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

A. Kazemi¹, N. Ghahreman¹*, M. Ghamghami¹, and A. Ghameshloo¹

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 (R²), 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 R² 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
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.

\[ \text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} \]  

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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Lat (°N)</th>
<th>Long (°E)</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Gorgan</td>
<td>Synoptic</td>
<td>36.84</td>
<td>54.43</td>
<td>100</td>
</tr>
<tr>
<td>2 Aliabad</td>
<td>Synoptic</td>
<td>36.9</td>
<td>54.86</td>
<td>126</td>
</tr>
<tr>
<td>3 Kalaleh</td>
<td>Synoptic</td>
<td>37.38</td>
<td>55.5</td>
<td>149</td>
</tr>
<tr>
<td>4 Gonbad</td>
<td>Synoptic</td>
<td>37.25</td>
<td>55.16</td>
<td>39</td>
</tr>
<tr>
<td>5 Aqala</td>
<td>Hydrometric</td>
<td>37.01</td>
<td>54.43</td>
<td>-12</td>
</tr>
<tr>
<td>6 Siah Ab</td>
<td>Hydrometric</td>
<td>36.83</td>
<td>54.55</td>
<td>-26</td>
</tr>
</tbody>
</table>
area (LAI= Leaf area/Ground area, m² 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): 

\[ ET = P - R \pm \Delta S \]  

Where, P, R, and \( \Delta S \) represent rainfall, Runoff and Soil water storage changes during the time period \( \Delta t \), respectively. Direct measurement of \( \Delta S \) is difficult, but in the watertight basins there are certain time periods (depending on the crop growth stage), where \( \Delta S \) 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:

\[ ET = P - R = WB \]  

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
Table 2. Name and description of the eight selected ET models.\(^a\)

<table>
<thead>
<tr>
<th>Type</th>
<th>Reference</th>
<th>Required variables</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy based</td>
<td>Turc (1961)</td>
<td>Rs, Ta</td>
<td>( \text{PET} = 0.013 \frac{T}{T+15} (R_s + 50) \cdot RH \cdot 50% ) (4)</td>
</tr>
<tr>
<td></td>
<td>Makkink (1957)</td>
<td>Ta, Z, Rs</td>
<td>( \text{PET} = a \frac{R_s}{\Delta + \gamma} - \beta, a = 0.61, \beta = 0.12 ) (5)</td>
</tr>
<tr>
<td></td>
<td>Priestly-Taylor</td>
<td>Ta, Z, Rn</td>
<td>( \text{PET} = \alpha \frac{R_s}{\Delta + \gamma} \cdot \lambda, \alpha = 1.26 ) (6)</td>
</tr>
<tr>
<td>Temperature based</td>
<td>Thornthwaite</td>
<td>Ta</td>
<td>( \text{PET} = 0.533L_a \left( \frac{10T_a}{I_a} \right)^\gamma ) (7)</td>
</tr>
<tr>
<td></td>
<td>Blaney-Criddle</td>
<td>Ta, ( \rho )</td>
<td>( \text{PET} = k \phi \left( 0.46T_a + 8.13 \right) ) \div 30 (8)</td>
</tr>
<tr>
<td>Mass Transfer</td>
<td>Penman (1948)</td>
<td>U2, Ta, Td</td>
<td>( \text{PET} = 0.455 \left( 1 + \frac{0.98}{100} U_2 \right) (e_s - e_a) ) (9)</td>
</tr>
<tr>
<td>Combined methods</td>
<td>Allen et al. (1998)</td>
<td>Ta, Z, Rn, G, Z, U2, Td</td>
<td>( \text{PET} = \frac{0.40\Delta (R_s - G) + 900}{T_a + 273} U_2 (e_s - e_a) ) (10)</td>
</tr>
<tr>
<td>Empirical methods</td>
<td>Wang (2007)</td>
<td>NDVI, Rn, Ta</td>
<td>( \text{ET} = 0.035R_n (a_0 + a_1 VI + a_3 T) ) (11)</td>
</tr>
</tbody>
</table>

\(^a\) 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 (\( ^\circ \)C); It is the total Index heat; \( a \): Empirical constant; RH: Relative Humidity; G: Soil flux (MJ m\(^{-2}\) d\(^{-1}\)); \( \gamma \): Psychrometric coefficient (kPa \( ^\circ \)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); \( \Delta \): Slope of saturation vapor pressure curve (kPa \( ^\circ \)C\(^{-1}\)); \( k \): Temperature correction factor; P\%: Sunny hours per month to sunny hours in one year.

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

\[
p(y|x_1, x_2, x_3, \ldots, x_k) = \sum_{k=1}^{k} p(y|x_k). p(x_k|y_t)
\]

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

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

\[
p(y|x_k) = \sum_{k=1}^{k} p(y|x_k), w_k
\]

(6)

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

\[
p(y|x_k) = g(y|\theta_k)
\]

(7)

Where, \( g \) refers to the Gaussian distribution and \( \theta_k = (\mu_k, \sigma_k, k = 1, \ldots, k) \) 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):

\[
p(y|x_1, x_2, \ldots, x_k) = \sum_{k=1}^{k} g(y|\theta_k), w_k
\]

(8)

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

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

(9)

In this equation, \( \Sigma (t) \) 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 \( (\Theta_k) \) 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.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (S_t - O_t)^2}
\]

(10)

Where, \( S_t \) and \( O_t \) are Simulated by the model and Observed values, respectively.

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

\[
Bias = \frac{\sum_{t=1}^{n} (S_t - O_t) \times 100\%}{\sum_{t=1}^{n} O_t}
\]

(11)

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

\[
\text{KGE} = 1 - \text{ED}(12)
\]

\[
\text{ED} = \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}
\]

(13)

\[
\alpha = \frac{\sigma_s}{\sigma_o}
\]

(14)

\[
\beta = \frac{\mu_s}{\mu_o}
\]

(15)

In above equations, ED is the Euclidean Distance from the ideal point, \( r \) is the
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, $\mu_o$ and $\sigma_o$ are the mean and standard deviation of the observations, respectively, and $\mu_s$ and $\sigma_s$ are the mean and standard deviations of the simulations, $\alpha$ is the percentage of changes in the simulated and observed values, and $\beta$ is the ratio of the mean values of simulation to the observations. If there is no simulation error, the values of the three components $r$, $\alpha$ and $\beta$ 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 ($\binom{8}{n} = 1,8$), considering at least two models for each of the combinations. The BMA-Best was selected based on KGE
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,
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 R² values for the BMA-Best model at all stations, except Aliabad, are greater than the R² values obtained for individual models and the BMA-All. However, difference between the greatest R² (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

<table>
<thead>
<tr>
<th>Station</th>
<th>Models</th>
<th>BIAS</th>
<th>RMSE</th>
<th>KGE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorgan</td>
<td>Makkink/Taylor</td>
<td>-0.0145</td>
<td>0.5024</td>
<td>-0.0695</td>
<td>0.1136**</td>
</tr>
<tr>
<td>Aliabad</td>
<td>Priestly-Taylor/Turc</td>
<td>0.0463</td>
<td>0.4419</td>
<td>-0.2182</td>
<td>0.0871*</td>
</tr>
<tr>
<td>Kalaleh</td>
<td>Priestly-Taylor/FAO-Penman-Monteith</td>
<td>0.0467</td>
<td>0.4378</td>
<td>-0.1894</td>
<td>0.1084**</td>
</tr>
<tr>
<td>Gonbad</td>
<td>FAO-Penman-Monteith/Turc</td>
<td>0.0141</td>
<td>0.4403</td>
<td>-0.2453</td>
<td>0.0486</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model/Equation</th>
<th>Weight in BMA-Best combination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gorgan</td>
</tr>
<tr>
<td>Blaney-Criddle</td>
<td>0</td>
</tr>
<tr>
<td>Makkink</td>
<td>0.51</td>
</tr>
<tr>
<td>Penman</td>
<td>0</td>
</tr>
<tr>
<td>FAO-Penman-Monteith</td>
<td>0</td>
</tr>
<tr>
<td>Priestly-Taylor</td>
<td>0.49</td>
</tr>
<tr>
<td>Thornthwaite</td>
<td>0</td>
</tr>
<tr>
<td>Turc</td>
<td>0</td>
</tr>
<tr>
<td>Wang</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 5. Weights given to all eight models in the BMA-All combination.

<table>
<thead>
<tr>
<th>Name</th>
<th>Weight in BMA-All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gorgan</td>
</tr>
<tr>
<td>Blaney-Cridle</td>
<td>0.1228</td>
</tr>
<tr>
<td>Makkink</td>
<td>0.1507</td>
</tr>
<tr>
<td>Penman</td>
<td>0.1416</td>
</tr>
<tr>
<td>FAO-Penman-Monteith</td>
<td>0.1314</td>
</tr>
<tr>
<td>Priestly-Taylor</td>
<td>0.1413</td>
</tr>
<tr>
<td>Thornthwaite</td>
<td>0.1389</td>
</tr>
<tr>
<td>Turc</td>
<td>0.1382</td>
</tr>
<tr>
<td>Wang</td>
<td>0.0353</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>KGE</th>
<th>Blaney-Cridle</th>
<th>Makkink</th>
<th>Penman</th>
<th>FAO-Penman-Monteith</th>
<th>Priestly-Taylor</th>
<th>Thornthwaite</th>
<th>Turc</th>
<th>Wang</th>
<th>BMA-All</th>
<th>BMA-best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorgan</td>
<td>-0.229</td>
<td>-0.118</td>
<td>-0.315</td>
<td>-0.429</td>
<td>0.010</td>
<td>-0.432</td>
<td>-0.495</td>
<td>-0.401</td>
<td>-0.166</td>
<td>-0.069</td>
</tr>
<tr>
<td>Aliabad</td>
<td>-0.342</td>
<td>-0.153</td>
<td>-0.227</td>
<td>-0.501</td>
<td>-1.125</td>
<td>-0.363</td>
<td>-0.294</td>
<td>-1.204</td>
<td>-0.329</td>
<td>-0.218</td>
</tr>
<tr>
<td>Kalaleh</td>
<td>-0.265</td>
<td>-0.127</td>
<td>-0.515</td>
<td>-0.098</td>
<td>-0.203</td>
<td>-0.394</td>
<td>-0.152</td>
<td>-1.162</td>
<td>-0.331</td>
<td>-0.189</td>
</tr>
<tr>
<td>Gonbad</td>
<td>-0.325</td>
<td>-0.122a</td>
<td>-0.458</td>
<td>-0.177</td>
<td>-0.209</td>
<td>-0.193</td>
<td>-0.320</td>
<td>-2.590</td>
<td>-0.344</td>
<td>-0.245</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorgan</td>
<td>0.000</td>
<td>0.054</td>
<td>0.024</td>
<td>0.057</td>
<td>0.070</td>
<td>0.014</td>
<td>0.051</td>
<td>0.005</td>
<td>0.016</td>
<td>0.114</td>
</tr>
<tr>
<td>Aliabad</td>
<td>0.075</td>
<td>0.093</td>
<td>0.087</td>
<td>0.093</td>
<td>0.052</td>
<td>0.022</td>
<td>0.058</td>
<td>0.046</td>
<td>0.151</td>
<td>0.087</td>
</tr>
<tr>
<td>Kalaleh</td>
<td>0.075</td>
<td>0.036</td>
<td>0.101</td>
<td>0.090</td>
<td>0.000</td>
<td>0.072</td>
<td>0.033</td>
<td>0.036</td>
<td>0.069</td>
<td>0.108</td>
</tr>
<tr>
<td>Gonbad</td>
<td>0.027</td>
<td>0.036</td>
<td>0.008</td>
<td>0.043</td>
<td>0.000</td>
<td>0.009</td>
<td>0.041</td>
<td>0.007</td>
<td>0.011</td>
<td>0.049</td>
</tr>
<tr>
<td>Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorgan</td>
<td>0.511</td>
<td>0.026</td>
<td>0.041</td>
<td>-0.311</td>
<td>-0.056</td>
<td>-0.155</td>
<td>-0.190</td>
<td>6.993</td>
<td>0.223</td>
<td>-0.014</td>
</tr>
<tr>
<td>Aliabad</td>
<td>0.341</td>
<td>-0.220</td>
<td>-0.054</td>
<td>0.067</td>
<td>0.126</td>
<td>0.006</td>
<td>-0.029</td>
<td>1.191</td>
<td>0.104</td>
<td>0.046</td>
</tr>
<tr>
<td>Kalaleh</td>
<td>0.246</td>
<td>-0.179</td>
<td>0.273</td>
<td>-0.056</td>
<td>0.159</td>
<td>0.046</td>
<td>-0.107</td>
<td>1.201</td>
<td>0.127</td>
<td>0.046</td>
</tr>
<tr>
<td>Gonbad</td>
<td>0.092</td>
<td>-0.220</td>
<td>0.514</td>
<td>0.008</td>
<td>0.166</td>
<td>-0.183</td>
<td>0.021</td>
<td>2.261</td>
<td>0.137</td>
<td>0.014</td>
</tr>
</tbody>
</table>

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    |

* The bolded values are the best findings.

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...
Bayesian Model Averaging and Evapotranspiration

**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 lysimetric 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

Bayesian Model Averaging and Evapotranspiration


کاربرد رهیافت میانگیری مدل‌ها در برآورد تبخیرتعرق در حوزه قره سو

گرگان رود

آ. کاظمی، ن. قهرمان، م. فقامی، آ. فشملو

چکیده

برآورده دقیق تبخیرتعرق بعنوان یک مولفه کلیدی در چرخه آبیاری، در مدیریت آب کشاورزی اهمیت زیادی دارد. در مطالعه حاضر، رهیافتی مبتنی بر میانگیری مدل‌ها به تخمین تبخیرتعرق شامل بلانی کریدل، ماکینگ، پنمن، پنمن-فائو، پریستلی تیلور، تورنت، تورنت وایت، تورنت وایت، تورنت وایت و وانگ و مقایسه آن با حالت کاربرد منفرد هریک از مدل‌ها با کار رفت تا مهارت این روش تدقیق
براورد تبخیرتعرق، نتایج کاربست این 8 مدل و 247 ترکیب آنها(بدون جایگزینی) با مقدار تبخیرتعرق حاصل از روش بیلان آب (بعنوان میانگین سنجش) مقایسه گردید. مطالعه با استفاده از داده‌های ماههای گرم سال در دوره 2005 تا 2014 حوزه آبریز گرگان رودخانه می‌تواند بر اساس گلستان انجام شد. عواملی که مدل‌های 8 مدل و تمام ترکیبات محتمل بر اساس 4 شاخص آماری شامل جذر میانگین مربعات خطا (RMSE)، ضریب تی وین (KGE)، و ارتباط ارزیابی و بر این اساس بهترین ترکیب (BMA-BEST) تعیین شد. بر مبنای مقدار حاصل از روش بیلان آب، در اکثر مواد و ترکیب بهترین عملکرد بهتری در قیاس با کاربرد منفرد هریک از مدل‌ها داشت. در محدود حالاتی که تک مدل‌ها عملکرد اندازه‌گیری بهتری داشتند، نتایج از نظر اماری معنی‌دار نبود. بطور متوسط ترکیب BMA-BEST، مقدار ضریب تی وین را به اندازه 50 درصد افزایش و RMSE را به میزان بیش از 70 درصد کاهش داد. بر اساس نتایج این پژوهش، روش میانگین گری مدل بیزی تخمینهای قابل اعتماد تری در قیاس با کاربرد منفرد مواد و جهت استفاده از آن توصیه می‌گردد. برای اطمینان از این مدل‌ها در اقلیم منطقه مطالعاتی می‌باشد.