J. Agr. Sci. Tech. (2021) Vol. 23(6): 1395-1409

Application of Bayesian Model Averaging (BMA) Approach for Estimating Evapotranspiration in Gorganrood-Gharesoo Basin, Iran

A. Kazemi[1], N. Ghahreman[1*], M. Ghamghami[1], and A. Ghameshloo[1]

ABSTRACT

Accurate estimation of Evapotranspiration (ET), as a key component in the hydrological cycle, is essential in agricultural water management. In the current study, an approach based on the Bayesian Model Averaging (BMA) was used to combine eight ET empirical models, namely, Blaney-Criddle, Makkink, Penman, FAO-Penman-Monteith, Priestly-Taylor, Thornthwaite, Turc and Wang to improve the accuracy of ET estimations compared to individual models. The results of eight models and 247 combinations of them (without replacement) were compared to the results of the Water Balance (WB) model as the reference of comparison. This study was performed using warm season (April-September) data of 2005-2014 from Gorganrood-Gharesoo Basin, north of Iran. The performance of the eight models and all possible combinations were evaluated based on four statistical metrics i.e. Root Mean Square Error (RMSE), Kling-Gupta (KGE), Coefficient of Determination (R^2), and Bias. Then, the best-performing combination, (BMA-Best), was determined. Based on the WB method, the BMA-Best combination had better performance than the single models according to most of the metrics. In a few cases in which individual models showed slightly better performance than BMA-Best combination, the differences were not statistically significant. On average, the BMA-Best combination increased the R^2 by more than 50% and decreased RMSE by more than 70%. According to results of the current study, BMA provides a more reliable estimation of ET and it is recommended for use rather than the individual models. Moreover, the BMA-best combination mostly consisted of energy-based ET models, suggesting that these models have a better performance in climatic conditions of the study area.

Keyword: Bias, EM algorithm, Statistical analysis, Water balance.

INTRODUCTION

Evapotranspiration (ET) is one of the most important variables of natural ecosystems as it connects water, carbon, and land surface energy exchanges. Therefore, accurate ET estimates in different spatial scales is essential for understanding the interactions between the earth's surface and the atmosphere and for resource management (Chen *et al.*, 2015). Direct measurement of ET is difficult and time consuming; therefore, various methods have been developed so far to estimate the ET. The common available methods are generally site-specific with promising application in small scales, such as irrigation networks. In a large region like a basin, the accuracy of these approaches is significantly decreased, mainly due to lack of adequate meteorological data (Bastiaanssen *et al.*, 2000).

Evapotranspiration (ET) is one of the most important but difficult components of the hydrologic cycle to quantify accurately (Gao, 2010). ET depends on the amount of soil water, climatic elements, and type of

[1] Department of Irrigation and Reclamation Engineering, University College of Agriculture and Natural Resources, University of Tehran, 31587-77871 Karaj, Islamic Republic of Iran.
* Corresponding author; e-mail: nghahreman@ut.ac.ir

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

plant (Ma *et al*. 2020). The Water Balance (WB) method is one of the best methods for estimating ET at a regional scale such as basins and can be used as a measure for evaluating other models (Zhao *et al*., 2013). WB concept is an explanation of the mass conservation law or the equation of continuity (Karongo and Sharma, 1997; Pourmeidani *et al*. 2020). Spatio-temporal variations of climatic conditions and soil moisture make the ET estimation more complicated. Hence, in many parts of the world, efforts have been made to provide empirical relationships between the ET and hydrological/meteorological data (Chun, 1989). Moreover, reliable data on runoff and interception required for WB calculation is scarce in Iran; therefore, any attempt for development of simplified and accurate alternative methods for estimation of ET is quite essential.

The alternative methods of ET estimation e.g. Makkink, Priestly-Taylor, Turc, and FAO-Penman-Monteith, are receiving an increasing attention by the various researchers (see Zhao *et al.,* 2013). In hydrological models, these methods are classified into two categories: 1) Methods that estimate water surface evaporation, soil evaporation, and vegetable transpiration separately, and then integrates them to obtain the basin evapotranspiration depending on the land use pattern, and 2) Those that estimate Potential Evapotranspiration (PET) and convert it into ET by applying the Soil Moisture Extraction Function (Zhao *et al.,* 2013).

In recent decades, for dealing with problems that involve the uncertainty existing in estimating ET, various statistical approaches have been proposed, among which the Bayesian Model Averaging (BMA) approach has shown promising results in various studies (Olson *et al.,* 2016; Dong *et al.,* 2013). In this approach, a combination of several given models with different weights is used. The BMA gives prior weights to the models, then, a weighted average of these estimates is calculated (Hinn *et al.,* 2020). The BMA method also has the ability to quantify uncertainty of input data, model structure and parameters, and improves the accuracy of the model (Najafi *et al.,* 2011). In case of ET estimation, BMA can be applied to select the best combination of methods estimating it based on statistical criteria.

BMA-based techniques have been widely used in studies of climate change (Duan and Phillips 2010), improving the accuracy of hydrological forecasts (Duan *et al.,* 2007), weather forecasts (Raftery *et al.,* 2005), forest biomass (Li *et al.,* 2008) and economics (Fernandez *et al.,* 2001). Several studies have confirmed better performance of BMA compared to other multi-model ensembles (e.g. Ellison 2004; Raftery *et al.,* 2005). Sloughter *et al.* (2007) quantified the probability prediction of precipitation by applying BMA. They showed that this method can give better estimates of the probability of high-precipitation events than logistic regression function. Duan *et al.* (2007) studied the hydrological predictions of the multi-modal ensembles by using the BMA in several basins of United States. The results showed that the BMA method produces more reliable probabilistic predictions than other approaches. Wang *et al.* (2012) predicted seasonal precipitation by multivariate statistical models using BMA. The results showed an improved skill in precipitation predictions. Chen *et al.* (2015) used BMA to estimate the ET in different regions of China. The findings showed that BMA reduces the bias and RMSE of estimates. The summary of the mentioned studies confirms that BMA might be considered as a helpful tool for achieving accurate estimations of ET as an important component of regional water management plans.

A comprehensive literature review revealed that despite of acceptable skill of BMA approach in environmental studies, the relevant researches in Iran are limited. Hence, the current study aimed to: (1) Combine eight ET estimation models using BMA approach, (2) Compare findings of all combinations as well as single models with values obtained from WB equation, and (3) To propose a suitable combination of estimation models for the study area.

MATERIALS AND METHODS

Data Description

The Gorganrood-Gharesoo Basin is located in Golestan Province in north of Iran (Figure 1). Golestan has three different climates by Köppen-Geiger classification (BSk, Csa and BSh) and is dominated by BSk.

Four synoptic stations and two hydrometric stations with long-term reliable data were selected for this study. The monthly values of soil moisture and discharge during warm season (April-September) for the period 2005-2014 were obtained from the Iran Water Resources Company (Water Resources Atlas, 2009). Moreover, precipitation, air temperature, wind speed, total sunshine, and air relative humidity were obtained from the Iran Meteorological Organization. Table 1 provides geographical characteristic of the study stations. Long term reliable observed data of radiation was not available in this region. Therefore, radiation data and vegetation indices i.e. Normalized Difference Vegetation Index (NDVI), Leaf Area Index (LAI) and albedo coefficient were retrieved from NASA's earth observation data set (https://neo.sci.gsfc.nasa.gov). NDVI is a widely used remote sensing vegetation index that represents the vegetation health and cover status (Tucker, 1979). It varies between -1 to +1.

$$NDVI = NIR - RED/NIR + RED \quad (1)$$

Where, NIR and RED are the Infrared and Red wavelength reflections, respectively. (Ghamghami *et al*. 2020)

LAI is a dimensionless quantity that characterizes plant canopies. It is defined as the area of the green leaf per unit of ground

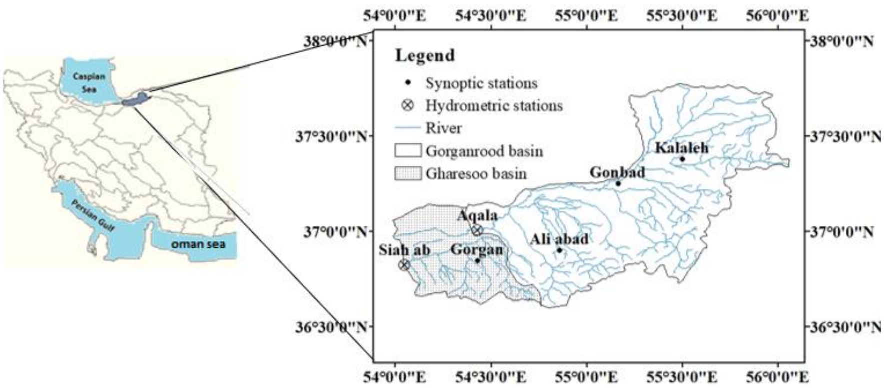

Figure 1. Location of selected synoptic and hydrometric stations in the study basin.

Table 1. Geographic characteristic of the study stations.

	Name	Type	Lat (°N)	Long (°E)	Elevation (m)
1	Gorgan	Synoptic	36.84	54.43	100
2	Aliabad	Synoptic	36.9	54.86	126
3	Kalaleh	Synoptic	37.38	55.5	149
4	Gonbad	Synoptic	37.25	55.16	39
5	Aqala	Hydrometric	37.01	54.43	-12
6	Siah Ab	Hydrometric	36.83	54.55	-26

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

area (LAI= Leaf area/Ground area, m^2 m^{-2}) in a broadband canopy (Bakhtiari *et al.*, 2020).

The Normalizing Process and Bias Elimination

Normality is an important assumption for many statistical techniques. For this purpose, data conversion functions can be used to normalize the distribution of abnormal data. One of the most important conversion functions is Box-Cox function (Box and Cox, 1964). Before applying the BMA, the pre-processing based on the Box-Cox converter was performed on original data to normalize values of the actual (obtained by water balance approach) and estimated ET values.

Moreover, bias affects hydrological variables such as ET, runoff, and snow melting. Hence, pre-processing techniques are necessary to eliminate the models' outputs bias prior to applying in impact assessment studies (Raneesh and Thampi, 2013). The procedure of bias elimination in the current study is based on linear transfer functions (Box and Cox 1964). The corrections were made for all months and stations to remove the temporal and spatial bias.

Reference Method

In the current study, Water Balance (WB) was considered as a reference method to evaluate the performance of other approaches. This method is based on the mass conservation law (European Commission, 2015). Therefore, the components of a WB equation could be precipitation, runoff, surface inflows, and underground inflows as input terms; and ET, soil water storage, surface outflows, and underground outflows as output terms. (Ibitoye *et al.* 2020). The geological characteristics of the study basin are such that the underground inflows (to basin) and outflows (from basin) are negligible. Such basins are known as "watertight basins" (Karongo and Sharma, 1997; Liu *et al.*, 2011; Chen *et al.*, 2015). The surface inflows and outflows were considered as base flow in the hydrograph curve. Therefore, the components of the WB are precipitation, runoff, ET, and soil water storage changes. Then, the ET can be estimated by the Equation (2):

$$ET = P - R \pm \Delta S \qquad (2)$$

Where, P, R, and ΔS represent rainfall, Runoff and Soil water storage changes during the time period Δt, respectively. Direct measurement of ΔS is difficult, but in the watertight basins there are certain time periods (depending on the crop growth stage), where ΔS is zero (Karongo and Sharma 1997; Liu *et al.*, 2011; Teuling *et al.*, 2009). In these periods, soil water content is eventually converted to the soil evaporation or plant transpiration. For these periods, the Equation (2) is modified as follows:

$$ET = P - R = WB \qquad (3)$$

The Equation (3) terms were calculated at a basin scale. For this purpose, using the data recorded at two hydrometric stations located in the basin outlet (Figure 1), the runoff volume of the basin during the study period was calculated by deduction of the base flow (surface inflows) in hydrograph curve. Then, the obtained volume was divided by area of the basin to estimate average runoff. The reason for using this approach, instead of Thiessen method, was lack of adequate hydrometric data (only two hydrometric stations with reliable data at basin outlet were available). Given the average depth of runoff and rainfall in each synoptic station, the ET was calculated. In other words, the synoptic stations were the representative points in which the methods were compared.

Methods of Estimating ET

Several methods, briefly described in Table 2 were used for estimating ET (Zhao

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

Table 2. Name and description of the eight selected ET models.[a]

Type	Reference	Required variables	Equation	
Energy based	Turc (1961)	Rs, Ta	$PET = 0.013\frac{T}{T+15}(R_s + 50), RH \rangle 50\%$	(4)
	Makkink (1957)	Ta, Z, Rs	$PET = \alpha\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda} - \beta, \alpha = 0.61, \beta = 0.12$	(5)
	Priestly-Taylor (1972)	Ta, Z, Rn	$PET = \alpha\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda}, \alpha = 1.26$	(6)
Temperature based	Thornthwaite (1948)	Ta	$PET = 0.533 L_a\left(\frac{10\overline{T_a}}{I_t}\right)^a$	(7)
	Blaney-Criddle (1950)	Ta, ρ	$PET = \frac{k\rho(0.46T + 8.13)}{30}$	(8)
Mass Transfer	Penman (1948)	U2, Ta, Td	$PET = 0.455\left(1 + \frac{0.98}{100}U_2\right)(e_s - e_a)$	(9)
Combined methods	Allen *et al.* (1998)	Ta, Z, Rn, G, Z, U2, Td	$PET = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T_a + 273}U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)}$	(10)
Empirical methods	Wang (2007)	NDVI, Rn, Ta	$ET = 0.035 R_n(a_0 + a_1 VI + a_2 T)$	(11)

[a] In the above equations, PET: The amount of Potential Evapotranspiration (mm d^{-1}); Rs: The short wave (MJ m^{-2} d^{-1}); Rn: The net Radiation (MJ m^{-2} d^{-1}); T_a: Mean air Temperature (°C); It is the total Index heat; a: Empirical constant; RH: Relative Humidity; G: Soil flux (MJ m^{-2} d^{-1}); γ: Psychrometric coefficient (kPa °C^{-1}); U_2: Wind speed at 2 meters height (m s^{-1}); e_s: Saturation vapor pressure at T_a (kPa); e_a: Vapor pressure at T_a (kPa); Δ: Slope of saturation vapor pressure curve (kPa °C^{-1}); k: Temperature correction factor; P%: Sunny hours per month to sunny hours in one year.

et al., 2013; Alexandris, 2008). The reasons for choosing these methods are: (1) Availability of required data, (2) Compatibility with the region's climate, and (3) Simplicity of mathematical calculations.

In the conceptual hydrological models, the ET is a function of Potential ET (PET) and available soil water. Soil moisture content is obtained by dividing Soil Moisture (SMT) by Soil field capacity Moisture Content (SMC) (Zhao *et al.*, 2013). The soil moisture extractions function is presented in equation 4. Kankash Omran Consulting Engineers. 2009.

$$\mathrm{ET} = \mathrm{PET}\left(\frac{\mathrm{SMT}}{\mathrm{SMC}}\right) \quad (4)$$

The corresponding monthly values were retrieved from Iran Water Resources Company bulletins.

Bayesian Model Averaging

Bayesian Model Averaging (BMA) is an approach to combine the forecast densities provided by different models and producing a new forecast Probability Density Function (PDF). It has been applied in various engineering problems such as hydrological modeling (Madadgar and Moradkhani, 2014). In this study, the BMA method was used to combine eight ET models (represented in Table 2) and achieve the best

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

combination for more accurate ET estimate. The BMA method includes a dependent variable y, the training data y_t and the sum of all predictions of the members X {x_1, x_2, x_3... x_k}. According to the law of total probability, PDF can be displayed as Equation (5):

$$p(y|x_1, x_2, x_3, \ldots, x_k) = \sum_{k=1}^{k} p(y|x_k).p(x_k|y_T) \quad (5)$$

In Equation (5) *y* and *K* represent ET variable and the number of ET estimation methods (i.e. 8), respectively.

$p(y|x_k)$ is the predictive PDF given by the simulation of x_k , p $(x_k|y_T)$ is the posterior probability of the model prediction x_k. y_T is the target data and T is the length of the data. In fact, p $(x_k|y_T)$ is a statistical weight (w_k). The magnitude of this weight indicates how much x_k agrees with y. The sum of the weights is equal to one, i.e. $\sum_{k=1}^{k} w_k = 1$.

$$p(y|x_1, x_2, \ldots, x_k) = \sum_{k=1}^{k} p\,(y|x_k).w_k \quad (6)$$

Before application of the BMA method, it seems reasonable to assume that p $(y|x_k)$ is a Gaussian distribution defined by mean (μ_k) and variance (σ_k^2).

$$p(y|x_k) = g(y|\theta_k) \quad (7)$$

Where, g refers to the Gaussian distribution and Θ_k = $\{\mu_k, \sigma_k, k = 1, \ldots, k\}$ is the parameter vector.

By combining Equations (6) and (7), the PDF of the probabilistic prediction of y in the BMA method can be expressed as Equation (8):

$$p(y|x_1, x_2, \ldots, x_k) = \sum_{k=1}^{k} g(y|\theta_k).w_k \quad (8)$$

The logarithm likelihood function was used to obtain both the weights w_k and the vector of the parameter Θ_k, where L is approximated using Equation (9).

$$l(\theta_1, \theta_2, \ldots, \theta_k) = \sum_{(t)} log \left[\sum_{k=1}^{k} g\,(y_t|\theta_k).w_k\right] \quad (9)$$

In this equation, Σ (t) is the total of ET and y_t is the target data at each point and time t. The BMA model calculates the weights (w_k) and parameter vector (Θ_k) while maximizing the logarithm of the likelihood function (Chen *et al.*, 2015).

EM Algorithm

Expectation–Maximization (EM) algorithm casts the maximum likelihood problem as a ‘‘missing data’’ problem (Chen *et al.*, 2015). The missing data may be considered as a latent variable that needs to be estimated. The EM algorithm alternates between the E (or Expectation) step and the M (or Maximization) step.

Statistical Analysis

To evaluate the performance of the model, four statistical criteria were used:

• The determination coefficient (R^2): It is used to analyze how differences in one variable can be explained by the difference in a second variable.

• The Root Mean Square Error (RMSE): is a frequently used measure of the differences between values predicted by a model or an estimator and the values observed.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \quad (10)$$

Where, S_i and O_i are Simulated by the model and Observed values, respectively.

• The relative bias represents the systematic bias of the simulation results.

$$\text{Bias} = \frac{\sum_{i=1}^{n}(S_i - O_i)}{\sum_{i=1}^{n} O_i} * 100\% \quad (11)$$

• The Kling-Gupta Efficiency index (KGE) was used to evaluate the overall performance of the model. The KGE is calculated as follows:

$$\text{KGE} = 1 - \text{ED} \quad (12)$$

$$\text{ED} = \sqrt{(\text{r} - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (13)$$

$$\alpha = \frac{\sigma_s}{\sigma_o} \quad (14)$$

$$\beta = \frac{\mu_s}{\mu_o} \quad (15)$$

In above equations, ED is the Euclidean Distance from the ideal point, r is the

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

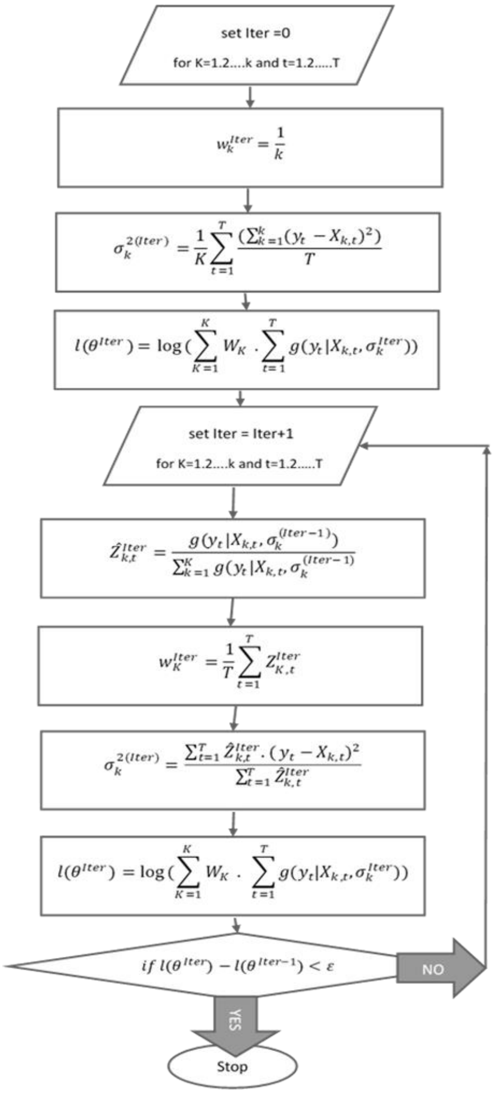

Figure 2. The flowchart of EM algorithm; Step 0: Initialize: Set iteration i= 0 and initially uniform weights, K is the total number of models.

correlation coefficient between simulations and observations, μ_o and σ_o are the mean and standard deviation of the observations, respectively, and μ_s and σ_s are the mean and standard deviations of the simulations, α is the percentage of changes in the simulated and observed values, and β is the ratio of the mean values of simulation to the observations. If there is no simulation error, the values of the three components r, α and β would be equal to 1. In current problem, the KGE is equal to 1.

Model Selection

In this study, two types of strategies were compared. In the first one, all eight models (BMA-All) were employed, and the second one used the best combination of models, hereafter denoted as (BMA-Best). There are 247 different combinations of the eight models (C_8^n , $n=1,8$), considering at least two models for each of the combinations. The BMA-Best was selected based on KGE

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

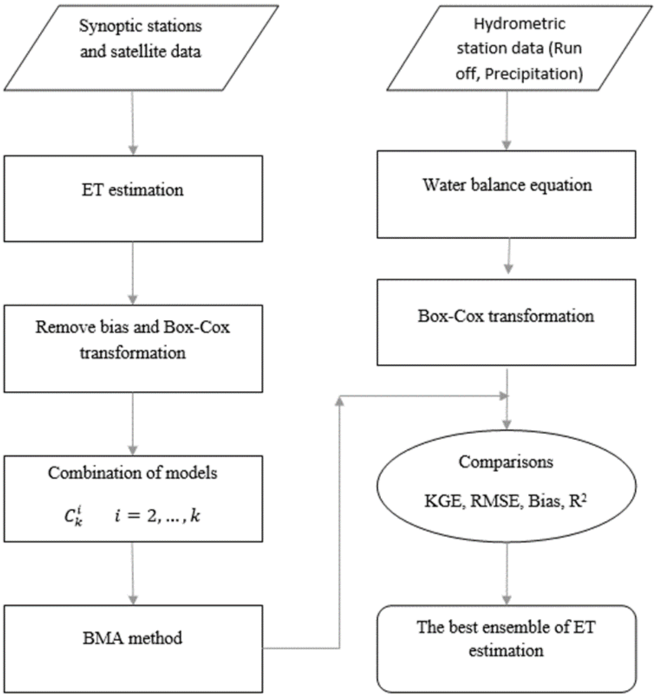

Figure 3. The flowchart of the applied algorithm.

values, which includes all their evaluation indices, namely, Bias, RMSE, and R^2 (Chen *et al.*, 2015). The so-called BMA-Best is awarded to a combination with the best KGE rather than a combination of the best single models. Figure 3 illustrates the flowchart of the applied algorithm in this study.

RESULTS

The results obtained from the BMA-Best combination are shown in Table 3. As seen in this table, combination of Makkink and Priestly-Taylor in Gorgan Station, Priestly-Taylor and Turc in Aliabad, Priestly-Taylor and FAO-Penman-Monteith at Kalaleh and FAO-Penman-Monteith and Turc at Gonbad Station had the lowest RMSE and the highest KGE compared to other 247 combinations. Hence, they were selected as the best combination at each station, or the BMA-Best. In terms of R^2, some combinations had the greater values than the BMA-Best, but no statistically significant difference.

In Table 4, the weights given to each model are shown for the Best-BMA and for each station. As it is clear, Blaney-Criddle, Penman, Thornthwaite, and Wang models were not selected in BMA-best combination. Therefore, they are not recommended for this basin. Moreover, Priestly-Taylor in Gorgan, Kalaleh and Aliabad Stations showed about 50% contribution in BMA-Best structure.

According to Table 4 and based on the weights given to each method, we can use a weighted average of Makkink and Priestly-Taylor models results in Gorgan, Priestly-Taylor and Turc in Aliabad, FAO-Penman-Monteith and Priestly-Taylor in Kalaleh, and the FAO-Penman-Monteith and Turc models in Gonbad Stations rather than individual models. Given that the weights are close to 0.5, the obtained average would be close to simple average.

Based on selected models in the BMA-Best combination, it can be concluded that,

Table 3. Evaluation indices of BMA-Best combination in 4 synoptic stations.

Station	Models	BIAS	RMSE	KGE	R2
Gorgan	Makkink Priestly-Taylor	-0.0145	0.5024	-0.0695	0.1136**
Aliabad	Priestly-Taylor Turc	0.0463	0.4419	-0.2182	0.0871*
Kalaleh	Priestly-Taylor FAO-Penman-Monteith	0.0467	0.4378	-0.1894	0.1084**
Gonbad	FAO-Penman-Monteith Turc	0.0141	0.4403	-0.2453	0.0486

Table 4. Weights given to the models in the BMA-Best combination.

	Weight in BMA-Best combination			
Model/Equation	Gorgan	Aliaba	Kalaleh	Gonbad
Blaney-Criddle	0	0	0	0
Makkink	0.51	0	0	0
Penman	0	0	0	0
FAO-Penman-Monteith	0	0	0.52	0.49
Priestly-Taylor	0.49	0.48	0.48	0
Thornthwaite	0	0	0	0
Turc	0	0.52	0	0.51
Wang	0	0	0	0

in this region, energy-based methods are superior compared to others.

The weights given to the combination of eight models (BMA-All) at each station are shown in Table 5. In four synoptic stations used for ET estimation, Wang model had the lowest weight, which might be attributed to its empirical nature.

Table 6 shows the results of two BMA approaches as well as individual models. According to Table 6, the RMSE values of the BMA-Best model at all stations is less than those obtained for individual models and the BMA-All. Moreover, the R^2 values for the BMA-Best model at all stations, except Aliabad, are greater than the R^2 values obtained for individual models and the BMA-All. However, difference between the greatest R^2 (Makkink) and that for BMA-Best in Aliabad was not statistically significant. Other criteria also indicated similar performance for BMA-best model compared to other models. Therefore, it can be concluded that, by combining these models at each station, better estimates of ET, i.e. closer agreement with the water balance approach, can be achieved. In other words, overall performance of the BMA-Best combination in each station would be more acceptable than others.

Figure 4 shows the monthly average ET values of the 8 models accompanied by the results of the BMA-Best and the BMA-All. It can be seen that the results of BMA-Best are strongly similar to those obtained from the Water Balance (WB) equation.

In this basin, the lowest and the highest ET, obtained from the best combination of models during the study period, are 28 and 36 mm.month^{-1}, respectively. These values were calculated by WB method equal to 28.7 and 32.7 mm month^{-1}. Figure 5 illustrates a reduced uncertainty when estimating ET for Gorgan Station during the study period. Top graph in this figure features variations range between two upper and lower models, i.e. the models for which the greatest (Wang model) and lowest (Turc

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

Table 5. Weights given to all eight models in the BMA-All combination.

Name	Weight in BMA-All			
	Gorgan	Aliabad	Kalaleh	Gonbad
Blaney-Criddle	0.1228	0.1307	0.1232	0.1193
Makkink	0.1507	0.1167	0.1124	0.1256
Penman	0.1416	0.1241	0.1456	0.1359
FAO-Penman-Monteith	0.1314	0.1517	0.1679	0.1684
Priestly-Taylor	0.1413	0.1391	0.1333	0.1337
Thornthwaite	0.1389	0.1280	0.1369	0.1276
Turc	0.1382	0.1527	0.1210	0.1539
Wang	0.0353	0.0570	0.0597	0.0355

Table 6. Results of evaluation of eight models and two BMA approaches at each station.

	model									
KGE	Blaney-Criddle	Makkink	penman	FAO-Penman-Monteith	Priestley-Taylor	Thornthwaite	Turc	Wang	BMA-All	BMA-best
Gorgan	-0.229	-0.118	-0.315	-0.429	**0.010**	-0.432	-0.495	-9.401	-0.166	-0.069
Aliabad	-0.342	-0.153	-0.227	-0.501	**-0.125**	-0.363	-0.294	-1.204	-0.329	-0.218
Kalaleh	-0.265	-0.127	-0.515	**-0.098**	-0.203	-0.394	-0.152	-1.162	-0.331	-0.189
Gonbad	-0.325	**-0.122**[a]	-0.458	-0.177	-0.209	-0.193	-0.320	-2.590	-0.344	-0.245
R^2										
Gorgan	0.000	0.054	0.024	0.057	0.070	0.014	0.051	0.005	0.016	**0.114**
Aliabad	0.075	**0.093**	0.087	0.093	0.052	0.022	0.058	0.046	0.151	0.087
Kalaleh	0.075	0.036	0.101	0.090	0.000	0.072	0.033	0.036	0.069	**0.108**
Gonbad	0.027	0.036	0.008	0.043	0.000	0.009	0.041	0.007	0.011	**0.049**
Bias										
Gorgan	0.511	0.026	0.041	-0.311	-0.056	-0.155	-0.190	6.993	0.223	**-0.014**
Aliabad	0.341	-0.220	-0.054	0.067	0.126	**0.006**	-0.029	1.191	0.104	0.046
Kalaleh	0.246	-0.179	0.273	-0.056	0.159	0.046	-0.107	1.201	0.127	**0.046**
Gonbad	0.092	-0.220	0.514	**0.008**	0.166	-0.183	0.021	2.261	0.137	0.014
RMSE										
Gorgan	0.801	0.512	0.512	0.636	0.533	0.542	0.556	8.830	0.580	**0.502**
Aliabad	0.669	0.668	0.522	0.448	0.470	0.496	0.442	1.662	0.504	**0.442**
Kalaleh	0.627	0.662	0.523	0.442	0.494	0.470	0.571	1.649	0.505	**0.438**
Gonbad	0.576	0.660	0.676	0.443	0.500	0.611	0.442	2.712	0.510	**0.440**

[a] The bolded values are the best findings.

model) values of ET were estimated, respectively. Bottom graph indicates uncertainty range at the 95% confidence estimated by the BMA-Best (gray section). In both graphs, observations (ET calculated by the WB model) are shown as red line. In addition, to keep inter-seasonal variability by the BMA-best, the uncertainties range was reduced. Findings confirmed that about 68, 71, 63, and 67% of the observed values fall into gray section for, respectively, Gorgan, Aliabad, Kalaleh, and Gonbad.

DISCUSSION

Accurate estimation of ET, especially at regional scales, is very important for improving land and water resource management, climate predictions, and drought monitoring. There are many

[DOR: 20.1001.1.16807073.2021.23.6.11.4]

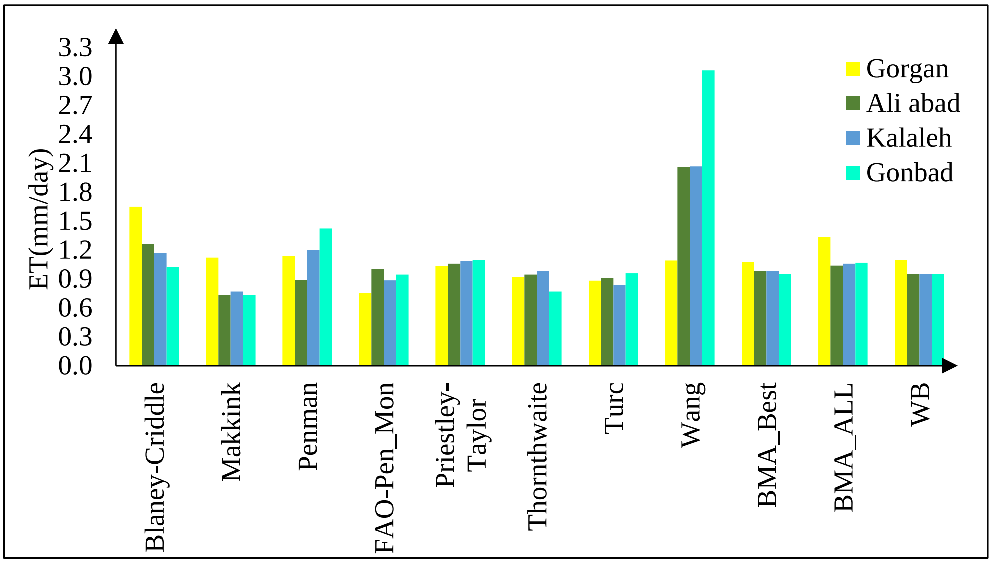

Figure 1. The average ET value obtained from eight individual models and two BMA approaches compared to the results of water balance approach.

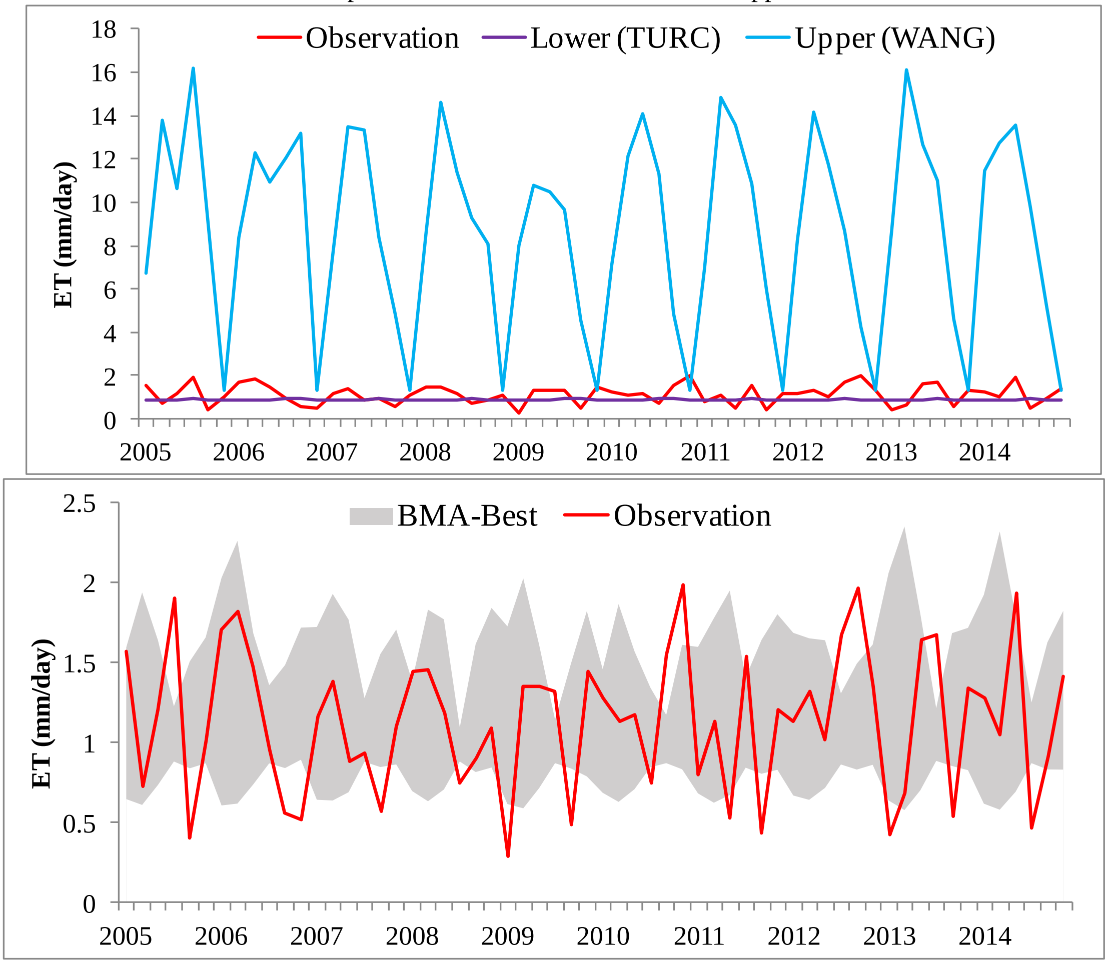

Figure 5. Uncertainties range associated with the application of individual models (top graph) and BMA technique (bottom graph) for Gorgan Station.

physical and empirical approaches for ET estimation. The uncertainty involved in using one of these methods instead of another i.e. statistical model uncertainty, is a major concern. It is quite important to find out how one can reduce the uncertainty among different methods. In this study, the skill of BMA method for combining single empirical models of estimation of ET was evaluated in Gorganrood-Gharesoo Basin, north of Iran. The results indicated that, generally, the BMA method improved the quality of ET probabilistic estimation compared to single empirical models. Besides, BMA showed a higher skill in reducing the uncertainty of estimations, which agrees with findings of Chen *et al.* (2015), Duan *et al.* (2007), and Raftery *et al.* (2005).

Assuming the study basin to be watertight, the results obtained from the simplified water balance model were considered as the actual ET and used for comparing empirical models. The simplified water balance used in the current study should be further evaluated in other regions by considering all involved terms for more scrutiny.

In all stations, the RMSE of the best combination, or the BMA-Best model, is lower than all single models and their combinations. Therefore, it can be concluded that combining these models at each station provide more accurate estimates of evapotranspiration. Moreover, the standard deviation and uncertainty range in BMA-Best is lower than other methods, which confirms its good performance in uncertainty reduction. Very similar results have been reported by Hao *et al.* (2019) and Sun *et al.* (2019).

CONCLUSIONS

According to findings of this case study, the best combination or The BMA-Best had the least error and, consequently, uncertainty than other single models. This proves that the BMA method may be recommended, as it provides higher prediction accuracy than an individual model. In addition to simplified water balance approach, which was selected as an evaluation metric, using lysimeteric or flux tower datasets, as a precise measurement of actual ET, might be also considered. The results of this study may be widely used in agricultural water management and planning.

REFERENCES

1. Alexandris, S., Stricevic, R. and Petkovic, S. 2008. Comparative Analysis of Reference Evapotranspiration from the Surface of Rainfed Grass in Central Serbia, Calculated by Six Empirical Methods against the Penman-Monteith Formula. *European Water*, **21(22):**17-28.
2. Allen, R. G., Pereira, L. S., Raes, D. and Smith, M. 1998. *Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements.* FAO Irrigation and Drainage Paper 56, Fao, Rome, **300(9):** D05109.
3. Bakhtiari, B., Ghahreman, N. and Afzali Goruh, Z. 2020. *Instruments and Methods of Observation in Grometeorology.* University of Tehran Press, 206 PP. (in Persian).
4. Bastiaanssen, W. G., Molden, D. J. and Makin, I. W. 2000. Remote Sensing for Irrigated Agriculture: Examples from Research and Possible Applications. *Agric. Water Manag.*, **46(2):** 137-155.
5. Blaney, H. F. and Criddle, W. D. 1950. Determining Water Requirements in Irrigated Areas from Climatological and Irrigation Data. Washington Soil Conservation Service, 48 PP.
6. Box, G. E. and Cox, D. R. 1964. An Analysis of Transformations. *Journal of the J. R. Stat. Soc. Series B Stat. Methodol.*, **26(2):** 211-252.
7. Chen, Y., Yuan, W., Xia, J., Fisher, J. B., Dong, W., Zhang, X. and Feng, J. 2015. Using Bayesian Model Averaging to Estimate Terrestrial Evapotranspiration in China. *J. Hydrol.*, **528**: 537-549..
8. Chun, Y. N. 1989. An Empirical Model for Estimating Evapotranspiration from Catchments. *IAHS Publ.*, **177:** 265-270.
9. Dong, L., Xiong, L. and Yu, K. 2013. Uncertainty Analysis of Multiple Hydrologic Models Using the Bayesian

Model Averaging Method. *J. Appl. Math.*, **2013(SI20):** 1-11.

10. Duan, Q. and Phillips, T. J. 2010. Bayesian Estimation of Local Signal and Noise in Multimodel Simulations of Climate Change. *J. Geophys. Res. Atmos.*, **115(D18).**
11. Duan, Q., Ajami, N. K., Gao, X. and Sorooshian, S. 2007. Multi-Model Ensemble Hydrologic Prediction Using Bayesian Model Averaging. *Adv. Water Resour.*, **30(5):** 1371-1386.
12. Ellison, A. M. 2004. Bayesian Inference in Ecology. *Ecol. Lett.*, **7(6):** 509-520.
13. European Commission. 2015. *Guidance Document on the Application of Water Balances for Supporting the Implementation of the WFD.* Final–Version 6.1– 18/05/2015, ISBN 978-92-79-52021-1, doi: 10.2779/352735
14. Fernandez, C., Ley, E. and Steel, M. F. 2001. Model Uncertainty in Cross Country Growth Regressions. *J. Appl. Econom.*, **16(5):** 563-576.
15. Gao, G. 2010. Changes of Evapotranspiration and Water Cycle in China during the Past Decades. PhD. Thesis, University of Gothenburg. http://hdl.handle.net/2077/21737
16. Ghamghami, M., Ghahreman, N., Irannejad, P., and Pezeshk, H. 2020. A parametric Empirical Bayes (PEB) Approach for Estimating Maize Progress Percentage at Field Scale. *Agric. For. Meteorol.* **281** (Feb): 107829. https://doi.org/10.1016/j.agrformet.2019.107829.
17. Hao, Y., Baik, J. and Choi, M. 2019. Combining Generalized Complementary Relationship Models with the Bayesian Model Averaging Method to Estimate Actual Evapotranspiration over China. *Agric. For. Meteorol.*, **279.**
18. Kankash Omran Consulting Engineers. 2009. The Study on Integrated Water Resources Management for Gorganrood-Gharesoo River Basin in the Islamic Republic of Iran, Water Balance in Gorgan Study.
19. Hinn, M., Gronau, Q. F.,van den Berg, D. and Wagenmakers E. 2020. A Conceptual Introduction to Bayesian Model Averaging. *Advances in Methods and Practices in Psychological Science,* **3(2)*:****200 –215*
20. Ibitoye, E., Adegun, I. K., Omoniyi, P. O., Ogedengbe, T. S., Alabi, O.O. 2020. Numerical Investigation of Thermo-physical Properties of Non-newtonian Fliud in a Modelled Intestine. *J. Biores. Bioprod.*, **5(3)**: 211–221.
21. Karongo, S. K. and Sharma, T. C. 1997. An Evaluation of Actual Evapotranspiration in Tropical East Africa. *Hydrol. Process.*, **11(5):** 501-510.
22. Li, Y., Andersen, H. E. and McGaughey, R. 2008. A Comparison of Statistical Methods for Estimating Forest Biomass from Light Detection and Ranging Data. *West. J. Appl. For.*, **23(4):** 223-231.
23. Liu, B., Hu, Q., Wang, W. P., Zeng, X. F. and Zhai, J. Q. 2011. Variation of Actual Evapotranspiration and Its Impact on Regional Water Resources in the Upper Reaches of the Yangtze River. *Quat. Int.*, **244(2):** 185-193.
24. Ma, H., Zhu, Q., Zhao, W. 2020. Soil Water Response to Precipitation in Different Micro-topographies on the Semi-arid Loess Plateau, China. *J. For. Res.*, 2020, **31(1)**: 245-256.
25. Madadgar, S. and Moradkhani, H. 2014. Improved Bayesian Multi-Modeling: Integration of Copulas and Bayesian Model Averaging. *Water Resour. Res.*, **50(12):** 9586-9603.
26. Makkink, G. F. 1957. Testing the Penman Formula by Means of Lysimeters. *Journal of the Institution of Water Engineers*, **11:** 277-288.
27. Najafi, M. R., Moradkhani, H. and Jung, I. W. 2011. Assessing the Uncertainties of Hydrologic Model Selection in Climate Change Impact Studies. *Hydrol. Process.*, **25(18):** 2814-2826.
28. Olson, R., Fan, Y. and Evans, J. P. 2016. A Simple Method for Bayesian Model Averaging of Regional Climate Model Projections: Application to Southeast Australian Temperatures. *Geophys. Res. Lett.*, **43:** 7661–7669.
29. Penman, H. L. 1948. Natural Evaporation from Open Water, Bare Soil and Grass. *Proc. Roy. Soc. London. Series A. Math. Phys. Sci.*, **193(1032):** 120-145.
30. Pourmeidani, A., Ghamghami, M., Olya, H., and Ghahreman, N. 2020. Determination of Suitable Regions for Cultivation of Three Medicinal Plants under a Changing Climate. *Environ. Processes*. **7**:89-108.
31. Priestley, C. H. B. and Taylor, R. J. 1972. On the Assessment of Surface Heat Flux and

Evaporation Using Large-Scale Parameters. *Mon. Weather Rev.*, **100(2):** 81-92.
32. Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. 2005. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Weather Rev.*, **133(5):** 1155-1174.
33. Raneesh, K. Y. and Thampi, S. G. 2013. Bias Correction for RCM Predictions of Precipitation and Temperature in the Chaliyar River Basin. *J. Climatol. Weather Forecasting*, **1(2):** 1-7.
34. Sloughter, J. M. L., Raftery, A. E., Gneiting, T. and Fraley, C. 2007. Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging. *Mon. Weather Rev.*, **135(9):** 3209-3220.
35. Sun, H., Yang, Y., Wu, R., Gui, D., Xue, J., Liu, Y. and Yan, D. 2019. Improving Estimation of Cropland Evapotranspiration by the Bayesian Model Averaging Method with Surface Energy Balance Models. *Atmosphere*, **10:** 188
36. Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L., Richardson, A.D., Wohlfahrt, G., Seneviratne, S.I. 2009. A Regional Perspective on Trends in Continental Evaporation. *Geophys. Res. Lett.*, **36(2).**
37. Tucker, C. J. 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.*, **8(2):** 127-150.
38. Turc, L. 1961. Estimation of Irrigation Water Requirements, Potential Evapotranspiration: A Simple Climatic Formula Evolved up to Date. *Ann. Agron.*, **12(1):** 13-49.
39. Thornthwaite, C. W. 1948. An Approach toward a Rational Classification of Climate. *Geogr. Rev.*, **38(1):** 55-94.
40. Wang, K., Wang, P., Li, Z., Cribb, M. and Sparrow, M. 2007. A Simple Method to Estimate Actual Evapotranspiration from a Combination of Net Radiation, Vegetation Index, and Temperature. *J. Geophys. Res. Atmos.*, **112(D15).**
41. Wang, Q. J., Schepen, A. and Robertson, D. E. 2012. Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging. *J. Clim.*, **25(16):** 5524-5537.
42. Water Resources Atlas. 2009. *Gorganrood-Gharesoo River Basin*. Ministry of Energy, Iran Water Resources Company, Tehran, Iran.
43. Zhao, L., Xia, J., Xu, C. Y., Wang, Z., Sobkowiak, L. and Long, C. 2013. Evapotranspiration Estimation Methods in Hydrological Models. *J. Geogr. Sci.*, **23(2):** 359-369. https://doi.org/10.1007/s11442-013-1015-9.

کاربست رهیافت میانگیری مدل بیزی در برآورد تبخیرتعرق در حوزه قره سو- گرگان رود

آ. کاظمی، ن. قهرمان، م. قمقامی، آ. قمشلو

چکیده

برآورد دقیق تبخیرتعرق بعنوان یک مولفه کلیدی در چرخه آبشناسی، در مدیریت آب کشاورزی اهمیت زیادی دارد. در مطالعه حاضر، رهیافتی مبتنی بر میانگین گیری بیزی برای ترکیب 8 مدل تجربی تبخیرتعرق شامل بلانی کریدل، ماکینک،پنمن، پنمن-مانتیث فائو، پریستلی تیلور، تورنت وایت، تورک و وانگ و مقایسه آن با حالت کاربرد منفرد هریک از مدلها بکار رفت تا مهارت این روش در تدقیق

براورد تبخیرتعرق ، ارزیابی شود. نتایج کاربست این 8 مدل و 247 ترکیب آنها(بدون جایگزینی) با مقدار تبخیرتعرق حاصل از روش بیلان آب (بعنوان معیار سنجش) مقایسه گردید. مطالعه با استفاده از داده های ماههای گرم سال در دوره 2005 تا 2014 حوزه آبریز گرگان رود-قره سو در استان گلستان انجام شد.عملکرد 8 مدل و تمام ترکیبات محتمل بر اساس 4 شاخص آماری شامل جذر میانگین مربعات خطا(RMSE)، کلینگ-گوپتا(KGE)، ضریب تعیین(R2) و اریبی ارزیابی و بر این اساس بهترین ترکیب (BMA-BEST) تعیین شد. بر مبنای مقدار حاصل از روش بیلان آب، در اکثر موارد، ترکیب بهینه عملکرد بهتری در قیاس با کاربرد منفرد هریک از مدلها داشت. در معدود حالاتی که تک مدلها عملکرد اندک بهتری داشتند، تفاوتها از نظر اماری معنی دار نبود. بطور متوسط ترکیب BMA-BEST، مقدار ضریب تعیین را به اندازه 50 درصد افزایش و RMSE را به میزان بیش از 70 درصد کاهش داد. بر اساس نتایج این پژوهش، رهیافت میانگین گیری مدل بیزی تخمینهای قابل اعتماد تری در قیاس با کاربرد منفرد مدلها بدست می دهد و بدین جهت استفاده از آن توصیه می گردد.بعلاوه ترکیب بهینه در اکثر موارد متشکل از معادلات مبتنی بر تابش بوده است که نشانگر عملکرد بهتر این مدلها در اقلیم منطقه مطالعاتی می باشد.