Eq. 1 as the objective function in order to increase the efficiency.

3.1.4. BBN downscaling method

BBNs were introduced in the late 1980s, and they have been used frequently in water resources planning and management (Govender et al., 2021; Phan et al., 2016; Xue et al., 2017). For instance, BBNs were applied to estimate missing rainfall data (Sun et al., 2017) and forecast flood peaks (Goodarzi et al., 2019). Khan and Coulibaly (2006) reported better performance of a BBN compared to an Artificial Neural Network for simulating daily river flow and reservoir inflow.

The BBN is a probabilistic graphical model consisting of nodes arranged in an acyclic directional graph. Nodes represent discrete or continuous random variables, and the direction of the arrows indicates whether or not the nodes are dependent on each other. The head node of the arrow is called the child node, and the start node is called the parent node so that the child node is dependent on the parent node. A node without parents and a node without children is called the root node and leaf node, respectively. If the nodes represent discrete random variables, then conditional probability tables (CPTs) are used to quantify the probabilistic influence of the parent nodes on the child node.

Values for the entries in the CPTs, where each entry represents the conditional probability of a child node state given its parent node states, are obtained by training or calibrating the BBN using observed gauge precipitation data from synoptic stations. Sensitivity analysis and cross-validation tests are used to validate the model. Validation tests are based on the Confusion matrix, the Accuracy criterion, and the Receiver Operating Characteristic (ROC) Curve (Fawcett, 2006).

The BBN for downscaling monthly precipitation forecasts is shown in Fig. 7. It consists of four nodes, i.e., one child node for downscaled precipitation and three parent nodes representing precipitation forecasts for the previous and current month, as well as the number of the current month. Downscaled precipitation and raw forecast precipitation were classified in two categories to perform probabilistic classification, shown in Table 2. Continuous forecast precipitation values can be obtained by multiplying the probability of each category into the mean of upper and lower limits of Table 2.

3.2. Hydrological rainfall-runoff model

The hydrological rainfall-runoff model was developed using two snow and soil moisture storages illustrated in Fig. 8. The parameters presented in Fig. 8 and their equations are shown in Table 3 and Table 4 for snow and soil moisture storages, respectively. Snow and soil moisture water balance equations are applied for each sub-basin on a daily scale. The input water balance component to the snow storage is precipitation, and its output water balance component is effective precipitation (P_e) (Table 3). In particular, effective precipitation (P_e) enters the soil moisture storage dependent on the air temperature and its comparison with freezing point (T_F) and melting point (T_M) (Fig. 8). The output components of soil moisture storage include actual evapotranspiration (ET_{ac}), irrigation ($Q_{ISW} + Q_{IGW}$), surface runoff (Q_{sur}), subsurface runoff (Q_{int}), and groundwater recharge (Q_r). As shown in Table 4, these output components are calculated by empirical functions based on relative soil water content (Z_1). The value of Z_1 varies between 0 (dry soil) and 1 (saturated soil).

The two-storage model contains six parameters: (1) freezing temperature of snow (T_F), (2) melting temperature of snow (T_M), (3) runoff resistance factor (RRF), (4) root zone conductivity (K), (5) soil water capacity (SWC), and (6) initial soil moisture in the root zone (Z_1). RRF controls surface runoff so that a significant value of RRF causes less surface runoff. This parameter depends on vegetation. Root zone conductivity (K) is a function of soil type, and initial soil moisture in the root zone (Z_1) is the percentage of occupied water ratio to the adequate total root zone storage at the beginning of the simulation.

We consider four sub-basins upstream of Bukan reservoir, corresponding to four rivers, namely ZRR upstream, Saruq river, Khoorkhooreh river, and Saqez river (Fig. 1b). Separate snow and soil moisture storages are considered for each sub-basin. There is no

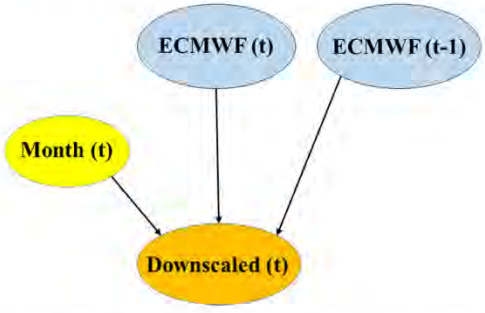


Fig. 7. The schematic of BBN for downscaling precipitation forecast.

Table 2
Classification of CPTs for each node in BBN for downscaling precipitation forecast.

Group	Name	Explanation	States	Time scale
Intervention variables	Month (t)	–	S1(Jan); S2(Feb); S3(Mar); S4(Apr); S5(May); S6(Jun); S7(Jul); S8(Aug); S9(Sep); S10(Oct); S11(Nov); S12(Dec)	Monthly
	ECMWF (t-1)	(mm)	S1(0–0.4); S2(0.4–2); S3(2–9.3); S4(9.3–42.2); S5(42.2–64.2); S6(64.2–85.9); S7 (85.9–114.2); S8(114.2–186.9)	
	ECMWF (t)	(mm)		
Objective variable	Downscaled (t)	(mm)	S1(0–1.3); S2(1.3–14.2); S3(14.2–38.9); S4(38.9–68.6); S5(68.6–189)	

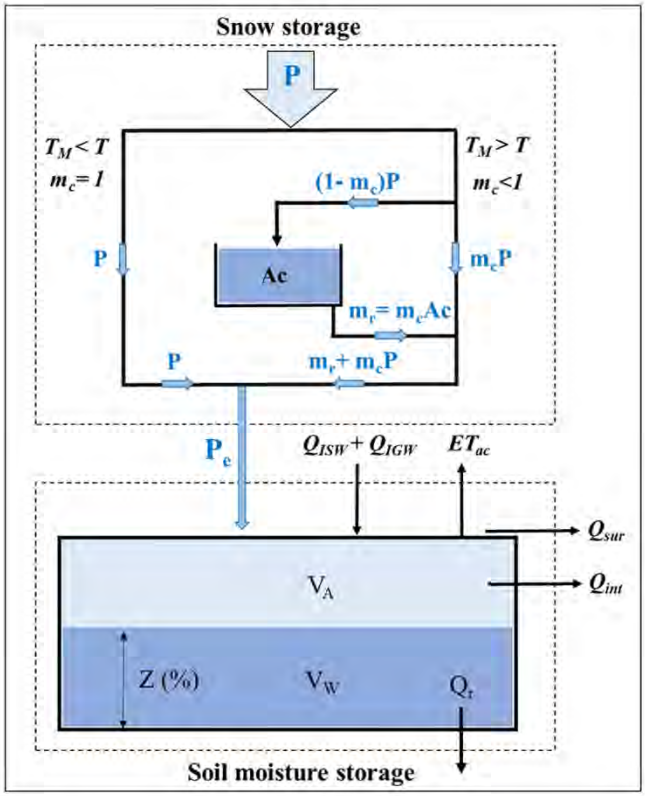


Fig. 8. Schematic diagram of the hydrological rainfall-runoff model consisting of snow and soil moisture storages. Variables are defined in Tables 3 and 4. Each water balance component is formulated on a daily scale.

Table 3
Snow storage water balance variables and equations, applied daily to each sub-basin (Fig. 1.b).

Variable	Dimension	Equation or data source
Effective precipitation	L/T	$P_e = m_cP + m_r$
Snowmelt coefficient	–	$m_c = \begin{cases} 0 & T < T_F \\ 1 & T > T_M \\ \frac{T - T_F}{T_M - T_F} & T_F < T < T_M \end{cases}$
Snow accumulation	L/T	$Ac_t = Ac_{t-1} + (1 - m_c)P_t$
Snow Melt	L/T	$m_r = m_cAc$
Total precipitation	L/T	P
Freezing point	°C	T_F: Calibration parameter
Melting Point	°C	T_M: Calibration parameter
Air (Observed) temperature	°C	T

Table 4
Soil moisture storage water balance variables and equations, applied daily to each sub-basin (Fig. 1.b) (Yates et al., 2005).

Variable	Dimension	Equation or data source
Storage change	L^3/T	$\frac{\Delta S_{rz}}{\Delta t} = SWC.A\frac{\Delta z}{\Delta t} = P_eA - ET_{act}A - Q_{sur} - Q_{int} - Q_r + Q_{ISW}$
Relative soil water content	–	$z_{t+1} = z_t + \Delta z$
Actual evapotranspiration	L/T	$ET_{ac} = (PET)K_c\left(\frac{5z1 - 2z^2}{3}\right)$
Surface runoff	L^3/T	$Q_{sur} = (Q_{ISW} + P_eA)z^{RRF}$
Subsurface runoff (Interflow)	L^3/T	$Q_{int} = fKz^2A$
Groundwater recharge	L^3/T	$Q_r = (1 - f)Kz^2A$
Irrigation with surface water	L^3/T	Q_{ISW}
Effective precipitation	L/T	P_e (This is calculated from Table 3)
Area for each sub-basin	L^2	A
Potential Evapotranspiration	L/T	PET: Penman-Monteith method using input meteorological data from synoptic weather station (Source: Iran Meteorological Organization, see https://data.irimo.ir/)
Crop coefficient	–	K_c
Runoff resistance factor	–	RRF: Calibration parameter
Preferred flow direction	–	$f = 1 \Rightarrow Q_r = 0$ and $Q_{int} = Kz^2A$ (There is no GW in the study area (Yekom Consulting Engineers, 2016))
Root zone conductivity	L/T	K: Calibration parameter
Soil water capacity	L	SWC: Calibration parameter
Initial Relative soil water content	–	Z_1: Calibration parameter

specific groundwater aquifer in the region due to the mountainous nature of the region (Yekom Consulting Engineers, 2016), and groundwater recharge is considered zero. The observed weather data, such as precipitation, temperature, relative humidity, wind speed, and sunshine hours, were obtained from the Meteorological Organization of Iran. The required land use information was extracted from Urmia Lake Restoration Program and Food and Agriculture Organization reports. Land use includes three rainfed, rangeland, and irrigated groups in the study area (Table 5).

3.3. Multiobjective calibration of hydrological rainfall-runoff model using MOPSO optimization algorithm

This study used two objective functions based on NSE (Eq. 8 and Eq. 9). The multiobjective perspective can reduce the uncertainty of determining the parameters of the hydrological model compared to the single-objective calibration process (Her and Seong, 2018; Schoups et al., 2005). Determining the objective functions is one of the essential steps in developing hydrological models. In some cases, due to the low diversity of observational data, lack of soil moisture data and groundwater level, and low discharge of rivers in the basin, a unique approach should be used to determine the objective functions. Approaches such as reverse flow (Pushpalatha et al., 2012) or logarithm of flow (Oudin et al., 2006; Roodari et al., 2020) are suitable for objective functions. For this reason, the approach of maximizing the simulation accuracy of observed flow and logarithm of flow were used as calibration objective functions in the present study. The calibration of the model was done using the observed temperature and precipitation input variables. Precipitation of upstream Bukan reservoir is very low from mid-spring to mid-autumn, as inflow to the reservoir is supplied by snowmelt during these dry months. To assess any potential trade-offs in the model in fitting low and high flows, we use two objective functions that capture these two flow regimes:

Table 5
Area and vegetation categories in each sub-basin of upstream Bukan reservoir (see the location of hydrometric stations in Fig. 1) (Yekom Consulting Engineers, 2016).

Sub-basin	Hydrometric station	Crop		Orchard		Rainfed	Rangeland
		Area [km^2]	No. categories	Area [km^2]	No. categories	Area [km^2]	Area [km^2]
Saruq river	**Safakhaneh (H1)**	37	4*	39.5	5**	312.5	2013.7
Khoorkhooreh river	**Darrepanbedan (H2)**	150.8	7***	37.5	5**	218.6	1622.7
ZRR upstream	**Polesaheb (H3)**					149.8	1224.1
Saqqez river	**Qabqabloo (H4)**					239.7	858.6
Total Area [km^2]		187.8		77		920.6	5719.1

* Crops include alfalfa, barley, vegetable, and wheat.
** Orchards include apple, conifer tree, grapes, stone fruit, and walnut.
*** Crops include alfalfa, barley, bean, sugarbeet, sunflower, vegetable, and wheat.