Modeling Earth Systems and Environment
https://doi.org/10.1007/s40808-019-00647-2

ORIGINAL ARTICLE

Can artificial neural networks estimate potential evapotranspiration in Peruvian highlands?

Wilber Laqui[1,2] · Ricardo Zubieta[1,3] · Pedro Rau[4] · Abel Mejía[1] · Waldo Lavado[1,5] · Eusebio Ingol[1]

Received: 29 June 2019 / Accepted: 10 September 2019

Abstract

Evapotranspiration (ETo) is one of the most important variables of the water cycle when water requirements for irrigation, water resource planning or hydrological applications are analyzed. In this context, models based on artificial neural networks (ANN) of the retro-propagation type can be an alternative method to estimate ETo in highland regions using a number of input variables limited. The objective of this study is to develop ANN models to estimate ETo for the Peruvian highlands using input variables such as maximum air temperature (T_{max}), minimum air temperature (T_{min}), hours of sunshine (Sh), relative humidity (Rh) and wind speed (Wv), as an alternative method to FAO Penman–Monteith method (FAO-PM56) and Hargreaves–Samani (HS). Daily climatic datasets recorded at 12 meteorological stations between 1963 and 2015 were selected in this study. For evaluation reason, the ETo calculated using the FAO-PM56 was also considered. The main input variable to ANN modeling is T_{max}, followed by Sh and Wv or combinations between them. Hargreaves–Samani (HS) showed a poor performance in the estimation of the ETo in the Peruvian highlands compared to the 13 ANN models. Additionally, it was determined that in stations with lower thermal amplitude (< 14.2 °C) the lowest performance levels are presented in the estimation of the ETo with HS equation, which does not occur markedly with the ANN models that they estimate adequately ETo. Therefore, ANN models represent a great option to replace the FAO-PM56 and HS method, when ETo data series are scarce.

Keywords Highlands · Artificial intelligence · ETo · Water requirement

Introduction

Evapotranspiration (ETo) is one of the main variables of the hydrological cycle affecting water requirements and depends on several climatic variables such as temperature, humidity, solar radiation, wind speed, among others. A better estimation of ETo is relevant in water resources studies such as estimation of water requirements of crops, rainfall-runoff modeling, evaluation of soil capacity and water balances among others (Adeloye et al. 2012; Antonopoulos and Antonopoulos 2017). Moreover, ETo estimations are also important to simulate stream flows (Laqui 2010; Lujano et al. 2015; Zubieta et al. 2018; Rau et al. 2017) and to evaluate the impact of land use changes on the hydrological responses (Rahimi Khoob 2008; Jain et al. 2008).

The Peruvian Altiplano constitutes an especial region due to extreme climates, high elevation (over 3820 m a.s.l.) and limited weather information, making it difficult to accurately estimate ETo. Garcia et al. (2004) evaluated the performance of several equations to calculate the ETo in the Bolivian

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40808-019-00647-2) contains supplementary material, which is available to authorized users.

✉ Wilber Laqui
20150840@lamolina.edu.pe

1 UNALM, Programa de Doctorado en Recursos Hídricos, Universidad Nacional Agraria La Molina, Lima, Peru

2 Escuela Profesional de Ingeniería Agrícola, Universidad Nacional del Altiplano, Puno, Peru

3 Instituto Geofísico del Perú, IGP, Subdirección de Ciencias de la Atmósfera e Hidrósfera, Lima, Peru

4 UTEC, Departamento de Ingeniería Ambiental, Universidad de Ingeniería y Tecnología, Lima, Peru

5 Servicio Nacional de Meteorología e Hidrología del Perú, Lima, Peru

highlands, suggesting that the use of the Thornthwaite (Tw) formula should be avoided, while the Hargreaves–Samani (HS) formula adequately estimates the ETo only in the north of the highlands (more humid) and that The FAO Penman–Monteith (FAO-PM56) equation gives good results in all the locations evaluated. Similarly, Chang et al. (2017) in the Tibetan highlands (over 3033) evaluated four representative methods, FAO-PM56, Priestley–Taylor (PT), HS, and Mahringer (MG). The results demonstrate that PT is the recommended method for estimating ETo in the Tibetan highlands, PM produces good results when sufficient information is available and HS proved to be a complementary method when information is scarce, while MG is an unsuitable method for The Tibetan highlands. Both studies show the limitations of some methods used in the estimation of ETo in the highlands; therefore, the question arises, what is the method that adequately estimates ETo in the Peruvian highlands under cold climates, high altitude and limited climatic data conditions?

The estimation of ETo depends on another factors: one of them is the availability of meteorological data (Zanetti et al. 2007). Therefore, many methods to estimate the ETo have been developed searching low cost alternatives as empirical and physical mathematical models (Huo et al. 2012; Yassin et al. 2016; Cobaner 2011). The Food and Agriculture Organization of the United Nations (FAO) and the World Meteorological Organization (WMO) recommend that the Penman–Monteith model (PM) as the standard method for calculating ETo (Citakoglu et al. 2014). The recent version of this method is known as the FAO-PM56 model (Allen 2000). However, the number of climatic variables required for its estimation is not always available. This is especially true in developing countries where reliable collection of wind speed, humidity, and radiation is limited or the number of the observation stations is usually very limited (Droogers and Allen 2002; Chang et al. 2017).

To decrease the uncertainty obtained from data scarcity in certain locations, an alternative method, which uses few weather inputs, is necessary. In last decade, computational models for estimating ETo such as artificial neural network (ANN) technique have been developed (Adeloye et al. 2012). This is a no lineal statistical method that can be used to problems that usually are not feasible to solve by mathematical, statistical or any conventional method (Laaboudi et al. 2012). This technique has been successful to model complex relationships in data series generally related to various areas of knowledge (Zanetti et al. 2007; Traore et al. 2010). Further, the models show a good generalization capacity when tested in locations that were not included in the training (Adamala 2018).

Recent studies have shown that ANN models can be useful to estimate ETo around different regions in the world, providing new information on the efficient performance of ANN models and their use in the planning of irrigation projects and water resources management in different environments in the absence of the appropriate meteorological variables for FAO-PM56 method (Antonopoulos and Antonopoulos 2017; Falamarzi et al. 2014; Laaboudi et al. 2012; Tabari and Hosseinzadeh Talaee 2013; Yassin et al. 2016). Nonetheless, its applicability has not been studied in the Altiplano of the Peruvian Andes, which presents cold climate and high altitude (Zolá and Bengtsson 2006; Pillco Zolá et al. 2018). Indeed, in this region the meteorological and hydrological can be scarce and even nonexistent, causing uncertainty about spatial a temporal distribution of the ETo. To evaluate the advantages of ANN models is necessary to investigate fully its applicability in the ETo estimation. The objectives of this study are: (i) to develop ANN models to estimate daily ETo in highland areas, as an alternative method in the absence of climatic data and limitations of representative methods for estimating ETo; (ii) to evaluate the performance of the ANN models by comparing to the ETo estimations from FAO-PM56 and HS method, and (iii) to identify the most influential parameters for estimating ETo.

Study area

The study area is located in the northern Titicaca Lake basin (TLB), which is located in the Peruvian/Bolivian Altiplano, (14°–17°S; 69°–71°W) (Fig. 1). The Altiplano is the largest and highest endorheic basin of the world, occupying the western part of Bolivia and the southeastern part of Peru, where the climate is classified as cold semi-arid, and characterized by having two well-marked seasons, a wet season during November–March period and a dry season during April–October period (Zolá and Bengtsson 2006). The annual cycle of dry and wet conditions of rainfall is associated with the seasonal expansion of the equatorial easterlies in the upper troposphere; meanwhile, interannual variability is predominantly related to changes in the mean zonal flow over the Altiplano (Garreaud et al. 2003). The Titicaca region is more humid than southern parts of the Altiplano (Pillco Zolá et al. 2018). The variations of climate have high influence on human activities as agriculture and water storage (Binford et al. 1997; Núñez et al. 2002). Thus, the Altiplano region is highly susceptible to present drought periods, since the rainfall during wet season is considered the major source of water to human consumption, agriculture, streamflow, and the recharge of the underground aquifers (García et al. 2003, 2007; Garreaud et al. 2009). Due to the high altitude of the Altiplano (between 3600 and 4000 m a.s.l.) and its location in the tropics, radiation strongly influences evapotranspiration (lower ETo), as a consequence of

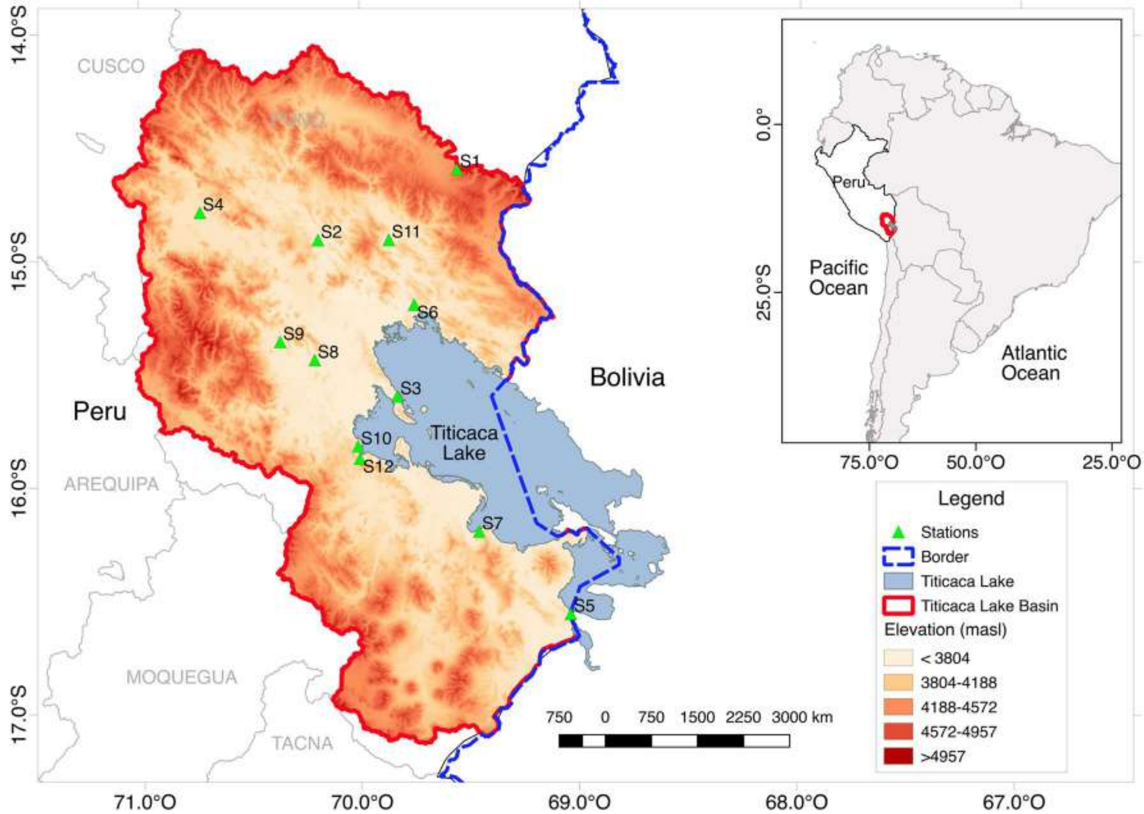

Fig. 1 Location of the study area. Green triangles indicate the location of weather stations used in the study

the reduced radiation and aerodynamic total energy in the Altiplano (García et al. 2004).

The Altiplano is one of the highest agricultural areas in the world. Due to the low levels of rainfall, high evapotranspiration rate and soils with low water retention capacity, water stress is a major constraint to crop production. Under these conditions, irrigation would be an asset to reduce the increased drought risk for agriculture (Garcia et al. 2004). Agriculture and cattle-raising activities, both focused on food production, are the main economic activities. Drought, floods and frost events are also significant factors (UNESCO, 2003).

Datasets and methodology

Datasets used

We use daily data series of 12 meteorological stations with data necessary to use FAO-PM56 method (Fig. 1). These data series are: maximum and minimum temperature (T_{max}, T_{min}) (°C), *sunshine hours* (Sh) (h), wind speed (Wv) (m s^{-1}) and relative humidity (Rh) (%). The data period for the stations has duration of 52 years between 1963 and 2015; only the days with complete records were used, so it was not necessary to complete the missing data. Likewise, data with grosser errors were excluded. Table 1 list their main characteristics, which include altitude, average value for T_{max}, T_{min}, Sh, Wv and Rh. In the last column shows the Aridity Index (AI), which is the ratio of potential evapotranspiration (ETo) and precipitation (P) (Huang et al. 2016). The S1 and S7 stations are in humid climate, S2 is in semi-arid climate and the others stations in dry sub-humid climate.

$$\mathrm{AI} = \frac{P}{\mathrm{ETo}}, \tag{1}$$

where P is the total annual precipitation (mm year^{-1}) and ETo is the total annual reference evapotranspiration calculated by the FAO-PM56 method (mm year^{-1}).

Moreover, ETo estimated using FAO-PM56 method was considered as the output target variable of the ANN models, as well as for the performance evaluation of the ANN models. This is usually considered in several studies (Antonopoulos and Antonopoulos 2017; Tabari and Hosseinzadeh Talaee 2013; Feng et al. 2017; Shiri 2017; Yassin et al. 2016; Kumar et al. 2008; Zanetti et al. 2007; KIŞI 2006; Traore et al. 2010), because it is the reference and standard equation to evaluate the results of mathematical models, as well as to provide results along the different climatic regions (Yassin et al. 2016), even Altiplano

Table 1 Characteristics of meteorological stations in the study area. Daily maximum temperature (T_{max}), daily minimum temperature (T_{min}), sunshine hours (Sh), daily relative humidity average (Rh), wind speed (Wv), evapotranspiration (ETo), total annual precipitation (P) and aridity index (AI)

Station	Lat (°S)	Long (°O)	Alt (m asl)	Period used	T_{max} (°C)	T_{min} (°C)	Ta (°C)	Sh (h)	Hr (%)	Wv (m s^{-1})	ETo (mm dia^{-1})	P (mm)	AI	CC
Ananea (S1)	14.68	69.53	4660	2002–2014	10.4	– 2.5	12.9	6.4	83.3	2.0	2.6	634.6	0.67	H
Azangaro (S2)	14.91	70.19	3830	2013–2015	17.0	0.6	16.4	7.7	66.2	3.1	3.6	597.5	0.45	S-a
Capachica (S3)	15.61	69.85	3819	1965–1969	16.0	1.1	14.9	8.4	68.8	2.2	3.4	776.1	0.63	DS-h
Chuquibambilla (S4)	14.80	70.73	3910	2005–2014	16.5	– 3.6	20.1	7.7	60.6	2.6	3.5	714.1	0.56	DS-h
Desaguadero (S5)	16.58	69.03	3860	1991–2014	15.0	1.0	14.0	7.8	62.8	5.2	3.7	731.7	0.54	DS-h
Huancane (S6)	15.20	69.76	3860	1990–2014	15.5	– 0.3	15.8	8.3	59.4	2.8	3.5	675.6	0.53	DS-h
Juli (S7)	16.20	69.46	3825	1990–2014	13.9	2.4	11.5	8.6	63.2	2.2	3.3	885.5	0.74	H
Juliaca (S8)	15.47	70.17	3826	2002–2013	17.8	– 1.2	19.0	8.3	74.9	1.4	3.3	620.4	0.52	DS-h
Lampa (S9)	15.36	70.37	3900	1994–2001	16.6	– 0.9	17.5	8.8	60.5	1.3	3.3	713.7	0.59	DS-h
Puno (S10)	15.83	70.02	3840	1964–2015	14.7	2.1	12.6	8.9	55.2	2.8	3.6	734.2	0.56	DS-h
Putina (S11)	14.91	69.87	3878	2006–2014	17.5	0.0	17.5	7.0	69.7	2.3	3.3	676.4	0.56	DS-h
Salcedo (S12)	15.88	70.00	3825	1963–1972	15.8	1.6	14.2	7.4	53.2	4.2	3.9	734.2	0.52	DS-h

The text in the bracket denotes the code of the climatic stations. *Lat* latitude, *Long* longitude, *Alt* altitude, T_{max} daily maximum temperature, T_{min} daily minimum temperature, *Ta* thermal amplitude, *Sh* sunshine hours, *Rh* daily relative humidity average, *Wv* wind speed, *ETo* evapotranspiration, *P* total annual precipitation and *AI* aridity index, *CC* climate classification, *H* humid, *S-a* semi-arid, *DS-h* dry semi-humid

conditions (Garcia et al. 2004). The PM-FAO56 method (Allen 2000) is summarized in the following equation:

$$\text{ETo FAO-PM56} = \frac{0.408\Delta\left(R_n G\right) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma\left(1 + 0.34u_2\right)}, \tag{2}$$

where ETo FAO-PM56 is the reference evapotranspiration calculated by the PM-FAO method56 (mm día^{-1}), Δ is the slope of the vapor pressure curve (kPa °C^{-1}), Rn is the net radiation on the crop surface (MJ m^{-2} día^{-1}), G is the soil heat flux density (MJ m^{-2} día^{-1}); γ = psychrometric constant (kPa °C^{-1}), T is the average daily air temperature (°C), u_2 is wind speed at 2 m height (m s^{-1}), e_s is the vapor saturation pressure (kPa), and e_a is the current vapor pressure (kPa).

And Hargreaves–Samani (HS) Eq. (1982, 1985):

$$\text{ETo HS} = 0.00135(T + 17.8)R_s, \tag{3}$$

$$R_s = k_{R_s} \sqrt{T_{max} - T_{min}} R_a, \tag{4}$$

where ETo HS is the reference evapotranspiration calculated by the Hargreaves–Samani method, T is mean daily air temperature (°C), R_s is the solar radiation (MJ m^{-2} per day), R_a is the extraterrestrial radiation (MJm^{-2} per day), T_{max} is the maximum daily air temperature (°C), T_{min} is the minimum daily air temperature (°C), and k_{Rs} is the adjustment coefficient 0.17 (°C$^{-0.5}$).

Methodology

ANNs are mathematical models, which have the ability to learn from examples and recognize data patterns between them, in order to adapt solutions over time and process information quickly (Jain et al. 2008). An ANN model usually consists of large number of layers of interconnected neurons, which is associated with weights representing the connection strengths and transfer or activation functions (Yassin et al. 2016). In this study, a typical multilayer ANN was used (Fig. 2). A layer was included for each ANN as input where the number of neurons is identical to the number of input variables. Meanwhile, the number of neurons in the output layer is equal to the unknown variables. The layers between the input and output layers are known as hidden layers. Indeed, ANNs consist of a single hidden layer in most cases (Antonopoulos and Antonopoulos 2017).

In the last decades, computational development has leaded improvements to the applicability of ANNs; they have become increasingly important due to their wide scientific application. ANNs are defined as interconnected processors, parallel distributed and formed by simple processing units, which have a natural tendency to store experimental knowledge and make it available for applications and are considered an effective tool for modeling

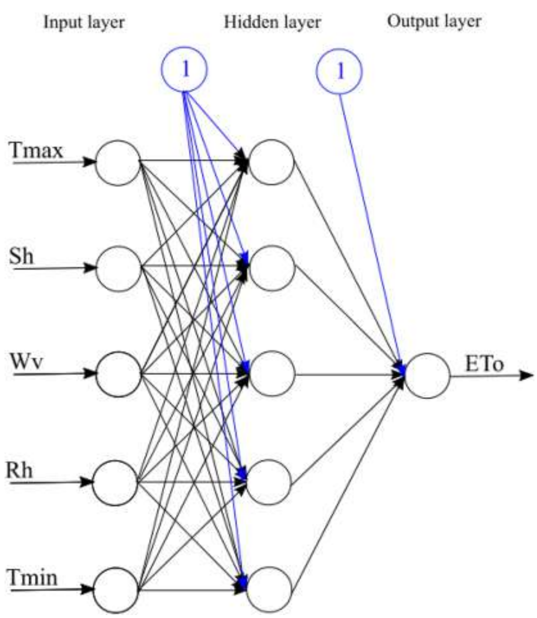

Fig. 2 Structure of ETo artificial neural network model, where T_{max}, T_{min}, Ra, Rh and Wv are the maximum daily temperature (°C), minimum daily temperature (°C), sunshine hours (h), relative humidity (%) and wind speed (m s^{-1}), respectively

nonlinear processes, since they require few inputs and can take input–output relationships without any understanding of the physical process involved (Haykin 1999).

To investigate the relationship between ETo and climatic variables such as T_{max}, Sh, Wv, Rh and T_{min}, preliminary processing using correlation coefficients were performed. This pre-processing shows a strong linear association between T_{max} and ETo using 12 stations ($r = \sim 0.7$, $p < 0.01$), followed by Sh, Wv, Rh and T_{min} variables with values of 0.47, 0.39, − 0.38 and 0.21 ($p < 0.01$), respectively (Table 2). The high correlation values of the T_{max} and Sh with ETo found in this study are consistent with the results of Garcia et al. (2004) who refer that the radiation term is more important for ETo in the Altiplano than that in sea level regions. Likewise, the daily thermal amplitude (difference between T_{max} and T_{min}) mainly in arid conditions permits a proper approximation of the radiation term influencing ETo at the Altiplano. This suggests that all models to be evaluated in the present study should consider T_{max} as the main input variable, because T_{max} defines the vapor holding capacity of the air. The relationships between the climatic parameters and the ETo indicate that there is a direct relationship between ETo and T_{max}, Sh, Wv, as found by Antonopoulos and Antonopoulos (2017) and Shiri (2017), and an inverse relationship with Rh (Fig. 3).

The principal component analysis (PCA) (Jolliffe 2002) was used to identify the climatic variables that can be considered the most important in the formulation of RNA models. A high correlation between T_{max} and Sh is observed in Fig. 3, suggesting that T_{max} could replace Sh in the ANN models formulation models, when there is a lack of Sh. Despite the cold weather in the Peruvian Altiplano, a weak correlation between ETo and T_{min} is observed, mainly in the stations located in the northern of the TLB. Garcia et al. (2004) showed that when the air temperature is close to T_{min}, the air is nearly saturated with water vapor and the relative humidity is close to 100%; therefore, it conditions the amount of ETo in the Altiplano mainly at sunrise.

To select a suitable ANN architecture (number of neurons as input layer), an ANN model of multilayer perception was considered. Indeed, there is not a usual methodology (Antonopoulos and Antonopoulos 2017; Jain et al. 2008; Wu et al. 2014); hence, different combinations of climatic parameters were used (T_{max}, T_{min}, Sh, Rh, Wv) as input variables to estimate the ETo using ANN models (Table 3). Finally, 13 ANN models were developed to test the performance of different combinations of inputs parameters, including the best climatic parameters associated with ETo, resulting from correlation analysis (CA) and PCA; the models with different input combinations evaluated in this work are listed in Table 3. M1, M2, M3, M4 and M5 are designed as temperature-based models,

Table 2 Correlation coefficients (r) between climatic variables (T_{max}, T_{min}, Sh, Rh, Wv) and ETo for 12 stations

Variable	Stations												Average
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	
Rh	-0.36	-0.50	-0.34	-0.26	-0.60	-0.42	-0.23	-0.35	-0.28	-0.36	-0.37	-0.51	-0.38
Sh	0.65	0.60	0.42	0.44	0.40	0.43	0.38	0.41	0.38	0.43	0.48	0.62	0.47
Tmax	0.58	0.76	0.73	0.73	0.79	0.70	0.71	0.68	0.72	0.71	0.64	0.78	0.71
Tmin	-0.01	0.01	0.46	0.06	0.24	0.28	0.39	0.33	0.27	0.35	-0.01	0.08	0.21
Wv	0.15	0.38	0.36	0.47	0.42	0.52	0.36	0.32	0.29	0.45	0.49	0.40	0.39

Legend
- Very high
- High
- Moderate
- Low
- Very low

The abbreviations S1 to S12 represent the code of the climatic stations, where S1: Ananea, S2: Azangaro, S3: Capachica, S4: Chucquibambilla, S5: Desaguadero, S6: Huancane, S7: Juli, S8: Juliaca, S9: Lampa, S10: Puno, S11: Putina, S12: Salcedo. T_{max} daily maximum temperature, T_{min} daily minimum temperature, *Sh* sunshine hours, *Rh* daily relative humidity average, *Wv* wind speed

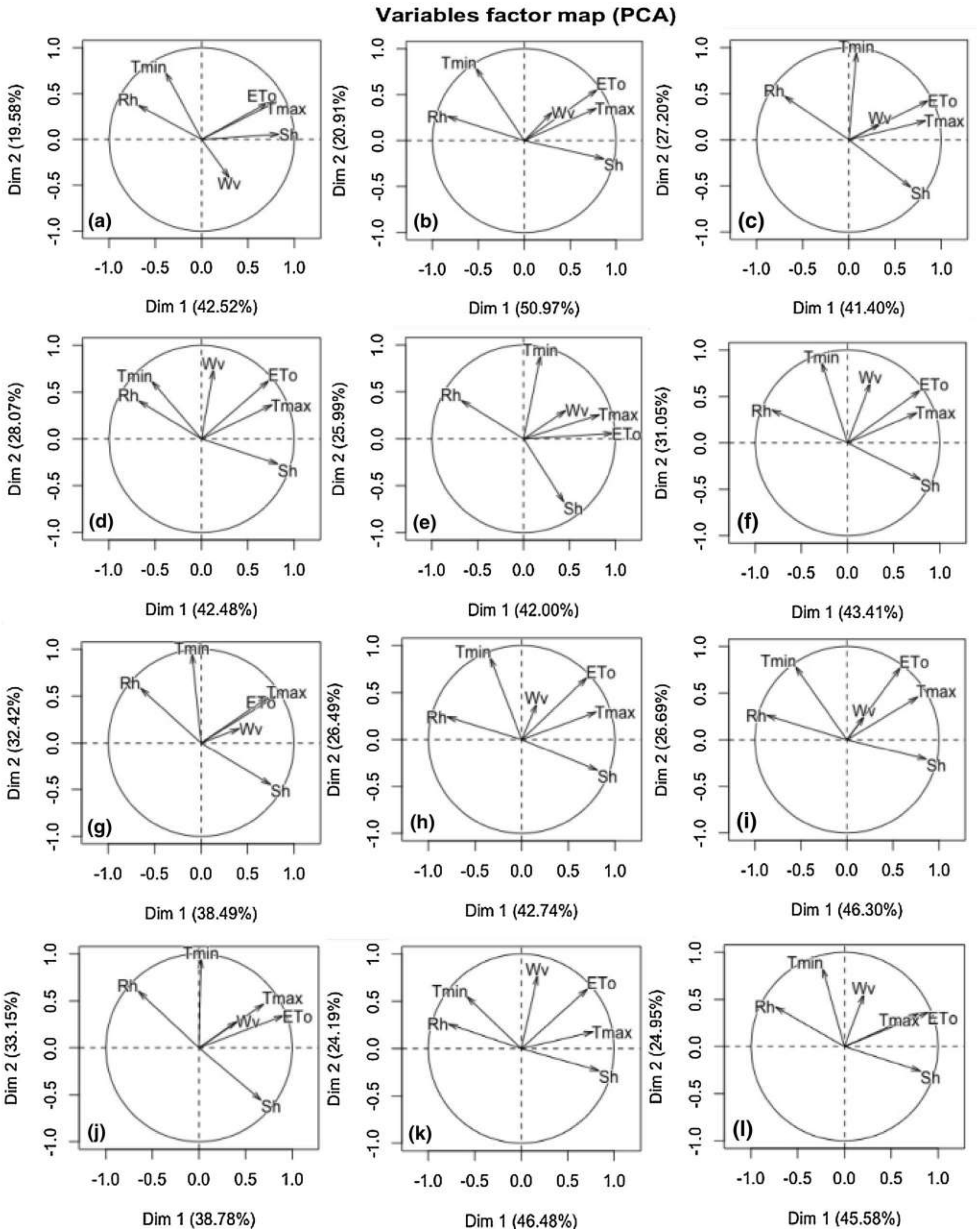

Fig. 3 Bi-plots obtained from principal component analysis using maximum daily temperature (T_{max}), minimum daily temperature (T_{min}), sunshine hours (Sh), relative humidity (Rh) and wind speed (Wv) for the stations such as: **a** Ananea, **b** Azángaro, **c** Capachica, **d** Chuquibambilla, **e** Desaguadero, **f** Huancané, **g** Juli, **h** Juliaca, **i** Lampa, **j** Puno, **k** Putina and **l** Salcedo

Table 3 Combinations of input variables for the development of ANN models for each model proposed

Model	Input variable				
	T_{max} (°C)	T_{min} (ªC)	Sh (h)	Rh (%)	Wv (m s^{-1})
M1	*				
M2	*		*		
M3	*				*
M4	*			*	
M5	*	*			
M6	*		*		*
M7	*			*	*
M8	*		*	*	
M9	*	*	*		
M10	*	*		*	
M11	*		*	*	*
M12	*	*		*	*
M13	*	*	*	*	*

The abbreviations M1 to M13 represent the code of the ANN model, where M1: model 1, M2: model 2, M3: model 3, M4: model 4, M5: model 5, M6: model 6, M7: model 7, M8: model 8, M9: model 9, M10: model 10, M11: model 11, M12: model 12 an M13: model 13, T_{max} daily maximum temperature, T_{min} daily minimum temperature, *Sh* sunshine hours, *Rh* daily relative humidity average, *Wv* wind speed

*Indicates each input variable used as parameter in the modeling

additionally M5 presenting similarity to the Hargreaves–Samani method (HS). Inserting a third input (Sh, Rh and Wv) forms the M6, M7, M8, M9 and M10 model variable. Then, the models M2 to M13 additionally to T_{max} consider Sh, Rh, Wv and T_{min}. Finally, M13 consider the all-climatic variables and its structure is similarity to the FAO PM56 method. The network was trained and tested for each combination summarized in Table 3.

To estimate ETo, it is very common to manage few climatic parameters, because of data scarcity for modeling. However, alternative ANN models were developed since they consider less number of climatic parameters. For instance, T_{max} is used to develop 13 ANN models. To develop ANN models that estimate the ETo, the Neuralnet package (Günther and Fritsch 2010) in R software (Ihaka and Gentleman 1996) was used, because R provides a flexible analysis toolkit where all the standard statistical techniques are built-in.

To minimize the influence of absolute scale, all inputs to the ANN models were normalized to fall between 0 and 1 (Feng et al. 2017). In this study, the following equation was used:

$$x_{norm} = \frac{x_0 - x_{min}}{x_{max} - x_{min}}, \tag{5}$$

where x_{norm}, x_0, x_{min}, and x_{max} are the normalized value, the original data, the minimum value, and the maximum value, respectively.

Likewise, the back propagation (BP) learning algorithm was used to train ANN models. According to Huo et al. (2012), this process basically involves forward propagation of input and backward propagation of errors. In the forward propagation, the effect of an applied activity pattern at the input layer was propagated layer by layer through the network. The activation value at *i*th neuron in *n*th layer a_i^n is given by the following equation:

$$a_i^n = \sum_{j=1}^{m} W_{ji}^n y_j^{n-1} + b_i^n \tag{6}$$

where W_{ji} is the weight of the link between the *i*th neuron in the *n*th layer and *j*th neuron in the $(n-1)$ layer, y_i is the output of the *j*th neuron in the $(n-1)$th layer, and b_i is the bias of the *i*th neuron in the *n*th layer.

The activation value of a neuron is used to obtain the output value of a neuron through the transfer function. The general functional form of the sigmoidal logistic transfer function, which is the most commonly used nonlinear transfer function, was used in the study. It is expressed by

$$f(t) = {}^{1}\!/_{(1 + \exp(-\beta t))}, \tag{7}$$

where *t* represents the weighted sum of input for a node in the hidden layer, and exp denotes the natural exponential function. The function value of each neuron in the output layer was obtained by propagating the effect of the input through layers. The goal of ANN is to establish a relationship of the form:

$$Y^m = f(X^n). \tag{8}$$

Usually, the network is trained by a BP algorithm, which adjusts the weights and biases to minimize the error function given by

$$E = \sum_p \sum_m (y_i - O_i)^2. \tag{9}$$

There is no established methodology for the selection of the appropriate network architecture before training (Antonopoulos and Antonopoulos 2017). In this study, the optimal number of neurons of the hidden layer was determined using trial and error method, the trial and error procedure started with one hidden neuron initially, and the number of hidden neurons was increased to 10 with a step size of 1 at each trial to find the architecture that optimizes the quadratic error in the output layer. Every ANN model was evaluated using ETo information calculated with the FAO-PM56 method for the 1963–2015 period (variable for each station). The training set comprising 70% of the total data was randomly chosen and remaining 30% was used for the test/validation phase.

To evaluate model performances, modeled ETo was compared to FAO-PM56 ETo. Five statistical parameters were selected for its evaluation such as mean absolute error (MAE), root mean square error (RMSE), Nash–Sutcliffe coefficient (NE) and correlation coefficient (r), defined as follows:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|}{n}, \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{n}}, \tag{11}$$

$$\mathrm{NE} = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2}, \tag{12}$$

$$r = \frac{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)\left(\hat{y}_i - \bar{\hat{y}}\right)}{\sqrt{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)^2}}, \tag{13}$$

where n is the number of observations, y_i stands for the observed value of ETo, $\hat{y}_i$ stands for the modeled value, $\bar{y}$ stands for the mean value of the observed ETo, and $\bar{\hat{y}}$ is the mean value of the modeled ETo. MAE is the average value of the absolute differences between the observed and modeled values; a low MAE implies a high performance of the model. The RMSE expresses the error and a small value of the RMSE suggests a better approximation. The NE coefficient provides information about the comparison of the relationships between the error models and the variance of the observed data. The optimal value of NS is 1, which represents a perfect fit. Finally, r measures the linear relationship between the estimated and calculated values, where the values closest to 1 indicate a better association between them.

Results and discussion

Performance of ANN models

Model performance statistics show that the average values over 12 stations of MAE range from 0.20 to 0.42 mm day^{-1}, RMSE ranges from 0.26 to 0.52, NS values from 0.5 to 0.9 and r values from 0.73 to 0.93 (Table 4). Figure 4a shows that the model 13 (M13) presents the best performance among the 13 models proposed in terms of errors, since the MAE and RMSE are smaller than others (MAE = 0.20, RMSE = 0.26). M1 presents the lowest values of efficiency with MAE = 0.42 mm day^{-1}, RMSE = 0.52 mm day^{-1}, NE = 0.53 and r = 0.73. The results suggest that the inclusion or exclusion of any of the input variables significantly affects the performance of the models (see Appendix A).

Moreover, these results show that the average values (10 models) of MAE range from 0.29 to 0.37 mm day^{-1}; values of RMSE range from 0.37 to 0.46; NE values range from 0.48 to 0.82 and r values range from 0.69 to 0.91 (Table 5). The best results are achieved at the station 12 (S12, 3825 m a.s.l.), while the lowest performance is shown at the station 1 (S1, 4660 m a.s.l) between 12 stations analyzed (Figs. 1, 4). This suggests that the localization of the climatic station significantly affects the performance of the models (see Appendix A). The stations located in semi-arid regions show better performance compared to the station located in semi-humid regions.

Comparison between ANN, FAO-PM56 and HS models

The comparison of the M13 estimated daily ETo and FAO-PM56 ETo (Fig. 5) shows a high correlation (r) between them; the values of r vary between 0.69 to 0.91 for S1 and S12, respectively. Likewise, it is observed that there are no marked overestimates and underestimations when compared to FAO-PM56 ETo (except S1, Fig. 5a). The results suggest

Table 4 Average performance statistics for each ANN model using different combinations as input variable during the test phase and Hargreaves–Samani method

Parameter	Models													
	HS	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
MAE	0.451	0.416	0.366	0.358	0.396	0.392	0.300	0.326	0.353	0.316	0.363	0.273	0.289	0.204
RMSE	0.554	0.522	0.465	0.448	0.500	0.497	0.381	0.412	0.452	0.407	0.461	0.349	0.372	0.264
NE	0.456	0.531	0.624	0.646	0.564	0.559	0.743	0.693	0.646	0.705	0.615	0.779	0.748	0.870
r	0.802	0.729	0.790	0.802	0.751	0.746	0.862	0.830	0.803	0.840	0.783	0.882	0.863	0.932

The abbreviations M1 to M10 represent the code of the ANN model, where M1: model 1, M2: model 2, M3: model 3, M4: model 4, M5: model 5, M6: model 6, M7: model 7, M8: model 8, M9: model 9, M10: model 10, M11: model 11, M12: model 12 and M13: model 13. *MAE* mean absolute error, *RMSE* root mean square error, *NE* Nash–Sutcliffe coefficient and correlation coefficient (*r*)

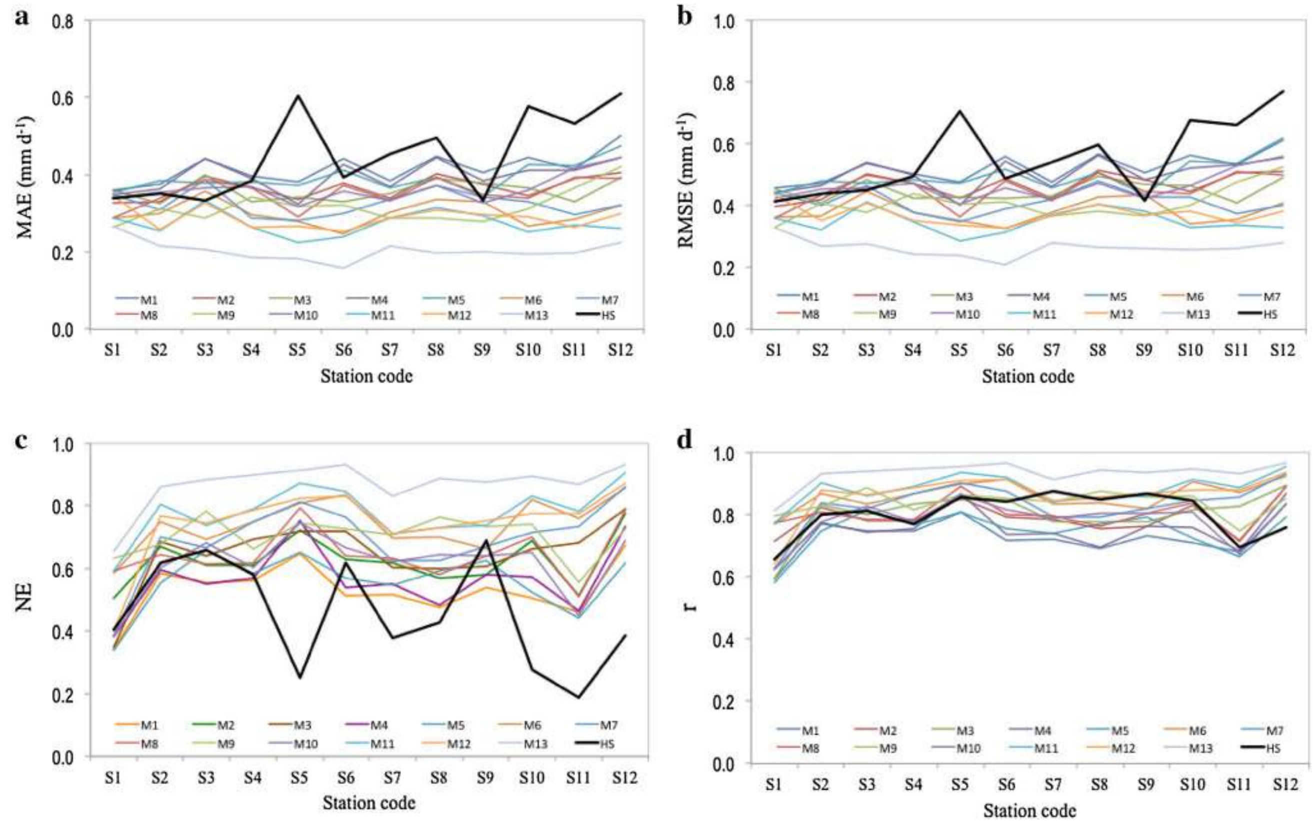

Fig. 4 ANN and HS models performance for each station analyzed, **a** MAE, **b** RMSE, **c** NE and **d** r

Table 5 Average statistic parameters for each meteorological station for ANN and HS model

Parameter	Stations											
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
MAE	0.321	0.316	0.362	0.326	0.301	0.326	0.323	0.358	0.331	0.333	0.343	0.377
RMSE	0.402	0.405	0.452	0.420	0.379	0.419	0.405	0.460	0.423	0.429	0.439	0.474
NE	0.473	0.685	0.678	0.683	0.770	0.708	0.644	0.645	0.665	0.699	0.615	0.786
r	0.686	0.829	0.823	0.825	0.877	0.838	0.801	0.800	0.815	0.833	0.779	0.890
MAE	0.338	0.351	0.333	0.385	0.604	0.394	0.453	0.496	0.333	0.578	0.532	0.609
RMSE	0.414	0.440	0.451	0.496	0.705	0.489	0.539	0.595	0.417	0.675	0.661	0.771
NE	0.404	0.618	0.658	0.577	0.250	0.617	0.379	0.426	0.690	0.279	0.187	0.386
r	0.656	0.798	0.812	0.770	0.855	0.843	0.875	0.847	0.868	0.845	0.695	0.757

The abbreviations S1 to S12 represent the code of the climatic stations, where S1: Ananea, S2: Azangaro, S3: Capachica, S4: Chuquibambilla, S5: Desaguadero, S6: Huancane, S7: Juli, S8: Juliaca, S9: Lampa, S10: Puno, S11: Putina, S12: Salcedo. *MAE* mean absolute error, *RMSE* root mean square error, *NE* Nash–Sutcliffe coefficient and correlation coefficient (*r*)

that the ANN models were not over-trained during calibration phase. Therefore, ETo estimates show greater consistency during test phase. This is concordant with the analysis of model performance done by Liu et al. (2012).

Likewise, the ability of ANN models to adequately estimate ETo in highland regions using different input variables is confirmed, highlighting their flexibility and the ability to use nonlinear models. This is consistent with different studies developed in the world (Cobaner 2011; Jain et al. 2008; Kumar et al. 2008; Rahimi Khoob 2008; Yassin et al. 2016; Zanetti et al. 2007). In addition, these models present the ability to reproduce the minimum and maximum ETo along the study area (Fig. 5). Therefore, our results represent an alternative option to replace the FAO-PM56 method in highland regions; since the lack of

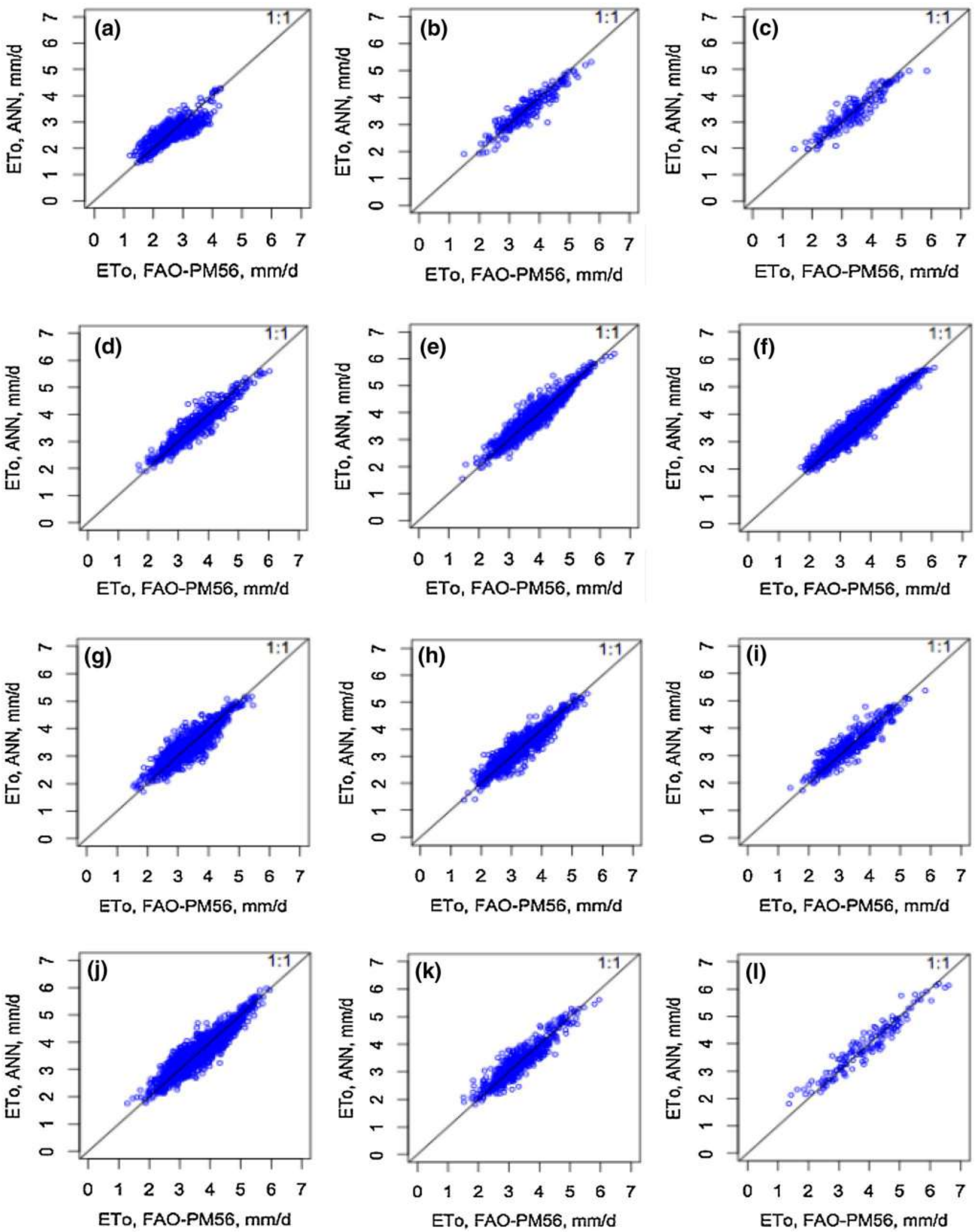

Fig. 5 Observed (FAO-PM56) and estimated (ANN) reference evapotranspiration (ETo) for the validation phase using stations such as: **a** Ananea, **b** Azángaro, **c** Capachica, **d** Chuquibambilla, **e** Desaguadero, **f** Huancané, **g** Juli, **h** Juliaca, **i** Lampa, **j** Puno, **k** Putina and **l** Salcedo

information required by this method, it could be a limiting factor for your application.

Hargreaves–Samani (HS) showed a poor performance in the estimation of the ETo in the Peruvian highlands, reaching values of MAE, RMSE, NE and r of 0.451, 0.554 mm day^{-1}, 0.456 and 0.802, respectively, which are less efficient compared to the 13 ANN models. These results show that the use of ANN models is a valid option for the estimation of ETo in highland conditions, since all ANN models presented better performance compared to HS model, which is the method commonly used to estimate the ETo for its simplicity and little information requirement.

The worst performance results (NE < 0.404) of the HS model were presented in stations located in humid (S1 and S7) and dry semi-humid (S5, S10, S11 and S12) climatic regions. Garcia et al. (2004) reported that the HS formula adequately estimates ETo in the northern (more humid) locations, which does not occur in the southern locations, which present conditions of greater aridity. The results obtained in this study suggest that what was indicated by Garcia et al. (2004) would not be completely fulfilled in the Peruvian highlands, since the stations that are located in humid and semi-humid climates as well as north of the Bolivian highlands have low levels performance. However, it was determined that in stations with lower thermal amplitude (< 14.2° C) the lowest performance levels are presented in the estimation of the ETo with the HS method, which does not occur markedly with the ANN models that they estimate adequately ETo in the three climatic regions, with a slight loss of efficiency in stations located in humid regions.

Influence of structure of ANN and climate parameters on the model performance of ANNs

The performances maximum is obtained by M13, M11, M6 and M3 with models that consider 5, 4, 3 and 2 input variables, respectively (NS = 0.87, 0.78, 0.74 and 0.65, respectively, see Table 4). This suggests that the number of input variables is associated to obtain a high performance from modeling. However, relevant results are not always obtained with complex structures in the ANN models, such that the case of M6 only considers three climatic parameters (T_{max}, Sh, Wv) as input variables (NE = 0.74). This also occurs with the M5 that considers two input variables (T_{max} and T_{min}), similar to those used by the HS equation, despite this the performance level of the ANN model (NE = 0.56) is higher than the one obtained by the HS method (NE = 0.46), which suggests the use of ANN models for the estimation of ETo under conditions of reduced availability of climatic information and high altitude regions. On the contrary, a greater number of neurons as input and hidden layers could generate a greater accumulation of errors (Liu et al. 2012).

In error terms, the results also indicate that the best models are M13, M11, M6, M3 and M1; these results suggest that a better modeling is achieved when the five climatic parameters (T_{max}, Sh, Rh, Wv and T_{min}) are included as input variables (Table 4). However, the climatic data unavailability would restrict the use of all climatological stations of the highland regions.

In highlands areas, as the Peruvian Altiplano, T_{max}, Sh and Wv climatic variables are the most important climatic parameters that determinate the ETo, the consideration of these three climatic parameters as input variables improve the model performance than those models where only considered the Rh and T_{min} parameters. Indeed, to achieve efficient ANN models in highland areas, it is essential to consider the T_{max}. It is possible to improve using Sh and Wv or combinations between them, also reported by Petković et al. (2015). Likewise, a high sensitivity of model using T_{max}, Sh and Wv is determined in this paper, due to the effects of the high altitude of the Altiplano (high solar radiation but low radiation term, moderate aridity, temperature and wind). This is partially concordant with ETo studies for the Bolivian Altiplano region developed by Garcia et al. (2003, 2004, 2007) and Chang et al. (2017).

The error and efficiency analyses suggest that T_{max} input variable in highlands areas has high influence for the ETo estimations, where the M1 model presented the better correlation ($r = 0.73$, $p < 0.01$) with FAO-PM56 ETo during the test phase (Fig. 4). This confirms the ability of the ANN models to adequately estimate the ETo in highlands areas using a small number of input variables (only T_{max}) (Antonopoulos and Antonopoulos 2017; Petković et al. 2015; Traore et al. 2010). On the contrary, the use of T_{min} (M11) improves slightly the performance of the model, despite the lower temperatures prevailing in the highlands areas as the Peruvian Altiplano.

The best results of the models were obtained when climatic stations are located in dry semi-humid regions (NE = 0.69) (see Fig. 1). The stations S1 and S7 presented the lowest correlations (NE = 0.56); the common characteristic of these two stations is that they are located in humid regions, where AI is greater than 0.65. The condition of more humid regions for S3 and S7 stations is due to the proximity to the Lake Titicaca, where the total annual precipitation is most significant and tends to decrease with distance from the lake (Roche et al. 1992). For S1, the low performance could be result of the advection events of airmass from Amazonian basin, which are presented with a decreasing frequency to the south of the TLB (Aceituno 1996).

These results suggest that both the proximity to water body and the altitude may condition the efficiency of the ANN and HS models to estimate adequately the ETo; this is confirmed by Garcia et al. (2004). However, in this study it was determined that stations that are located in the humid region (S1 and S7) showed lower performance compared to stations located in dry semi-humid and semi-arid regions, where the value of the AI is between 0.2 and 0.65. According to the results of this study could be affirmed that the degree of climatic drought (Aridity Index) have a significant influence in the performance of the ETo, which does not apply to the use of the HS method, where thermal amplitude is the conditioning factor in the performance of the method in conditions of the Peruvian highlands. The influence of these factors on the estimation of the ETo could not be categorically affirmed, since the number of stations located at 3910 m a.s.l and in humid regions used in this research is limited.

Conclusions

The results indicate that ANN models can be used successfully to estimate ETo from the combinations of five daily climatic variables such as maximum temperature (T_{max}), minimum temperature (T_{min}), sunshine hours (Sh), relative humidity (Rh) and wind velocity (Wv) or fewer in highland regions as the Peruvian Altiplano, since the performance levels of all ANN models exceeded that obtained by means of the HS equation. To estimate adequately ETo using ANN models, it is essential to consider climatic variables such as T_{max}, associated with Sh and Wv or combinations between them; this confirms that the radiation term is more important for ETo in highlands areas. In addition, other climatic parameters (Rh and T_{min}) can be used as input variables to improve the modeling, despite the lower temperatures prevailing in the region. In general, the M13 model that considers five input variables presents a better performance for ETo estimations.

With respect to Hargreaves–Samani (HS) method, it was determined that in stations with lower thermal amplitude (< 14.2 °C) the lowest performance levels are presented in the estimation of the ETo, which does not occur markedly with the ANN models that they estimate adequately ETo, without considering the thermal amplitude, which is a peculiarity of the highlands areas.

Having determined the climatic variables of greater influence in the estimation of the ETo, the results obtained by M11, M6, M3 and M1 models that consider 4, 3, 2 and 1 input variable, respectively, suggest the ability of the ANN models to adequately model the ETo using a number of input variables limited in highland regions. The M5 model that considers two input variables (T_{max} and T_{min}), similar to those used in the HS equation, presents a better performance (NE = 0.56) than the HS method (NE = 0.46), which suggests the use of ANN models for the estimation of ETo under conditions of reduced availability of climatic information and high altitude regions. Therefore, ANN models represent a great option to replace the FAO-PM56 and HS methods, when ETo data series are scarce. The highest performance of ANN models suggests that the stations located in dry semi-humid regions show better performance compared to the stations located in humid regions, which could be attributed to the presence of a water body and the advection events of air mass. This suggests that the proximity to the water body, altitude and climatic condition (aridity index) may condition the efficiency of ANN models to estimate adequately ETo.

ETo is a key element in the efficient management for water resources in agriculture on Altiplano regions. Therefore, the ANN models can be considered a new opportunity for an adequate estimation of ETo over highland regions where reliable collection of wind speed, humidity, and radiation is limited. To improve their applications, future research could emphasize on developing ANN models to highland regions. To evaluate its applicability over the highland regions, a greater number of climatic stations for its analysis are required (currently not available). This would allow to investigate confirm the influence of altitude in the efficiency of ANN models along the highland regions.

Acknowledgements The authors would like to thank *Servicio Nacional de Meteorología e Hidrología* (SENAMHI) for providing us the climatic information used in this investigation. They also would like to acknowledge *Programa de Doctorado en Recursos Hídricos of the Universidad Nacional Agraria La Molina* for providing support in the development of this research.

Compliance with ethical standards

Conflict of interest This statement is to certify that all Authors have contributed significantly to the work, have read the manuscript, attest to the validity and legitimacy of the data and its interpretation, and agree to its submission to the Neural Computing & Applications Journal. We warrant that the article is the Authors' original work. We warrant that the article has not received prior publication and is not under consideration for publication elsewhere. All authors agree that author list is correct in its content and order and that no modification to the author list. Therefore, we have no conflicts of interest to disclose. On behalf of all Co-Authors, the corresponding Author shall bear full responsibility for the submission.

References

Aceituno P (1996) Elementos del Clima en el Altiplano Sudamericano. Revista Geofísica-IPGH 44:37–55

Adamala S (2018) Temperature based generalized wavelet-neural network models to estimate evapotranspiration in India. Inf Process Agric 5:149–155. https://doi.org/10.1016/j.inpa.2017.09.004

Adeloye AJ, Rustum R, Kariyama ID (2012) Neural computing modeling of the reference crop evapotranspiration. Environ Model Softw 29:61–73. https://doi.org/10.1016/j.envsoft.2011.10.012

Allen RG (2000) Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. J Hydrol 229:27–41. https://doi.org/10.1016/S0022-1694(99)00194-8

Antonopoulos VZ, Antonopoulos AV (2017) Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Comput Electron Agric 132:86–96. https://doi.org/10.1016/j.compag.2016.11.011
Binford MW, Kolata AL, Brenner M, Janusek JW, Seddon MT, Abbott MB, Jason H (1997) Climate variation and the rise and fall of an Andean civilization. Quat Res 47:235–248
Chang Y, Wang J, Qin D, Ding Y, Zhao Q, Liu F et al (2017) Methodological comparison of alpine meadow evapotranspiration on the Tibetan Plateau, China. PLoS One 12(12):e0189059. https://doi.org/10.1371/journal.pone.0189059
Citakoglu H, Cobaner M, Haktanir T, Kisi O (2014) Estimation of monthly mean reference evapotranspiration in Turkey. Water Resour Manag 28:99–113. https://doi.org/10.1007/s11269-013-0474-1
Cobaner M (2011) Evapotranspiration estimation by two different neuro-fuzzy inference systems. J Hydrol 398:292–302. https://doi.org/10.1016/j.jhydrol.2010.12.030
Droogers P, Allen RG (2002) Estimating reference evapotranspiration under inaccurate data conditions. Irrigat Drain Syst 16(1):33–45. https://doi.org/10.1023/A:1015508322413
Falamarzi Y, Palizdan N, Huang YF, Lee TS (2014) Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). Agric Water Manag 140:26–36. https://doi.org/10.1016/j.agwat.2014.03.014
Feng Y, Peng Y, Cui N, Gong D, Zhang K (2017) Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Comput Electron Agric 136:71–78. https://doi.org/10.1016/j.compag.2017.01.027
Garcia M, Raes D, Jacobsen SE (2003) Reference evapotranspiration and crop coefficient of quinoa (Chenopodium quinoa Willd) in the Bolivian Altiplano. Agric Water Manag 60:119–134
Garcia M, Raes D, Allen R, Herbas C (2004) Dynamics of reference evapotranspiration in the Bolivian highlands (Altiplano). Agric For Meteorol 125(1–2):67–82. https://doi.org/10.1016/j.agrformet.2004.03.005
Garcia M, Raes D, Jacobsen SE, Michel T (2007) Agroclimatic constraints for rainfed agriculture in the Bolivian Altiplano. J Arid Environ 71:109–121
Garreaud R, Vuille M, Clement C (2003) The climate of the Altiplano: observed current conditions and mechanisms of past changes. Palaeogeogr Palaeocl 194:5–22. https://doi.org/10.1016/S0031-0182(03)00269-4
Garreaud RD, Vuille M, Compagnucci R, Marengo J (2009) Present day South American climate. Palaeogeogr Palaeocl 281:180–195
Günther F, Fritsch S (2010) Neuralnet: training of neural networks. R J 2(1):30–38
Haykin S (1999) Neural networks: a comprehensive foundation. Pearson Education, Delhi
Huang H, Han Y, Cao M, Song J, Xiao H (2016) Spatial-temporal variation of aridity index of China during 1960–2013. Adv Meteorol 2016:1–10. https://doi.org/10.1155/2016/1536135
Huo Z, Feng S, Kang S, Dai X (2012) Artificial neural network models for reference evapotranspiration in an arid area of northwest China. J Arid Environ 82:81–90. https://doi.org/10.1016/j.jaridenv.2012.01.016
Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5(3):299–314
Jain SK, Nayak PC, Sudheer KP (2008) Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrol Process 22:2225–2234. https://doi.org/10.1002/hyp.6819
Jolliffe IT (2002) Principal component analysis. Springer, New York. https://doi.org/10.1007/b98835
Kişi Ö (2006) Generalized regression neural networks for evapotranspiration modelling. Hydrol Sci J 51:1092–1105. https://doi.org/10.1623/hysj.51.6.1092
Kumar M, Bandyopadhyay A, Raghuwanshi NS, Singh R (2008) Comparative study of conventional and artificial neural network-based ETo estimation models. Irrig Sci 26:531–545. https://doi.org/10.1007/s00271-008-0114-3
Laaboudi A, Mouhouche B, Draoui B (2012) Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions. Int J Biometeorol 56:831–841. https://doi.org/10.1007/s00484-011-0485-7
Laqui W (2010) Aplicación de Redes Neuronales Artificiales a la Modelización y Previsión de Caudales Medios Mensuales del Rio Huancané. Revista Peruana Geo-Atmosférica (RPGA), SENAMHI-Perú 2:30–44
Liu Z, Peng C, Xiang W, Deng X, Tian D, Zhao M, Yu G (2012) Simulations of runoff and evapotranspiration in Chinese fir plantation ecosystems using artificial neural networks. Ecol Model 226:71–76. https://doi.org/10.1016/j.ecolmodel.2011.11.023
Lujano E, Felipe O, Lujano A, Quispe J (2015) Validación de la precipitación estimada por satélite TRMM y su aplicación en la modelación hidrológica del rio Ramis Puno Perú. Revista de Investigaciones Altoandinas 17:2. https://doi.org/10.18271/ria.2015.116
Núñez L, Grosjean M, Cartagena I (2002) Human occupations and climate change in the Puna de Atacama, Chile. Science 298:821–824
Petković D, Gocic M, Trajkovic S, Shamshirband S, Motamedi S, Hashim R, Bonakdari H (2015) Determination of the most influential weather parameters on reference evapotranspiration by adaptive neuro-fuzzy methodology. Comput Electron Agric 114:277–284. https://doi.org/10.1016/j.compag.2015.04.012
Pillco Zolá R, Bengtsson L, Berndtsson R, Martí-Cardona B, Satgé F, Timouk F, Bonnet M, Mollericom L, Gamarra C, Pasapera J (2018) Modeling Lake Titicaca daily and monthly evaporation. Hydrol Earth Syst Sci Discuss. https://doi.org/10.5194/hess-2018-127
Rahimi Khoob A (2008) Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. Irrig Sci 27:35–39. https://doi.org/10.1007/s00271-008-0119-y
Rau P, Bourrel L, Labat D, Frappart F, Ruelland D, Lavado W, Dewitte B, Felipe O (2017) Hydroclimatic change disparity of Peruvian Pacific drainage catchments. Theor Appl Climatol 10:100. https://doi.org/10.1007/s00704-017-2263-x
Roche M, Bourges J, Cortes J, Mattos R (1992) Climatology and hydrology of the Lake Titicaca basin. In: Dejoux C, Iltis A (eds), Lake Titicaca: a synthesis of limnological knowledge. Kluwer Academic, Dordrecht, vol 68, pp 63–88 (Monographiae Biologicae). ISBN 0-7923-1663-0
Shiri J (2017) Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. Agric Water Manag 188:101–114. https://doi.org/10.1016/j.agwat.2017.04.009
Tabari H, Hosseinzadeh Talaee P (2013) Multilayer perceptron for reference evapotranspiration estimation in a semiarid region. Neural Comput Appl 23:341–348. https://doi.org/10.1007/s00521-012-0904-7
Traore S, Wang Y-M, Kerh T (2010) Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. Agric Water Manag 97:707–714. https://doi.org/10.1016/j.agwat.2010.01.002
UNESCO (2003) Water for people, water for life. UNESCO, New York
Wu W, Dandy GC, Maier HR (2014) Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water

quality modelling. Environ Model Softw 54:108–127. https://doi.org/10.1016/j.envsoft.2013.12.016

Yassin MA, Alazba AA, Mattar MA (2016) Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agric Water Manag 163:110–124. https://doi.org/10.1016/j.agwat.2015.09.009

Zanetti SS, Sousa EF, Oliveira VP, Almeida FT, Bernardo S (2007) Estimating evapotranspiration using artificial neural network and minimum climatological data. J Irrig Drain Eng 133:83–89. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:2(83)

Zolá RP, Bengtsson L (2006) Long-term and extreme water level variations of the shallow Lake Poopó, Bolivia. Hydrol Sci J 51:98–114. https://doi.org/10.1623/hysj.51.1.98

Zubieta R, Laqui W, Lavado W (2018) Modelación hidrológica de la cuenca del río Ilave a partir de datos de precipitación observada y de satélite, periodo 2011–2015, Puno-Perú. Revista de Tecnología y Ciencias del Agua

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.