XVII^th World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)
Québec City, Canada June 13-17, 2010

DEVELOPMENT OF REFERENCE EVAPOTRANSPIRATION MODELS WITH LIMITED DATA BY USING ARTIFICIAL NEURAL NETWORKS STRUCTURES

A.MARTÍNEZ[1], M.A. MORENO[2], J.F. ORTEGA[3], J.M. TARJUELO[4], J.A. DE JUAN[5]

[1] Regional Centre of Water Research (CREA), Universidad de Castilla-La Mancha, Ctra. de Las Peñas km. 3.4, 02071 Albacete, Spain. Tel: +34 967 599 304; Fax: +34 967 599 269, angel.mromero@uclm.es
[2] M.A. Moreno, miguelangel.moreno@uclm.es
[3] J.F. Ortega, jose.ortega@uclm.es
[4] J.M. Tarjuelo, jose.tarjuelo@uclm.es
[5] J.A. de Juan, arturo.juan@uclm.es

CSBE1050 – Presented at Section I: Land and Water Engineering Conference

ABSTRACT The analysis of evapotranspiration (ET) measured at ground level has been mainly considered for agricultural activities, with the aim of determining the water requirements and the supply of water in an appropriate manner. Irrigation activity uses 70% of freshwater in the world, which denotes the importance of this use. ET_0 calculation is a key factor for water balance and crop production, water resource management, irrigation scheduling, and environmental assessment. ET_0 can be obtained with indirect measurements, with high-cost micrometeorological techniques, or can also be estimated with mathematical models. The Food and Agriculture Organization of the United Nations (FAO) has proposed the use of the Penman-Monteith equation as the standard for estimating ET_0 and for evaluating other equations. However, some of the weather variables that are required by this model are often not available in irrigation. In such cases, a simple empirical Hargreaves equation is mainly used. The main objective of this paper is the generation of a model that estimates monthly ET_0 values in Castilla-La Mancha region (Spain) from available termo-pluviometic weather stations with monthly data that can be applied to historical time series. The procedures to estimate the ET_0 values in this study are in one hand, the use of Hargreaves equation, which requires the maximum and the minimum monthly temperature differences, together with extraterrestrial global solar radiation data; and on the other hand, the application of Artificial Neural Network (ANN) structures with the same climate data. Calibration and validation of the ANN models are performed by using data from the years 2000 to 2008 of 44 complete agroclimatic weather stations. The monthly average of the Penman-Monteith daily ET_0 is compared with the monthly values of ET_0 using Hargreaves and the ANN models. The obtained relative errors are: 1) spatial error (spatial validation) of 17.0% and 13.21% for Hargreaves equation, and ANN model respectively, 2) temporal errors (temporal validation) of 13,8% and 12,0% for Hargreaves equation, and ANN model respectively.

Keywords: ET_0, limited data, Hargreaves, Artificial Neural Networks (ANN).

INTRODUCTION The climate has always had a very important role on human's health. Actually, it has determined the location of agricultural production sites, recreational areas, urban development, and industrial areas. The availability of different climatic parameters permits to characterize the climate of a region, being necessary several climate data to carry out a complete study. Several studies have been based on the use of different climate data or indicators calculated by using basic climatic parameters. However, these studies usually present limitations on their development due to problems with data availability and quality (Elías and Ruiz, 1981; de León, 1988; Allen *et al*., 1998; Fount, 2000).

Different authors have proposed algorithms and techniques to perform data generation in different areas of knowledge (Allen *et al*., 1998; Allison, 2001). Some of them have tried to estimate climatic parameters, such as global radiation and reference evapotranspiration (ET_0) from basic climatic data or ANN structures (Popova *et al*., 2006; Landeras *et al*., 2008; Gavilán *et al*., 2008).

The analysis of evapotranspiration (ET) measured at ground level has been mainly considered for the agricultural activity, in order to determine the water requirements and the supply of water in an appropriate manner (Calera, 2005). Irrigation activity uses 70% of freshwater in the world, which denotes the importance of this use (FAO, 2002). Accurate estimates of reference evapotranspiration (ET_0) are required for responsible water use in irrigation (Trajkovic, 2007). ET_0 determination is a key factor for water balance and crop production, water resources management, irrigation scheduling, and environmental assessment. ET_0 can be obtained with indirect measurements, with high-cost micrometeorological techniques, or can also be estimated with mathematical models (Allen *et al*., 1991). The Food and Agriculture Organization of the United Nations (FAO) has proposed the use of the Penman-Monteith equation as the standard for estimating ET_0 and for evaluating other equations (Trajkovic, 2007). However, some of the weather variables that are required by this model are often not available in irrigation. In such cases, a simple empirical Hargreaves equation is mainly used (Popova *et al*., 2006; Trajkovic, 2007).

The main objective of this paper is the generation of a model that estimates monthly ET_0 values in Castilla-La Mancha region (Spain) from available termo-pluviometic weather stations with monthly data that can be applied to historical time series.

MATERIAL AND METHODS The proposed methodology can be summarized in the Figure 1. The first step consists in selecting the termo-pluviometic data that best estimate the ET_0. From 44 complete weather stations, and with time series from 2000 to 2008, data quality tests are applied. Then, 40 weather stations and time series from 2000-2007 are selected for ANN training. With the remainder 4 stations and the year 2008 data, spatial and temporal validation is performed. Thus, the proper ANN structure can be obtained by using surface polts.

The case study The case study is located in C-LM Region (Spain) (Fig. 2), which is a semiarid region with an extension of 79,462 km^2. In C-LM Region, 132 weather stations are available with historical time series of monthly average, maximum, and minimum temperature data (Tav, Tmax, and Tmin), together with precipitation data, which would be useful to characterize the local climate. In addition, in this Region there are 44 complete agroclimatic weather stations (Fig. 2) which are located in irrigated areas and are incluyed in the Agroclimatic Information Service for Irrigation (SIAR, "Servicio de Información Agroclimática para el Regadío") network. Generation of the ANN models to estimate ET_0 from monthly temperature data is performed by using these data. In addition the Figure 2 shows distribution of annual average temperature in the Region.

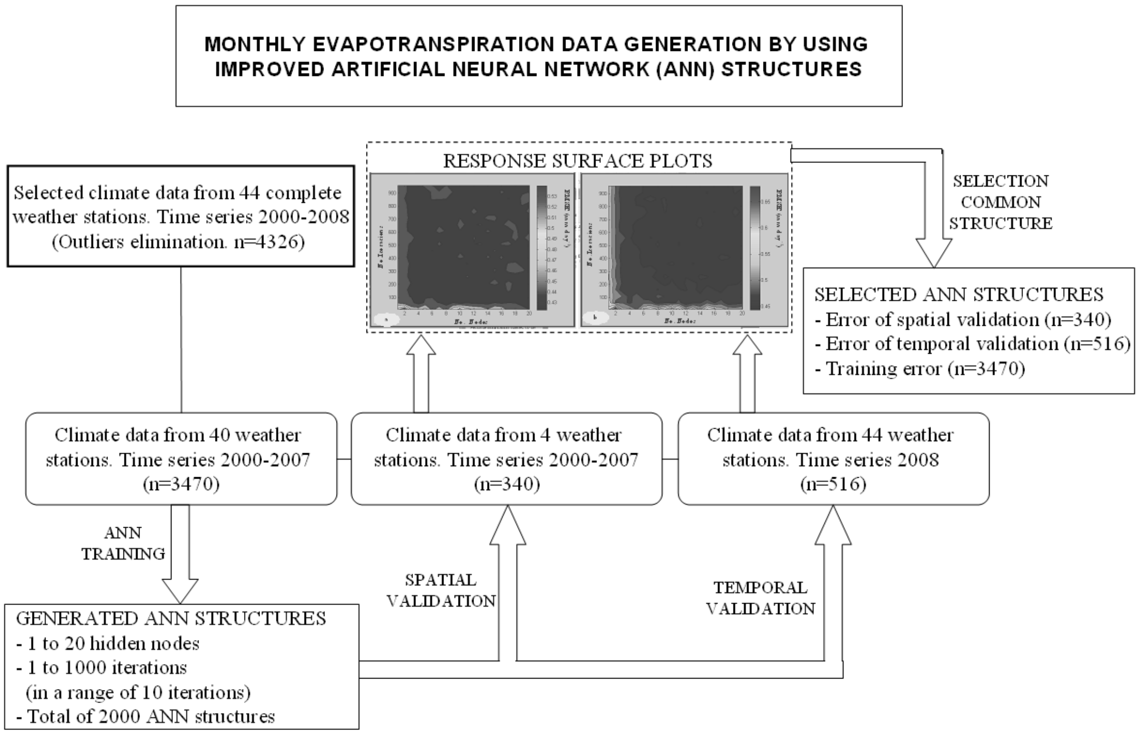

Figure 1. Diagram of the proposed methodology.

Calibration and validation of the Artificial Neural Network models In the proposed methodology, Artificial Neural Networks (ANNs) are used (Schalkoff, 1997; Nabney, 2002) to generate ET_0 in the C-LM Region by using temperature, precipitation, and extraterrestrial solar radiation data (Ra) (values for Ra on the 15^{th} day of the month, Allen *et al.*, 1998). The general structure of the used ANN is shown in Figure 3. Three layers are defined by three groups of neurons. Those layers are: input, hidden, and output nodes. Each neuron implements a local function. The input nodes implements a linear function (Eq. (1)), the hidden nodes a nonlinear function (Eq. (2)) and the output nodes implements a linear function (Eq. (4)) (Fig. 3). The equation (1) shows the activation function (weighted sum) and equation (2) the squashing function (hyperbolic tangent). In this case, in which a regression problem is tried to solve, the hidden to output layer is used the linear activation function (Eq. (3)) (Fig. 3) (Nabney, 2002).

$$S_j = w_{0j} + \sum_{i=1}^{N} w_{ij} \cdot x_i \qquad (1)$$

$$y_j = \tanh(S_j) \qquad (2)$$

$$S_k = v_{0k} + \sum_{j=1}^{H} v_{jk} \cdot y_j \tag{3}$$

Thus, the result of the ANN is described by Eq. (4). In this case (regression problem), $Z_k = S_k$.

$$Z_k = f(S_k) \quad \text{for k= 1…M} \tag{4}$$

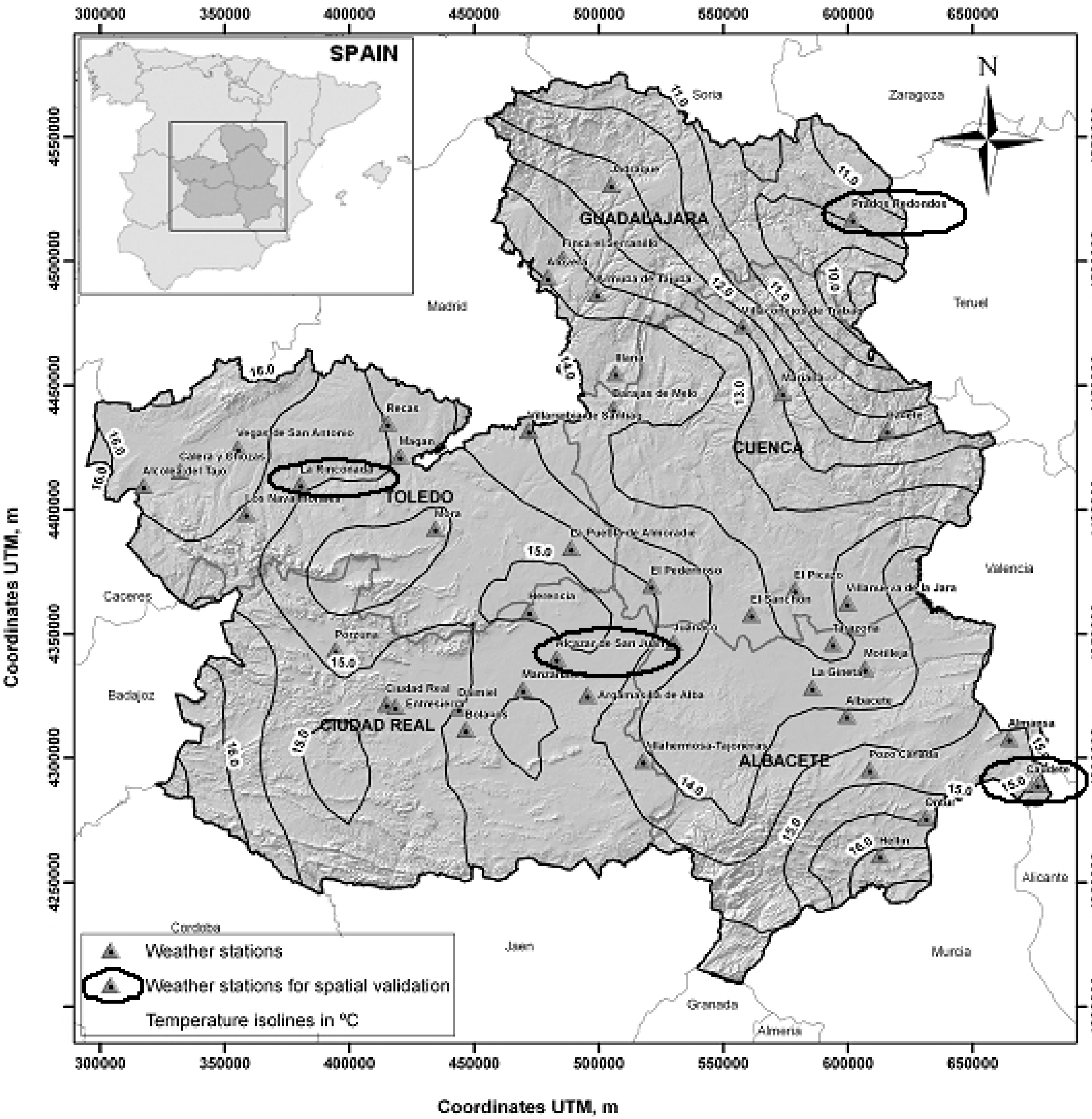

Figure 2. Location of the available and utilized complete weather stations in the Castilla-La Mancha (C-LM) Region during the time series 2000-2008. Distribution of annual average temperature.

The ANNs were trained (calibration process) under supervision using scaled conjugate gradient algorithm with the main objective of minimizing the error of the estimation

(Nabney, 2002). The input layer is composed by basic climatic data such as temperature (average, maximum, and/or minimum temperature), and extraterrestrial solar radiation data. Two thousand ANN structures with different number of nodes (1 to 20) and with different number of iterations (1 to 1,000 in a range of 10) were trained with monthly data of 40 of the 44 SIAR weather stations for the period covering the years 2000 to 2007 (Figs. 2 and 4). The validation process consists on applying these 2,000 ANN structures calibrated previously to a dataset and study the generated errors. In order to decide the improved ANN structure, response surface plots are used to detect, graphically, the areas with the lowest error in both graphs. Thus the validation of the ANN models are performed from two points of view (Figs. 1 and 2): 1) spatial validation (applying the calibrated ANN structures to 4 weather stations for period 2000- 2007) and 2) temporal validation (applying the calibrated ANN structures to 44 weather stations for period 2008). Data used in validation process were not used in calibration process.

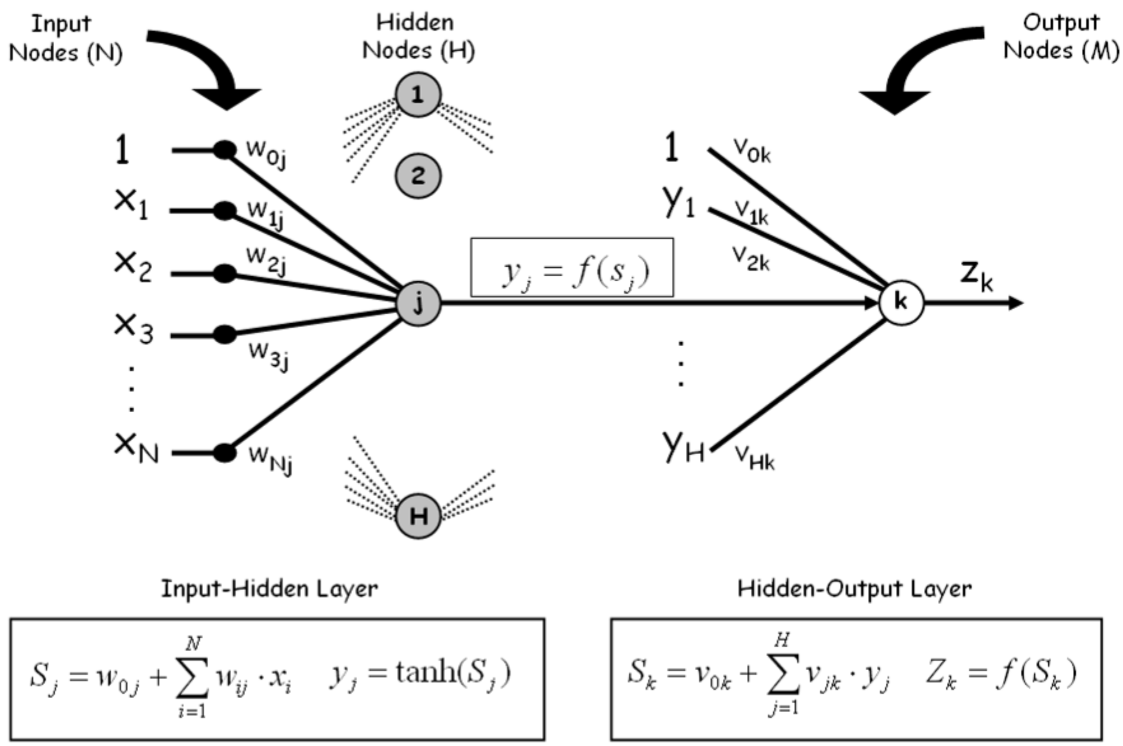

Figure 3. Utilized Artificial Neural Network structure and related equations.

Goodness of fit of the ANN models. To determine the goodness of fit or the ANN models and Hargreaves equation (eq. (5)) different statistic parameters (Willmott, 1982) are used such as: coefficient of determination (R^2), Root Mean Square Error (RMSE), Relative Error (RE), and the Similarity Rate (SR) (eqs. (6) to (8). In both cases, reference values (observed variable) are the averages monthly ET_0 values that have been calculated by applying Penman-Monteith equation with daily date.

The Hargreaves method calculate the ET_0 values by using temperaturas and Ra values, which depends of the latitud (eq. (5)):

$$ET_0 = 0{,}0023 R_a \left(T_{av} + 17{,}8\right)\sqrt{(T_{max} - T_{min})} \qquad (5)$$

where: ET_0, is the referente evapotranspiration ($mmday^{-1}$); R_a, is the extraterrestrial solar radiation data ($MJm^{-2}day^{-1}$); T_{av}, is the monthly average temperature average (ºC); T_{max}, is the monthly maximum temperature average (ºC); T_{min}, is the monthly minimum temperature average (ºC).

$$RMSE = \left[n^{-1} \sum (P_i - O_i)^2\right]^{1/2} \qquad [6]$$

where: RMSE is the root mean square error; n, is the number of observations; P_i and O_i are the predicted and observed values, respectively.

$$RE = (RMSE / O_{med}) \cdot 100 \qquad [7]$$

where: RE, is the relative error, estimated as a percentage of the average value of the variable; O_{med}, is the average value of the variable observed.

$$SR = 1 - \left[\sum (P_i - O_i)^2 / \sum \left((P_i - O_{med}) + (O_i - O_{med})\right)^2\right] \qquad [8]$$

where: SR, is the similarity rate expressed as a relative measure of the difference between variables. (if SR = 1, there is a perfect agreement between P and O).

RESULTS The set of variables (termo-pluviometric variables) that permits to obtain a good fitting with the ET_0 data by using ANN structures are Tmax, Tmin, and Ra. The response surface plots show the RMSE of all ANN structures (2,000 in this case) in function of the number of iterations and hidden nodes. As an example, Figure 4 show the response surface plots of the spatial (Fig. 4a) and temporal (Fig 4b) validation of the estimation of ET_0 by using as input variables Tmax, Tmin, and Ra. It can be shown that the common minimum RMSE in the estimation of ET_0 is obtained with a structure of between 8-16 hidden nodes and between 200-600 iterations.

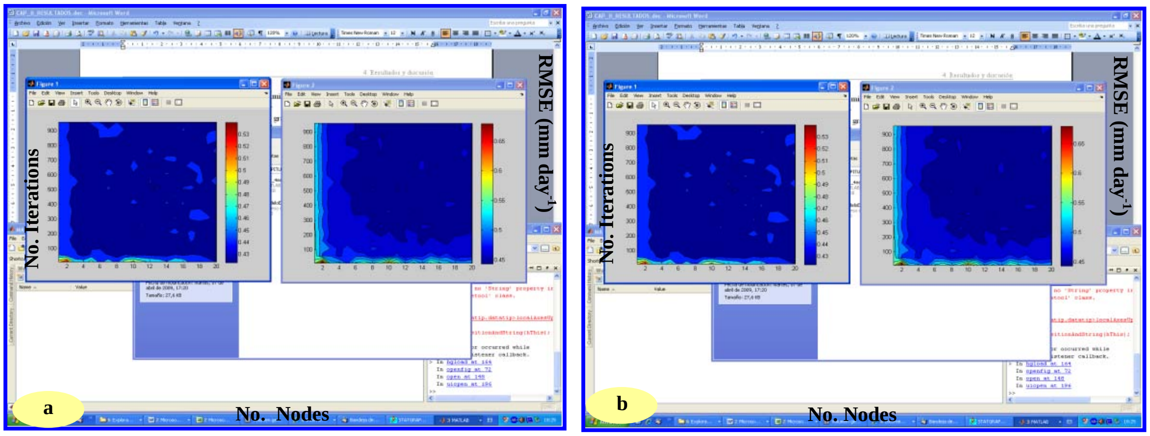

Figure 4. Response surface plots of the evolution of the root mean square error (RMSE) based on Artificial Neural Network structure used for estimating the evapotranspiration of reference in Castilla-La Mancha (C-LM): a) spatial validation for 4 weather stations (time series 2000-2007), b) temporal validation for 2008 year (44 weather stations).

The selected structure for estimating ET_0 is defined by 12 nodes and 500 iterations. Table 1 summarize the statistic parameters analyzed after the estimation of ET_0 data, for spatial and temporal validation, from average monthly Tmax, Tmin, and Ra data by using this structure. In addition this Table shows the statistical parameters by using Hargreaves equation. The error (RMSE) is lower by using ANN *vs*. Hargreaves model. For spatial validation data the RMSE values are 0.40 mm day^{-1} and 0.46 mm day^{-1}, respectively. For temporal validation these values are 0.42 mm day^{-1} and 0.55 mm day^{-1}. ANN shows ERs values of 12.02 % and 13.21 % while in Hargreaves equation they are 13.82 % and 16.99 % for spatial and temporal validation data.

Table 1. Estimated monthly evapotranspiration of reference (ET_0) in Castilla-La Mancha (C-LM) by using selected ANN structure (12 nodes and 500 iterations) from averages of maximum and minimum temperatures and extraterrestrial solar radiation, and, by using Hargreaves equation. Error of spatial and temporal validation.

Statistical parameters	Artificial Neural Network (ANN)		Hargreaves equation	
	Spatial validation	Temporal validation	Data utilized by spatial validation in ANN	Data utilized by temporal validation in ANN
Data number	340	516	340	516
Average x	3.34	3.21	3.34	3.21
Average y	3.28	3.22	3.49	3.38
Average x /average y	0.98	1.00	1.04	1.05
R^2	0.96	0.95	0.95	0.93
R^2 adjust	0.96	0.95	0.95	0.93
RMSE	0.40	0.42	0.46	0.55
RE	12.02	13.21	13.82	16.99
SR	0.99	0.99	0.99	0.98

where: x, are observed ET_0 values (mm day^{-1}); y, are predicted ET_0 values (mm day^{-1}); R^2, is determination coefficient; RMSE, is root mean square error (mm day^{-1}); RE, is relative error (%) and SR, is similarity rate.

CONCLUSION. The use of ANNs could be interesting for the generation of climate parameters from other available climatic data. For determining the optimal ANN structure in the estimation of ET_0 it is necessary to calibrate the model with part of the available data and validate the results with climatic data from stations not used in the calibration process of the ANN, considering spatial and temporal characteristics.

In C-LM Region, the estimated ET_0 data from basic temperature and Ra monthly data, were estimated with relative errors of: 1) spatial error of 12.02% 2) temporal errors of 13.21%. Hargreaves equation shows higher errors than the spatial and temporal errors obtained with the validation process of the ANN models, which are estimated in 13.82% and 16.99% respectively.

In C-LM Region, with limited data, ET_0 could be estimated by using ANN structures. This methodology shows better goodness of fit than Hargreaves equation.

Acknowledgements. We would like to thank to the CICYT for funding the national project (AGL2007-66716-C03-03), and the Education Regional Government of C-LM for funding the project (PCI08-0117) in which this work has been framed.

REFERENCES

Allen, R.G., L.S. Pereira, D. Raes and M. Smith, 1998. Crop Evapotranspiration. Guidelines for Computing Crop water Requirements. FAO Irrigation and Drainage paper No. 56. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

Allen, R.G., W.O. Pruitt and M.E. Jensen 1991. Enviromental requirements for lysimeter. In: R.G. Allen, W.O. Pruitt, T.A. Howell, I.A. Walter and M.E. Jensen (eds.). Lysimeters for Evapotranspiration and Environmental Measurements (170-181). American Society of Civil Engineers (ASCE), New York, NY, USA.

Allison, P.D. 2001. Missing Data. Series: Quantitative Applications in the Social Sciences. Sage Publications, Inc., University of Pennsylvania, Philadelphia, PA, USA.

Calera, A. 2005. La evapotranspiración: concepto y metodología de cálculo. En: F.J. Martín de Santa Olalla, P. López Fuster and A. Calera (eds.). Agua y Agronomía (163-238). Mundi-Prensa, Madrid, España. [In Spanish].

de la Casa, A., G. Ovando and A. Rodríguez 2003. Estimación de la radiación solar global en la provincia de Córdoba, Argentina, y su empleo en un modelo de rendimiento potencial de papa. Revista de Investigaciones Agropecuarias (RIA) 32, 45-62. [In Spanish].

de León, A., A. Arriba and M.C. de La Plaza 1988. Caracterización Agroclimática de la Provincia Albacete. Ministry of Agriculture, Fisheries and Food, Technical Secretary-General, Madrid, Spain. [In Spanish].

Elías, F. and L. Ruiz 1981. Estudio Agroclimático de la Región de Castilla-La Mancha. Department of Agricultural of “Junta de Comunidades de Castilla-La Mancha”, Toledo, Spain. [In Spanish].

FAO 2002. Agua y cultivos. Logrando el uso óptimo del agua en la agricultura. Organización de las Naciones Unidas para la Agricultura y Alimentación (FAO), Roma, Italia. [In Spanish].

Fount, I. 2000. Climatología de España y Portugal (2ª Ed.). University of Salamanca, Salamanca, Spain. [In Spanish].

Gavilán, P., J. Estévez and J. Berengena 2008. Comparison of Standardized Reference Evapotranspiration Equations in Southern Spain. J. Irrig. Drainage Eng-ASCE 134, 1-11.

Landeras, G., A. Ortiz and J.J. López 2008. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain) Agric. Water Manage. 95, 553 – 565.

Nabney, I.T. 2002. Netlab. Algorithms for Patter Recognition (4ª Ed.). Springer Verlag, London, United Kingdom.

Popova, Z., M. Kercheva and L.S. Pereira 2006. Validation of de FAO methodology for computing ET_0 with limited data. Application to South Bulgaria. Irrig. and Drain. 55, 201-215.

Schalkoff, R.J. 1997. Artificial Neural Networks. McGraw Hill, London, United Kingdom.

Trajkovic, S. 2007. Hargreaves *vs.* Penman-Monteith under Humid Condition. J. Irrig. Drainage Enage. Eng-ASCE 133, 47-52.