Knowledge extraction from artificial neural networks(ANN) trained to estimate daily reference evapotranspiration in southeastern of rolling pampas of Argentina

Andrea Inés Irigoyen[1,2]

[1] Unidad Integrada Balcarce: Facultad de Ciencias Agrarias UNMdP-INTA Balcarce,

Ruta Nac. 226,km 73.5 (7620) Balcarce, Argentina

[2] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

Corresponding author: irigoyen.andrea@inta.gob.ar

Abstract. Evapotranspiration is an important component of hydrologic balance and represent essential information for irrigation scheduling and water resources planning. Sometimes, the use of the recommended Penman-Monteith method is restricted by the lack of input variables and, therefore, empirical methods become essential. The study aimed: a) to develop and evaluate the performance of models based on artificial neural networks (ANN) to estimate daily values of reference evapotranspiration ($ET0_{PM}$) with a limited number of input variables and b)to apply methods of knowledge extraction based on connection weights and sensitivity analysis to better understanding of ANN. Daily evapotranspiration values computed following the Penman-Monteith equation ($ET0_{PM}$), were used as target outputs for the implementation of the ANN. Data of global radiation (Rg), net radiation (Rn) and extraterrestrial radiation (RTA) were alternated in combinations with air temperature (Ta), vapor pressure deficit (DPV) and wind (u) as inputs to networksof type multilayer perceptron. Also, combinations with basis in RTA and minimum and maximum air temperatures(Tmin, Tmax) were tested. The ANN with best performance for each combination of inputs were retained to evaluate the performance based on multi-criteria analysis. According to the results, it can be concluded that it is possible to estimate accurately daily $ET0_{PM}$ values. Air temperature and deficit of pressure vapor were found to be more effective than wind velocity in modelling ET0, whichever the radiation (Rn, Rg or RTA) used as input. A decomposition method based on Garson's algorithm was applied to quantify the relative importance for each input variable. Sensitivity analysis was also performed to identify relevant inputs and quantify the risk of a certain combination of inputs on target values. The application of complementary proceduresin evaluation of ANN models is discussed, paying attentionespecially on detection of the better predicting variables and analysis of errors.

Key words: Radiation, Deficit Pressure Vapor, Synaptic Weight, Decomposition Method, Sensitivity Analysis

1Introduction

Evapotranspiration is an important component of hydrologic balance and represent essential information for irrigation scheduling and water resources planning. Several models were developed to predict ET0 from meteorological elements and the most recommended model is Penman-Monteith (PM) procedure presented in [1]. Sometimes, the use of the standard method is restricted by the lack of input variables and, therefore, empirical methods become essential.

Because of need of alternative methods for dealing with missing data, some models based on regression have already been evaluated for climate local conditions of southeastern of rolling pampas of Argentina [2][3] [4] Regardless of an acceptable approximation to estimate mean values on 10-days period, a better approximation for daily scale is required. In this sense, the capacity of Artificial Neural Networks (ANN) to solve approximation problems could be a feasible alternative.The ANN are mathematical models inspired frombiological neurons, with computational capacity to solve problems of approximation, prediction and optimization [5][6].A schematic representation of neuron model is given in the Fig. 1.

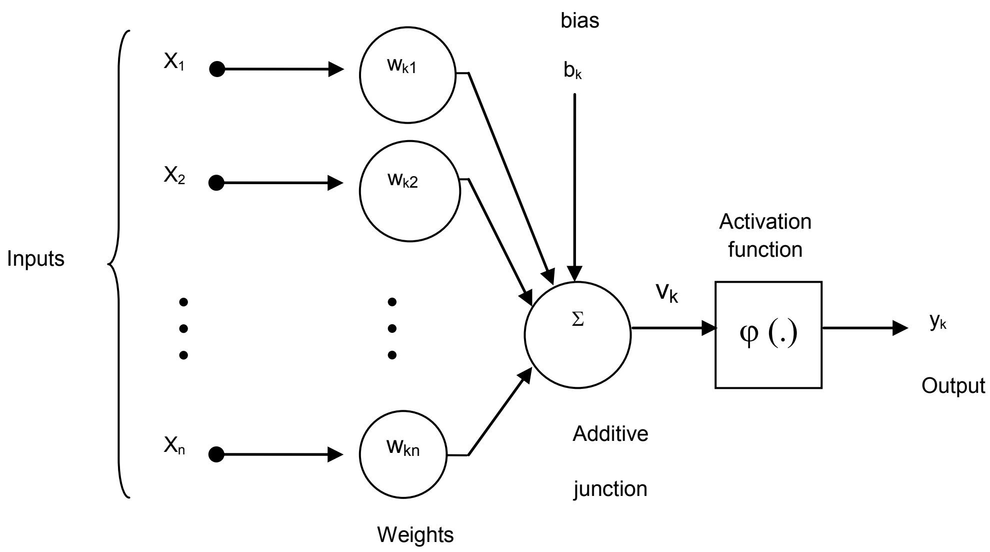

Figure 1. Model of artificial neuron adapted from [5].

Mathematically, the artificial neuron can be described by the equation:

$$y_k = \varphi\left(\sum_{i=1}^{n} x_i w_{ki}\right) + b_k \quad \textbf{(1)}$$

Where yk is the output neuron; φ is the activation function; X1, X2, ..., Xn are the input signals; and wij are the synaptic weight of k neuron and b_k is the bias. In ANN architecture, the neurons are arranged in layers and are interconnected. The summing junction is used to sum up all inputs weighed by the synaptic weights of each neuron. The strength of connection between the two neurons in adjacent layers is represented by what is known as a 'synaptic weight'. The role of the bias is either increasing or reducing the influence of each value for the activation of the neuron. On the other hand, the activation function restricts the output amplitude of each neuron and adds the nonlinear components to the model.

The characteristics of a neural network are: structure or architecture, training algorithm and activation functions. The development of a model based in neural network consists in the definition of these characteristics. To solve approximation problems a supervised training is carried out. Inputs and target outputs are provided to the ANN. Training or learning of a ANN with a defined structure is achieved by adjusting the weights of the neurons through an iterative algorithm that minimizes the error between the predicted and the target outputs. The procedure used to carry out the training process is called the training algorithm. There are many different training algorithms with different performance.The training algorithm stops when a specified condition, or stopping criterion, is satisfied. For the other hand, activations functions for the hidden units are needed to introduce non-linear components [5][6].

The multilayer perceptron network (MLP) is one of the most commonly feed-forward used ANN. A MLP network consists of one input layer, one or more hidden layers and one output layer. The radial basis function (RBF) network is also a feed-forward type of ANN. The property of locality is the main reason why the RBF network can be learned much faster than the MLP [5].

A major drawback often associated with ANN is to be deficient in understanding the knowledge learnt by the trained network.Since the assimilated knowledge from data during training is represented by the network topology, the activation functions and the synaptic weights, some methods are proposed to extract knowledge based on analysis of synaptic weights or sensitivity analysis [7] [8][9].

The objectives were: a) to develop and evaluate the performance of models based on artificial neural networks (ANN) of type multilayer perceptron (MLP) to approximate daily values of reference evapotranspiration ($ET0_{PM}$) with a limited number of input variables and b) to apply methods of knowledge extraction based on connection weights and sensitivity analysis to better understanding of ANN.

Therefore, the major contributions of this work are: (i) identify alternative methods to estimate daily reference evapotranspiration; (ii) discuss the ANN applicability for defining and selecting strategies of estimation from limited climatic data and (iii) interpret the ANN from simple knowledge extraction methods.

The remaining of the paper is structured as follows: Section 2 details related works. Section 3 describes the estimation of daily reference evapotranspiration following the standard method Penman-Monteith. Section 4 describes the development of the ANN models and their evaluation in comparison with the standard method. A simple procedure of knowledge extraction from ANN is detailed in Section 5. And, finally, the conclusions and future works are summarized in Section 6.

2 Related works

The application of neural networks in environmental problems is relatively newer than in other research areas, but is becoming popular because of their ability of capturing nonlinear relationships between the variables, and hence, providing key advantages over traditional statistical techniques. Specifically, the applications of ANN in water resources modeling is increasing. Estimation of reference evapotranspiration (ET0), the basic step toward the calculation of crop water requirements, is a case.

Despite the reference in literature about adequate performance of ANN to approximate evapotranspiration under different climate conditions [10] [11] [12] [13] [14] [15] [16] [17] [18][19] [20] [21], only some studies carried out the estimation with limited variables.In most of studies, wind speed, relative humidity, air temperature and solar radiation were used as predictors. Also, minimum and maximum temperatures,extraterrestrial radiation, and the maximum sunshine hourswere good predictors [15] [21].

The MLP networksare usually trained to estimate evapotranspiration. However, good performances were reported with radial basis networks [20] [21].

No information about interpretation of ANNis reported in most of cases.The study of how uncertainty of inputs affects outputs of ANN is neither described. Some physical interpretation was given in estimating hourly values of ET0 from ANN, following the Garson´s method [7]. Extraction of knowledge was also performed by analyzing the connection weightsfor a case of study on soil waterduring maize crop season [22].

3Target outputs of ANN: meteorological data and estimation of daily reference evapotranspiration

The southeastern region of rolling pampas is characterized with a climate of the Cfb humid-subhumid type, according Köppen classification. The present study is focused at Balcarce, Buenos Aires Province, Argentina (37° 45' S, 58º 18´ W, 130 m altitude).

Meteorological data were obtained from a conventional weather station localized at Experimental Station of Instituto de Tecnologia Agropecuaria INTA Balcarce. The site includes observations of daily maximum and minimum air temperatures (Tmax, Tmin), relative humidity (RH), wind speed (u), and sunshine duration (SH). Measurements were made at a height of 2 m above the soil surface.

The reference evapotranspiration values, that are target outputs for the artificial neural networks (ANN), were computed on the daily basis of Penman-Monteith method ($ET0_{PM}$) for the period 1971-2000, following the recommendations in [1]:

$$ET0_{PM} = \frac{0.408\delta(Rn - G) + \gamma \dfrac{900}{Ta + 273} u(es - ea)}{\delta + \gamma(1 + 0.34u)} \quad (1)$$

where the $ET0_{PM}$ is reference crop ET calculated using the PenmanMonteith-FAO56 method (mm d^{-1}), Rn is the daily net radiation (MJ $m^{-2}d^{-1}$), G is the daily soil heat flux (MJ $m^{-2}d^{-1}$), Ta is the mean daily air temperature at a height of 2 m (°C), u is the daily mean wind speed at a height of 2 m (m s^{-1}), es is the saturation vapor pressure (kPa), ea is the actual vapor pressure (kPa), δ is the slope of the saturation vapor pressure versus the air temperature curve (kPa°C^{-1}), and γ is the psychrometric constant (kPa°C^{-1}).

In this study, the daily values of γ, Rn, es and ea were calculated using the equations (for albedo, α=0.23 for green vegetation surface) given by Allen et al. (1998). The soil heat flux (G) was assumed to be zero over the calculation for time step of 24 h [3]. The measured RH, Tmax and Tmin values were used to calculate ea and es.

At Balcarce, the daily reference evapotranspiration ($ET0_{PM}$) shows a seasonal pattern with maximum occurring in January (4.9 mm d^{-1}) and minimum in July (0.8 mm d^{-1}). Relative contribution of radiation term is dominant with values about 70% from October to March.In addition, the vapor pressure deficit (DPV), calculated from the difference between es and e, was the best correlated variable to evapotranspiration in all months[23].xx

The series on study was finished at 2000 due the increasing missing data in more actual series for some of driving variables to give $ET0_{PM}$. At present and as routine procedure, multiple regression models [2] [4] are used to complete the missing data, at the local weather station.

4 Development and evaluation of ANN

In order to specify a network structure, the relevant input variables and the appropriate number of hidden units respect to samples have chosen. It has been shown that only one hidden layer is required to approximate any continuous function [24]. Models of type MLP with one hidden layer and one output were utilized in this study; therefore, the size of each network was defined by the number of inputs and nodes in the hidden layer.

Different combinations for two and three inputs were evaluated. As relevant inputs were firstly regarded the inputs to Penman-Monteith model. Two alternative inputs to substitute the effect of net radiation (Rn) are proposed. Then, net radiation (Rn), solar radiation (Rg) and radiation on top of the atmosphere or extraterrestrial radiation (RTA) were combined in the input layer with mean air temperature (Ta), deficit of pressure vapor (DPV) and wind velocity (u)to develop the models. The criteria of mandatory input of some variable linked to available energy to evaporate (Rn, Rg or RTA), in all ANN, was due to the predominant contribution of radiation term in reference evapotranspiration, when estimated by Penman-Monteith model[23]. It should be noted that it is not only important to find the best model to approximate $ET0_{PM}$, but a set of models for situations with missing data and then to propose a strategy for selecting the optimal models under different situations of available meteorological data.

The daily global radiation (Rg) values were obtained from relative sunshine hours, according a model adjusted for local conditions [25]. The relative sunshine was obtained as the quotient between actual sunshine hours and theoretical sunshine for each day of the year in the location. The radiation on top of the atmosphere (RTA) or extraterrestrial radiation, which only needs latitude data and day of the year, was also combined with maximum daily temperature (Tmax) and minimum daily temperature (Tmin), similar as inputs in Hargreaves method [26]. The difference between maximum and minimum air temperature is related to the degree of cloud cover and can be used as an indicator of the fraction of extraterrestrial radiation that reaches the earth's surface.

Two types of transformed of sigmoid activation functions (i.e. logistic and hyperbolic tangent) were applied in the hidden layer and linear ones in the output layer. The sigmoid response, in general, allows a network to map a nonlinear process. A linear function was used in output. The training was carried out under conjugated algorithm of errors propagation using the daily values of $ET0_{PM}$(period 1971-2000) as target output in the ANN.

Lack of generalization can be caused by overfitting. A very common technique to avoid this defect is an early stopping criterion that ends training before convergence.So, each ANN architecture was trained under automatic early stopping criterion associated to cross validation method [27]. For this reason, the data set was split into sets for training (period 1971-1988), validation (period 1995-2000) and test (period 1989-1994) to apply cross-validation.

To investigate if partitioned subsets could disruptany pattern of data, as descriptive statistical characterization as correlation analysis were performed for each set. In order to evaluate the hypotheses of equality of frequencies distributions between values of training set with test and validation ones, respectively the non-parametrical Kolmogorov-Smirnov test was applied ($p < 0.05$).Frequencies distribution of $ET0_{PM}$ valuesfrom training set did not differ from respective test and validation sets. Therefore, the requirement of series to be part of the same population was attended. This fact is relevant because cross-validation was applied during training process, and this method is sensitive to the way that available data are divided [27]. Median values for target and inputs for the three sets are presented in Table 1.Further details of target and inputs values over complete series 1971-2000 (i.e. seasonal patterns, contribution of radiation and aerodynamics termsto ET_{PM} and correlation between ET_{PM} andinputs)are given in [23].Besides, relationships among target and inputs of ANN models was not different in the three subsets of data, as showed by the Pearson correlation coefficients reported in Table 2. For this reason, it is deduced that changes in subsets are not statistically evident.

Table 1. Median values of daily reference evapotranspiration ($ET0_{PM}$) and inputs discriminated by sets used for training, validation and test of ANN models.

		Target	Inputs							
Set	n	$ET0_{PM}$	Rn	Rg	RTA	DPV	Ta	Tmax	Tmin	u
		mm d^{-1}	------ MJ m^{-2} d^{-1} --------			kPa	---------- °C -----------------			ms^{-1}
Training	6474	2.4	7.5	14.5	29.7	0.41	14.3	19.2	7.5	2.3
Validation	2158	2.3	7.4	13.7	29,6	0.37	14,1	19.5	8.6	1.8
Test	2120	2.1	7.4	13.8	29.7	0.43	14.3	19.6	8.7	2.0

Table 2. Pearson correlation coefficients (r)between target and inputs[1] discriminated by training, validation and test sets.

		r coefficients for $ET0_{PM}$ and inputs of ANN							
Set	n	Rn	Rg	RTA	DPV	Ta	Tmax	Tmin	u
Training	6474	0.92	0.89	0.85	0.84	0.75	0.81	0.57	0.19
Validation	2158	0.95	0.93	0.87	0.81	0.75	0.81	0.57	0.14
Test	2120	0.94	0.92	0.87	0.83	0.77	0.81	0.61	0.13

[1]Units of correlated variables are the same as Table 1.

The selection of ANN architectures was based on the application of a selected algorithm integrated on the IPS (Intelligent Problem Solver) of the Neural Network module of Statistica Software [28]. The inputs and the outputs of data sets were automatically normalized to improve the performance of ANN models. Conjugate of retropropagation of errors algorithm, a second-order nonlinear optimization technique, was used in training process. The software provides two random methods for initializing the weights (normal and uniform distributions). The normal method, followed in this work, initializes the weights using normally distributed values, within a range whose mean is zero and standard deviation equal to one.The software also possibilities the application of random or bootstrap sampling different from the used in this paper (cross-validation with subset sampling). There was no intention in this work to evaluate such variations. However, in future research the application can be tested over the ANN with better performances.

Following the automated network search (ANS), the five models with the lowest cross-validation error were retained (over 2000 ANN for each combination of inputs) and then, the ANN with best performance for each combination was chosen and evaluated. However, weights and estimates from five better ANN models for each combination of inputs has been saved to future studies.

In Table 3 are described the ANNtrained to estimate daily values of $ET0_{PM}$ following: a) the variables used as input, b) the sequence n-m-x, where n is the number of inputs, m is the number of neuron at hidden layer and x is the number of outputs; c) the activation function; d) number of free parameters. The maximum number of neurons at hidden layer was fixed at 10. In this case, training was not limited by the example cases (n= 6474). Some of the ratios of the number of training sample to the number of connection weights cited in literature ranged from 2 to 30 [27]. Nevertheless, the ANN 16 with highest number of free parameters (51) did not present inconvenient with relation to examples cases.

A more reduced number of free parameters were needed to approximate the process if radiation (Rn, RTA or Rg) was combined with some driving variable of the aerodynamic component of evapotranspiration (DPV or u), whereas than the input of another variable of radiation component (Ta) resulted in ANN with more parameters, except for Rg.This can be explained due multicollinear variables require more sized structure in the network due the presence of mutual information.

Table 3. Description of artificial neural networks (ANN) trained to estimate daily values of reference evapotranspiration ($ET0_{PM}$) at Balcarce, Argentina.

ANN	Inputs	Structure	Activation in hidden layer	Number of free parameters
1	Rn Ta	MLP 2-7-1	Logistic	29
2	Rn DPV	MLP 2-3-1	Hyperbolic Tangent	13
3	Rn Ta DPV	MLP 3-6-1	Hyperbolic Tangent	31
4	Rn u	MLP 2-3-1	Hyperbolic Tangent	13
5	Rn Ta u	MLP 3-4-1	Hyperbolic Tangent	21
6	Rg Ta	MLP 2-3-1	Hyperbolic Tangent	13
7	Rg DPV	MLP 2-3-1	Logistic	13
8	Rg Ta DPV	MLP 3-8-1	Hyperbolic Tangent	41
9	Rg u	MLP 2-3-1	Logistic	13
10	Rg Ta u	MLP 3-6-1	Logistic	31
11	RTA Ta	MLP 2-7-1	Logistic	29
12	RTA DPV	MLP 2-4-1	Logitic	17
13	RTA Ta DPV	MLP 3-3-1	Logistic	16
14	RTA u	MLP 2-5-1	Logistic	21
15	RTA Ta u	MLP 3-3-1	Hyperbolic Tangent	16
16	RTA Tmax Tmin	MLP 3-10-1	Logistic	51

The evaluation of ANN performance to estimate daily values of reference evapotranspiration was based on comparison of their performance estimates from FAO-56 ($ET0_{PM}$). Multi-criteria analysis wasappliedwith basis on root mean of square error (RMSE), mean absolute error (MAE), mean bias error (MBE) and regression coefficients (a, b, R^2) between estimates from ANN and measured values. The Student test was used to statistically evaluate the value of either the intercept (H0: a=0) or slope of the straight line (H0: b=1) at the 5% probability level. To assess the capacity of generalization of the ANN, descriptions of performance are given over both validation and test sets.

From regression analysis and errors of estimation between outputs of the ANN and $ET0_{PM}$ values was possible to distinguish some combinations of variables with better performance. The analyses were carried out on both data sets (validation and test). In Table 4 are reported the results on validation set. The a and b parameters obtained by regression analyses between the target output and estimates from all ANN did not differ significantly from 0 and 1, respectively, being possible to infer that ET estimated from ANN did not differ from reference evapotranspiration ($ET0_{PM}$).

In general, the input of DPV provides a better performance, whichever the type of radiation used. The MAE values ranged from 0.2 to 0.6 mm d^{-1} were equivalent to 9 and 22% of observed mean values of validation series. Furthermore, the ANN models with DPV did not imply structures with high number in hidden layer. The combination of RTA with Tmax and Tmin did not improve the performance respect the model with DPV. The RTA was not input in the six best ANN of the group when ranked in function of minor RMSE. The difference in RMSE between the best ranked ANN with RTA (ANN13) was about 19% and 49% and RTA in comparison to their analogue models with Rg (ANN8) and Rn (ANN3), respectively. The last ANN ranked in function of minor RMSE (ANN14) increased 62% the RSME respect their analogue combination with Rn (ANN4). The RMSE increased 20% when RTA was combined with Ta instead Tmax and Tmin (ANN11 vs ANN16). A better explanation from this combination can be associated with the humidity description from difference in maximum and minimum temperatures, following to [26].

Table 4. Errors of estimation of the ANN trained to approximate daily reference evapotranspiration (ET_{0PM}) for the validation set (pairs of data=2120)

ANN	Model MLP	a mm d^{-1}	b	R^2	RMSE mm d^{-1}	MAE mm d^{-1}	MBE mm d^{-1}
1	$ET0_{PM}$(Rn Ta)	-0.0402	0.9333	0.93	0.4790	0.3848	0.2205
2	$ET0_{PM}$(Rn DPV)	0.0168	0.9734	0.96	0.3155	0.2231	0.0506
3	E $T0_{PM}$ (Rn Ta DPV)	-0.0115	0.9782	0.96	0.3045	0.2164	0.0669
4	$ET0_{PM}$ (Rn u)	-0.0302	0.9750	0.91	0.4825	0.3620	0.0939
5	$ET0_{PM}$(Rn Ta u)	0.0115	0.9549	0.93	0.4303	0.3205	0.1051
6	$ET0_{PM}$(Rg Ta)	-0.0318	0.9425	0.93	0.4571	0.3600	0.1849
7	$ET0_{PM}$(Rg DPV)	0.0230	0.9863	0.93	0.4280	0.3292	0.0111
8	$ET0_{PM}$ (Rg Ta DPV)	0.0052	0.9757	0.94	0.3819	0.2866	0.0564
9	$ET0_{PM}$ (Rg u)	-0.0033	1.0031	0.88	0.5413	0.4138	-0.0043
10	$ET0_{PM}$(Rg Ta u)	0.0002	0.9717	0.94	0.3923	0.2960	0.0719
11	$ET0_{PM}$(RTA Ta)	-0.0403	0.9239	0.83	0.7191	0.5287	0.2479
12	$ET0_{PM}$(RTA DPV)	-0.0657	0.9955	0.92	0.4596	0.3311	0.0772
13	$ET0_{PM}$ (RTA Ta DPV)	-0.0854	1.0088	0.92	0.4544	0.3312	0.0629
14	$ET0_{PM}$ (RTA u)	-0.0351	0.9506	0.76	0.7826	0.5540	0.1657
15	$ET0_{PM}$(RTA Ta u)	0.0132	0.9408	0.82	0.6858	0.4977	0.1421
16	$ET0_{PM}$(RTA Tmax Tmin)	0.0061	0.9303	0.87	0.5983	0.4505	0.1794

RMSE: root mean square error; MAE: mean absolute error; MBE: mean bias error.

In Table 5 the results on test set are reported. The accuracy of the model on the test data gives a realistic estimate of the performance of the model on completely unseen data and to confirm the actual predictive power of the network. The a and b parameters obtained by regression analyses between the target output and estimates from ANN did not differ significantly from 0 and 1, respectively, being possible to infer that ET estimated from ANN did not differ from reference evapotranspiration ($ET0_{PM}$).The same ANN ranking according RMSE values was maintained for test evaluation. The losses on generalization (RMSE of validation – RMSE of test) varied between 0 and 11%. The ANN 16 was the model that showed more decline in predictive power. In general, the input of DPV improved the performance, whichever the radiation used. The MAE values ranged from 0.2 to 0.6 mm d^{-1} were equivalent to 8 and 22% of observed mean values of test series. The combination of RTA with Tmax and Tmin did not improve the performance respect model with DPV.

The RTA was not input in the six best ANN of the group when ranked in function of minor RMSE. The difference in RMSE between the best ranked ANN with RTA (ANN13) was about 20% and 56% in comparison to their analogue models with Rg (ANN8) and Rn (ANN3), respectively. In addition, the last ANN ranked in function of minor RMSE (ANN14) increased 52% the RMSE respect their analogue combination with Rn (ANN4), similarly to reported on validation set. The RMSE increased 19% when RTA was combined with Ta instead Tmax and Tmin (ANN11 vs ANN16). A better explanation from this combination can be associated with humidity description from difference in maximum and minimum temperatures.

If daily u values are missing, the estimations from DPV based ANN are more recommended than those from temperature-based models. For the other hand, if daily DPV values are missing, the estimations from temperature models are more recommended than those from u. In cases without Rn and Rg values, the application of ANN with RTA is suggested, but ever with DPV as input.

In estimations with basis on linear multiple regression that were previously adjusted at Balcarce [2], from RTA in combinations with Tmin and Tmax and precipitation, the MAE values ranged between 0.51 and 0.65 mm d^{-1}. However, the estimations for averaged 10-day values were between 0.25 and 0.35 mm d^{-1}. Some improvements have been reported [4] when RTA and DPV were used (MAE= 0.36 mm^{-1}). In this work, the ANN 12 (RTA and DPV) reduced in 40% the errors of estimation in comparison to that errors obtained from regression models previously adjusted [4]. On the other hand, ANN 16 showed minor errors for daily values than the Hargreaves method (same inputs) tested to estimate monthly values [3].

Table 5. Errors of estimation of the ANN trained to approximate daily reference evapotranspiration (ET_{0PM}) for the test set (pairs of data=2158)

ANN	Model MLP	a	b	R^2	RMSE mm d^{-1}	MAE mm d^{-1}	MBE mm d^{-1}
1	$ET0_{PM}$(Rn Ta)	0.1933	0.9764	0.92	0.4798	0.3814	-0.1323
2	$ET0_{PM}$(Rn DPV)	0.0706	0.9936	0.97	0.3075	0.2318	-0.0539
3	E $T0_{PM}$ (Rn Ta DPV)	0.0987	0.9888	0.97	0.2973	0.2237	-0.0697
4	$ET0_{PM}$ (Rn u)	0.2693	0.9014	0.89	0.5243	0.3999	-0.0146
5	$ET0_{PM}$(Rn Ta u)	0.1564	0.9591	0.93	0.4396	0.3404	-0.0509
6	$ET0_{PM}$(Rg Ta)	0.1727	0.9735	0.92	0.4728	0.3728	-0.1042
7	$ET0_{PM}$(Rg DPV)	0.1601	0.9469	0.93	0.4325	0.3388	-0.0229
8	$ET0_{PM}$ (Rg Ta DPV)	0.1321	0.9690	0.94	0.3872	0.2962	-0.0520
9	$ET0_{PM}$ (Rg u)	0.3176	0.8496	0.86	0.6057	0.4635	0.0709
10	$ET0_{PM}$(Rg Ta u)	0.1340	0.9573	0.93	0.4197	0.3192	-0.0237
11	$ET0_{PM}$(RTA Ta)	0.4885	0.8777	0.82	0.7108	0.5192	-0.1725
12	$ET0_{PM}$(RTA DPV)	0.2719	0.9354	0.92	0.4653	0.3377	-0.1050
13	$ET0_{PM}$ (RTA Ta DPV)	0.2815	0.9268	0.92	0.4634	0.3390	-0.0925
14	$ET0_{PM}$ (RTA u)	0.6700	0.7782	0.76	0.7939	0.5858	-0.0971
15	$ET0_{PM}$(RTA Ta u)	0.4441	0.8658	0.82	0.6917	0.5086	-0.0974
16	$ET0_{PM}$(RTA Tmax Tmin)	0.3255	0.9005	0.86	0.5987	0.4456	-0.0684

RMSE: root mean square error; MAE: mean absolute error; MBE: mean bias error.

Errors reported in this works are yet lightly larger than from ANN under other environmental conditions de ET0 and driving variables [15][19]. Further efforts could be made in future to regard components of seasonality from easily available variables, such as maximum sunshine hours (from Julian data) as suggested by [15] [21]. Losses of estimation from ANN with Rn and Rg as input to ANN with basis in RTA are not different from reported previously in anotheratmospheric environment [19]. In addition, networks of type BRF can be regarded with limited data [21]. In this case, deviation of estimates ranged from -1 a 0.1 %from Penman-Monteith values.

5Extraction of knowledge from ANN models

Once the ANN were trained on a specific network topology, then the modeling of attributes process using ANN involved the extracting knowledge from each network. The embedded knowledge is in the form of connection weights. Garson's method [7] was performed from adjusted synaptic weights of each ANN. The contribution of each input neuron to the output (c_{ijo}) was computed via each hidden neuron as the product of the input-hidden connection (w_{ij}) and the hidden-output connection (w_{jo}):

$$c_{ijo} = w_{ij} \times w_{jo} \qquad \textbf{(2)}$$

The relative contribution of each input k to hidden neuron j can be expressed as:

$$r_{ijo} = \frac{|C_{ijo}|}{\sum_{K=1}^{m} |C_{kjo}|} \quad \textbf{(3)}$$

The total contribution of input i is:

$$S_i = \sum_{J=1}^{n} r_{ijo} \quad \textbf{(4)}$$

Finally, the relative contribution of each input is:

$$RI = S_i / \sum_{k=1}^{m} Sk \quad \textbf{(5)}$$

In Table 6 is presentedthe relative contribution (RI) of the inputs to each ANN. Despite the importance of radiation component in reference evapotranspiration values from Penman-Monteith method [23], the contribution of Rn was not predominant in the models tested with reduced number of variables (ANN1 to ANN5). The relative contribution of aerodynamic components (DPV and u) was similar when Rn was regarded in input (ANN2 and ANN4), but did not for models with Rg (ANN 7 vs ANN9) or RTA (ANN12 vs ANN14). When Ta was input with DPV or u, the RI values of Rn decreased (ANN3 and ANN5). In general, the contribution relative of Rg tended to increase in each model (ANN6 to ANN10) respect the same combination with Rn and other variables (ANN1 to ANN5). It was conspicuous the contribution the one variable to model in ANN12 (RTA) and ANN15 (u). The RI of RTA was minor when air temperature was input as maximum and minimum daily values than the average value (ANN16 vs ANN11).

Table 6. Relative contribution of inputs (RI) to neural network to approximate daily reference evapotranspiration ($ET0_{PM}$).

					RI				
ANN	Inputs	Rn	Rg	RTA	Ta	Tmax	Tmin	DPV	u
1	Rn Ta	0.36			0.64				
2	Rn DPV	0.41						0.59	
3	Rn Ta DPV	0.16			0.47			0.37	
4	Rn u	0.45							0.55
5	Rn Ta u	0.21			0.41				0.38
6	Rg Ta		0.39		0.61				
7	Rg DPV		0.46					0.54	
8	Rg Ta DPV		0.41		0.20			0.39	
9	Rg u		0.61						0.39
10	Rg Ta u		0.42		0.38				0.20
11	RTA Ta			0.55	0.45				
12	RTA DPV			0.80				0.20	
13	RTA Ta DPV								
14	RTA u			0.33					0.67
15	RTA Ta u			0.11	0.06				0.83
16	RTA Tmax Tmin			0.24		0.48	0.28		

Sensitivity analysis was performed following procedures from the Neural Network module in STATSOFT [28] The program test how the neural network response (predictions) and, hence, the error rates would increase or decrease if each of inputs variables were to undergo a change. The data set is submitted to the network repeatedly, with each variable in turn replaced with its mean value from the training sample, and the resulting network error is recorded. The trained ANN were more sensitive to radiation than to the other variables (Table 7). Besides, the models performed with RTA (ANN 11 to ANN16) were less sensitive to radiation than those with Rg and Rn (ANN1 to ANN10). Within group of radiation based ANN, the estimation of $ET0_{PM}$ resulted more sensitive to DPV than Ta or u for models with two inputs (ANN2 vs ANN1 and ANN4, ANN7 vs ANN6 and ANN9, ANN12 vs ANN11 and ANN14).

Table 7. Relative sensitivity of each artificial neural network trained to approximate daily reference evapotranspiration ($ET0_{PM}$) to each input parameter.

		Sensitivity to input (fraction)							
ANN	Inputs	Rn	Rg	RTA	Ta	Tmax	Tmin	DPV	u
1	Rn Ta	0.79			0.21				
2	Rn DPV	0.68						0.32	
3	Rn Ta DPV	0.64			0.06			0.30	
4	Rn u	0.87							0.13
5	Rn Ta u	0.70			0.18				0.12
6	Rg Ta		0.72		0.28				
7	Rg DPV		0.63					0.37	
8	Rg Ta DPV		0.58		0.12			0.30	
9	Rg u		0.85						0.15
10	Rg Ta u		0.64		0.24				0.12
11	RTA Ta			0.68	0.32				
12	RTA DPV			0.55				0.45	
13	RTA Ta DPV			0.48	0.10			0.42	
14	RTA u			0.78					0.22
15	RTA Ta u			0.55	0.27				0.18
16	RTA Tmax Tmin			0.46		0.39	0.15		

6Conclusionsand future work

This study evaluated the performance of 16 ANN models to approximate reference evapotranspiration against the standard Penman Monteith method, under the climatic conditions in the southeastern region of rolling pampas, Argentina. The case is analyzed to illustrate the use of the neural network technique and demonstrate its capabilities of effectively analyzing and predicting the reference evapotranspiration at Balcarce.

Air temperature and deficit of pressure vapor were found to be more effective than wind velocity in modelling ET0, whichever the radiation (Rn, Rg or RTA) used as input.The results give helpful data and documentation to choose strategies to select the more exact ET0 estimations under limiting data conditions. Future efforts could be attempt to model the cases without radiation, via either adding seasonal components or by using adifferent type of network from MLP. In addition, further studies from the region may be required to reinforce the conclusions drawn from this study.

A description of the knowledge that was learned by the ANN during their training was obtained by applying simple knowledge extraction methods. An advantage of this method is that additional information about the model performance is obtained, including the relative contribution of inputs via analysis of connection weights in the ANN.

The developed ANN models are useful to the precise agricultural water management, regional water resources planning, and other hydrological modeling related studies that can aid in more proficient and viable water resources management. Furthermore, techniques of knowledge extraction could be carried out in further studies to determine the types of problems where artificial neural networks would yield better results than other methods. The results reported here also contribute to coping with problems of scarce or missing data and thus can be used to guide priorities for data acquisition.

Acknowledgements

The author thanks the financial support from Universidad Nacional de Mar del Plata.

References

1. Allen, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration. Guides for computing crop water requirements. FAO Irrig. Drain. Paper 56. FAO, Rome, Italy (1998)
2. Irigoyen, A.I., Della Maggiora, A.I., Gardiol, J.M.: Evapotranspiración de referencia a partir de modelos con mínimo requerimiento de datos: Estimación de valores diarios. Proc. XIII Congresso Brasileiro de Agrometeorología. Santa María, Brasil (2003)
3. Irigoyen, A.I., Della Maggiora, A.I., Gardiol, J.M.: Evapotranspiración de referencia a partir de modelos con mínimo requerimiento de datos: Estimación de valores medios mensuales. Proc. XIII Congresso Brasileiro de Agrometeorología. Santa María, Brasil (2003)
4. Irigoyen, A.I., Della Maggiora, A.I., Gardiol, J.M.: Estimación de la evapotranspiración de referencia a partir de un modelo sencillo basado en el déficit de vapor para la localidad de Balcarce. (Arg.). X Reunión Argentina de Agrometeorología, Mar del Plata, Argentina (2004)
5. Haykin, S.: Redes Neurais: princípios e prática. Trad. M. Engel. Porto Alegre. Bookman (2001)
6. Braga A.P., Carvalho, A.P.L., Ludermir, T. B.: Redes Neurais Artificiais: Teoria e aplicações. Rio de Janeiro: LCT editora (2007)
7. Garson, G.D.: Interpreting neural network connection weights. Artificial Intelligence Expert 6 (1991) 47-51
8. Saltelli, A.: Sensitivity analysis for importance assessment. Risk Analysis 22 (2002) 579-590
9. Gevrey, M., Dimopoulos, I. Lek, S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol. Model 160 (2003) 249-264
10. Odhiambo, L. O., Yoder, R.E., Yoder, D.C. Hines, J.W.: Optimization of fuzzy evapotranspiration model through neural training with input-output examples. Trans. ASAE 44(2001) 1625-1633
11. Kumar, M., Raghuwanshi, N.S., Singh, R. Wallender,R.R., Pruitt, W.O.: Estimating evapotranspiration using artificial neural network. Journal of Irrigation andDrainage Engineering, 128 (2002) 224-233
12. Silva, A. F.; Costa, L. C.; Sediyama, G. Previsão da evapotranspiração de referência utilizando redes neurais. Engenharia na Agricultura 14 (2006)93-99
13. Trajkovic, S.; Todorovic, B.; Stankovic, M.: Forecasting reference evapotranspiration by artificial neural networks. Journal of Irrigation and Drainage Engineering 129(2003) 454-457
14. Arca, B., Beniscasa, F., Vincenzi, M.: Evaluation of neural network techniques for estimating evapotranspiration. Consultado en: <http://server.ss.ibimet.cnr.it/~arca1/papers>(2004)
15. Zanetti, S.S, Sousa, E.F, Oliveira, V.P.S., Almeida, F., Bernardo, S.: Estimating evapotranspiration using artificial neural network and minimum climatological data. Journal of Irrigation and Drainage Engineering 133 (2007) 83-89
16. Landeras G, Ortiz-Barredo A, Lopez JJ.: Comparison of artificialfor daily reference evapotranspiration estimation in the BasqueCountry (Northern Spain). Agricultural Water Management 95 (2008) 553–565.
17. Jain, S. K., Nayak, P.C., Sudheer, K.P.: Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation Hydrol. Process. 22 (2008) 2225-2234
18. Rahimikhoob, A.: Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. Theor. Appl. Climatol. 101 (2009) 83-91
19. Irigoyen, A.: Modelagem da evapotranspiracão de referência e da evapotranspiracão de limeira acida com aplicação de técnicas de regressão e redes neurais artificiais, Tese Doutorado USP (2010)
20. Sudheer, K. P., Gosain, A. K., and Ramasastri, K. S.:Estimating actual evapotranspiration from limited climatic data using neural computing technique. J. Irrig. Drain. Eng. 129 (2003) 214-218.
21. Trajkovic, S.: Temperature-based approaches for estimating reference evapotranspiration. J. Irrig. Drain. Eng. 131 (2005) 316-323.
22. Irigoyen, A.I., Della Maggiora, A.I.: Knowledge extraction from artificial neural networks: case study on total soil water during maize crop season. Borges de Campos, A.; Gabriels, D.; Luján, D.L.(ed.) Impacts of Agroecosystems on Environment. CH. I. Environmental sustainability of agrosystems. UNESCO Chair of Eremology, Ghent University, Belgium and by International Centre of Eremology (ICE), Ghent University, Belgium. (2015) 29-34
23. Irigoyen, A.I., Della Maggiora, A.I., Gardiol, J.M.: Variabilidad intra e interanual de la evapotranspiración de referencia estimada según el modelo Penman-Monteith en la localidad de Balcarce (Argentina). Revista Brasileira de Agrometeorologia 14 (2006) 85-96
24. Cybenko, G.: Approximations by superpositions of a sigmoidal function. Mathematical of Control, Signal and Systems 2 (1989) 303-324
25. Gardiol, J.M., Della Maggiora, A.I., Irigoyen, A.I.: Estimación de la radiación global para la localidad de Balcarce. Proc. VII Congreso Latinoamericano e Ibérico de Meteorología. Buenos Aires (1996)
26. Hargreaves, G.H., Samani, Z. A.: Estimating potential evapotranspiration. J. Irrig. and Drain Engr., ASCE, 108 (1982) 223-230
27. Maier, H., Dandy, G.C.: Neural network based modelling of environmental variables: A systematic approach. Mathematical and Computer Modelling 33(2001) 669-682

28. Statsoft. Statistica Neural Networks Module (computer program). Vers. 9. Statsoft, Inc. Tulsa, 2009. 1 CD-ROM (2009)