Expert Systems with Applications 38 (2011) 14190–14195

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Modelling net radiation at surface using "*in situ*" netpyrradiometer measurements with artificial neural networks ☆

Antonio Geraldo-Ferreira [a,b], Emilio Soria-Olivas [c], Juan Gómez-Sanchis [c,*], Antonio José Serrano-López [c], Almudena Velázquez-Blazquez [d], Ernesto López-Baeza [b]

[a] *Fundação Cearense de Meteorologia e Recursos Hídricos, Fortaleza, Ceará, Brazil*
[b] *Climatology from Satellites Group, Department of Physics of the Earth and Thermodynamics,University of Valencia, Spain*
[c] *Electronic Engineering Department, E.T.S.E, University of Valencia, C/ Dr Moliner 50, 46100 Burjassot, Valencia, Spain*
[d] *Royal Meteorological Institute of Belgium, Bruxelles, Belgium*

A R T I C L E I N F O

Keywords:
Neural networks
Modelization
Net radiation
Radiometer

A B S T R A C T

The knowledge of net radiation at the surface is of fundamental importance because it defines the total amount of energy available for the physical and biological processes such as evapotranspiration, air and soil warming. It is measured with net radiometers, but, the radiometers are expensive sensors, difficult to handle, that require constant care and also involve periodic calibration. This paper presents a methodology based on neural networks in order to replace the use of net radiometers (expensive tools) by modeling the relationships between the net radiation and meteorological variables measured in meteorological stations. Two different data sets (acquired at different locations) have been used in order to train and validate the developed artificial neural model. The statistical results (low root mean square errors and mean absolute error) show that the proposed methodology is suitable to estimate net radiation at surface from common meteorological variables, therefore, can be used as a substitute for net radiometers.

1. Introduction

Net radiation is a fundamental parameter that governs the climate of the lower layers of the atmosphere and it depends critically on the structure and composition of the atmosphere and the presence of clouds, in addition to surface features such as albedo, emissivity, temperature, humidity and thermal properties of the underlying soil. Thus net radiation is a fundamental quantity for analyzing the evolution of climate, from both local and global perspective. It is the driving force of physical and biological processes such as evapotranspiration, the latter being used to optimize the quality and yield of crops, water resources planning, weather forecasting, etc. (Bennie, Wiltshire, Hill, & Baxter, 2008; Ji, Kang, Zhao, Zhang, & Jin, 2009; Li et al., 2009). Despite its importance, the net radiation is measured only in a very few number of standard weather stations because net radiometers are expensive instruments and require constant care in the field, so that the net radiation measurements can be reliable. Hence, this quantity is difficult to obtain due to the cost of net pyrradiometers. This paper presents a methodology for modeling net radiation using artificial neural networks. After an initial period collecting data in order to train the network with real samples, the neural network model can be used as an estimator of net radiation samples for a given area without using net radiometers at all times. The strategy here is to train the neural network model using the net radiation collected "in situ" over a representative period and then use that model and not the net radiometer. There are a large number of linear and nonlinear models who perform modeling of the net radiation at surface but using as input the incoming solar radiation or the net radiation components separately (downwelling shortwave radiation, reflected shortwave radiation, downwelling and upwelling longwave radiation). But the root of the problem remains; radiometers are needed to obtain these input variables to the model (Alados, Foyo-Moreno, Olmo, & Alados-Arboledas, 2003; Daughtry et al., 1990; Kohsiek et al., 2007). This problem can be avoided by using as input parameters, in the neural networks developed to model the net radiation at surface, the most common meteorolog-

☆ This work was supported by the Programme Alβan, the European Union Programme of High Level Scholarships for Latin America, scholarship no. E05D058998BR-Antonio Geraldo Ferreira and also for Remote Sensing Techniques for the Observation of Environmental Parameters in the Valencia Community Autonomous Region for 2007-2009,contract of the Department for Environment, Water, Planning and Housing, General Directorate for Climate Change, Generalitat Valenciana, also this work was supported by the Spanish Ministerio de Educacion y Ciencia under grant TIN2007-61006:Aprendizaje por Refuerzo Aplicado en Farmacocinetica (RLA2AP).

* Corresponding author.
E-mail address: juan.gomez-sanchis@uv.es (J. Gómez-Sanchis).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.231

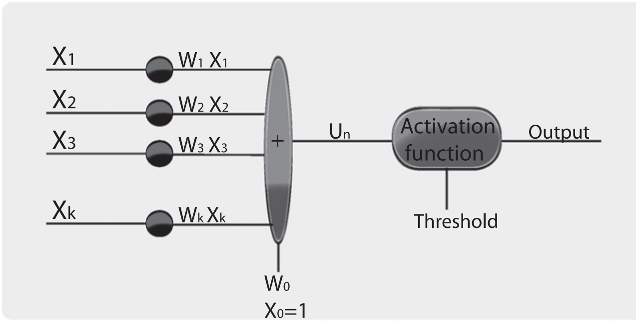

Fig. 1. Scheme of a neuron.

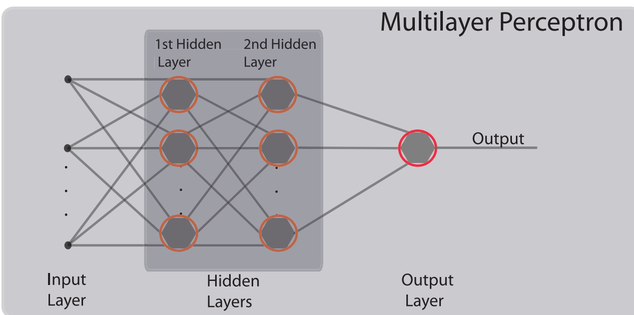

Fig. 2. Multilayer perceptron scheme. Dotted lines show the feedback in the model; in this paper this type of connections was not used.

ical variables collected in the majority of weather stations, around the word, including those that send, daily, information to the *Global Telecommunication System (GTS)*. The GTS is defining as: *The co-ordinated global system of telecommunication facilities and arrangements for the rapid collection, exchange and distribution of observations and processed information within the framework of the World Weather Watch* citegts. These variables are: wind speed, air temperature, atmospheric pressure and humidity.

The following sections describe the neural model used, the multilayer perceptron. After this, the datasets and the variables involved in the problem will be presented. Finally, we will present the results and conclusions obtained in the study.

2. Multilayer perceptron

The multilayer percepron (MLP) has been the neural network model used in this study. It consists of some individual process elements called neurons, which are arranged in a series of layers. Fig. 1 shows the structure of these neurons.

This neuron is constituted, in its first part, by a multiplier, which multiplies the inputs by a series of coefficients called synaptic weights. The objective of learning algorithm is to obtain the optimum values for the synaptic weights (Haykin, 2009). In the next part of the neuron we will find the activation function that gives nonlinear behavior to the neural network model. Fig. 2 shows the scheme of a MLP.

The number of neurons in the input and output layers is defined by the problem addressed. The user is responsible for choosing the number of hidden layers and neurons. There are many demonstrations of the fact that the multilayer perceptron with a hidden layer is an universal modelization tool of continuous functions. In the case of discontinuous functions two hidden layers are required (Reed & Marks, 1999). It is important to highlight that there are rules that guide to designer on the number of hidden neurons in each layer, however there is not work to set this number accurately. In a large number of applications a "trial and error" strategy is used in order to obtain the number of neurons in the hidden layer (Haykin, 2009).

The operation of the neural network is given by the values of synaptic weights. The learning algorithm is the procedure by which the neural model obtain the optimal parameters for solve the problem. There are many learning algorithms but, when choosing a particular one of them is necessary to consider the following features that any learning algorithm should fulfill (Bishop, 1995): Effectiveness, robustness, independence from initial conditions, high generalization ability, low computational cost and simplicity. The aim of any learning algorithm is to obtain as an error (defined

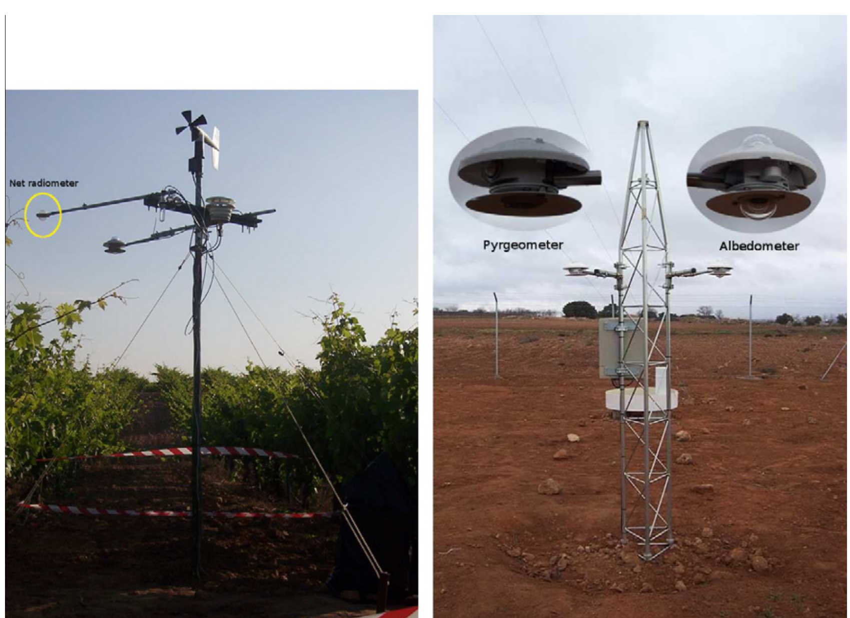

Fig. 3. (left) Net radiometer used to measure net radiation at surface in the FESEBAV experiment and (right) radiometer used at VAS that measure net radiation components separately (downwelling and reflected shortwave radiation, downwelling and upwelling longwave radiation).

Table 1
Basic statistic of FESEBAV and VAS data sets. WS: Wind speed; AT: air temperature; AP: atmospheric pressure; RH: relative humidity; RN: net radiation.

	WS (m/s)	AT (°C)	AP (mb)	RH (%)	RN (W/m^2)
FESEBAV (N = 13,248)					
Maximum	5.05	40.05	937.00	99.30	741.30
Minimum	0.00	8.82	916.00	6.89	−73.30
Mean	1.21	21.99	926.12	54.74	144.94
Standard deviation	0.73	6.57	3.77	25.8	213.24
VAS (N = 23,616)					
Maximum	8.30	36.50	938.00	95.00	1011.15
Minimum	0.00	4.00	914.00	8.00	−114.2
Mean	1.90	19.50	925.44	53.89	136.36
Standard deviation	1.42	6.07	4.06	21.78	244.34

as the difference between actual signal and the neural network output) a value of zero. There are two different kinds of algorithms (Haykin, 2009):

- *On-line*. In this type of learning, error is calculated by the neural network model for each pattern in the data set. The synaptic weights are updated using the error of each pattern.
- *Batch*. In this type of learning, error is calculated by the neural network model for all patterns. After this, the synaptic weights are updated using the average error for all patterns.

The learning algorithm used in this paper has been Levengert–Macquart algorithm which presents a good compromise between speed of convergence, steady-state error and complexity (Bishop, 1995; Haykin, 2009). Random synaptic weights have been used for each architecture in order to avoid the problem of local minima; the authors have initialized 100 times each neural architecture.

3. Data sets

In order to validate our approach were used two data sets obtained from two surface areas with the same land use (vineyard crop) but with different land cover (vineyard and bare soil). Methodology and sensors employed to data collection are described below.

- **Data set 1 (FESEBAV).** The first data set corresponds to data collected during the field campaign called FESEBAV 2007 (*Field Experiment on Surface Energy Balance Aspects over the Valencia Anchor Station area*) conducted from June 19th to September 18th 2007. In this experiment a mobile weather station (EMM) was installed in a field of vines (latitud 39° 31′ 23″N and de longitude 1° 17′ 22″ W, at an altitude of 796 m above sea level), in the study area of the Valencia Anchor Station (VAS), near the town of Caudete de las Fuentes (Utiel-Plana de Requena), Valencia, Spain, with the goal of collecting the data necessary for the surface energy balance studies of the crops. The net radiation was measured with a CN1 Net Pyrradiometer (*Middleton & Co. Pty. Ltd.*), air temperature with a probe PT 100/3 (*Campbell Scientific Ltd.*), the relative humidity with a probe HMP45C (*Campbell Scientific Ltd.*) and wind speed with an anemometer *RM Young 05103* (*R.M. Young Company*). All sensors were installed at 2 m over the surface, except for the wind that was set to 2.10 m. The sensors were integrated into *Campbell CR1000* datalogger and were scheduled to collect data every second. The data recorded every second was stored as 10 min averages.
- **Data set 2 (VAS).** The second data set was obtained at the meteorological station known as VAS (latitude 39° 34′ 15″ N and longitude 1° 17′ 18″ W, lying at an altitude of 813 m above sea level), a reference meteorological station used to calibrate and

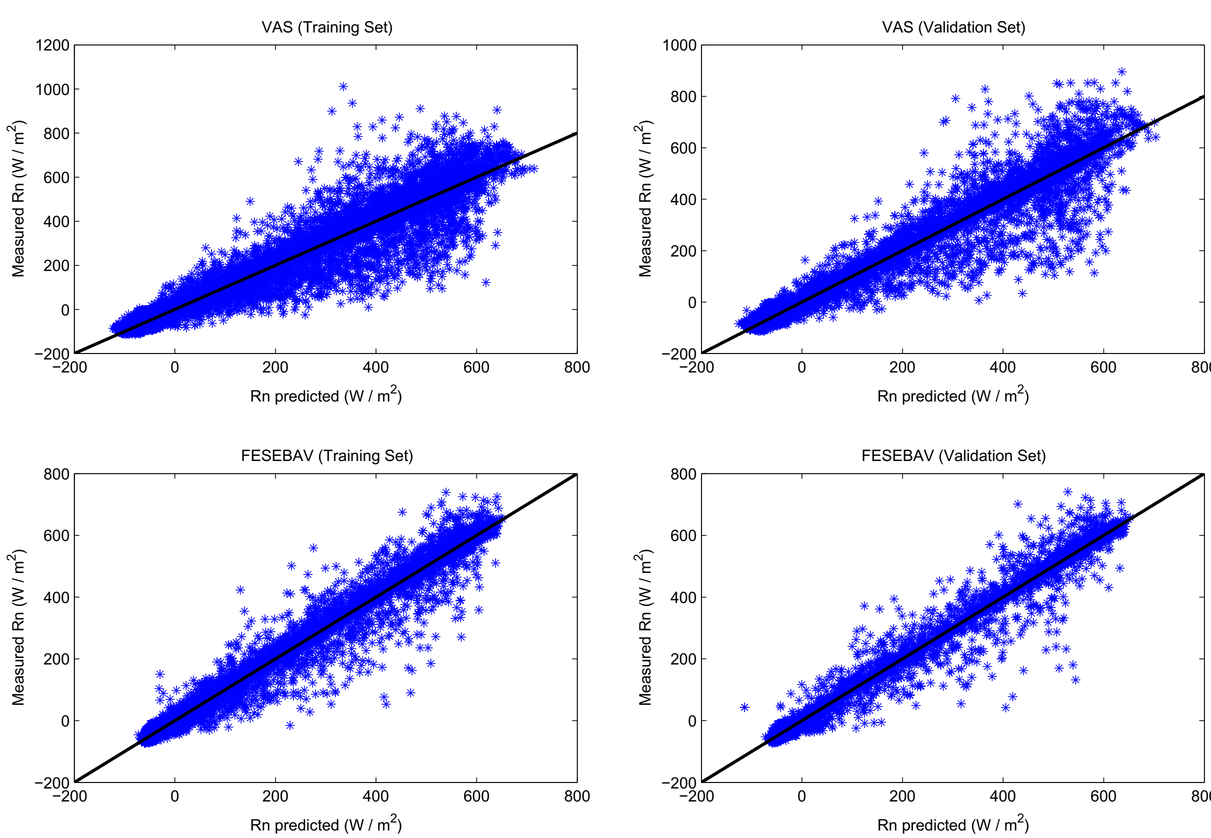

Fig. 4. Representation of net radiation predicted by the neural network model and the actual values of net radiation. In the same graphs we can see the linear regression in both data sets.

Table 2
Performance indices for FESEBAV data set.

FESEBAV data set	MAE (W/m^2)	RMSE (W/m^2)	ME (W/m^2)	*a*	*b*
Training set $N = 8832$	19.46	35.56	−0.38	0.97	3.73
Validation set $N = 4416$	21.65	39.88	0.027	0.97	4.46

Table 3
Performance indices for VAS data set.

VAS data set	MAE (W/m^2)	RMSE (W/m^2)	ME (W/m^2)	*a*	*b*
Training set $N = 15{,}744$	34.55	61.36	0.65	1.00	0.30
Validation set $N = 7872$	36.47	65.07	−0.26	0.99	0.46

validate remote sensing missions low spatial resolution. The station is installed in an isolate characterized by bare soil and surrounded by vineyards. Details on the VAS and meteorological parameters listed can be obtained at http://www.uv.es/anchors/Estacion.html. VAS data used in the study are those relating to the months of May, June and July 2007 and 2008. Wind speed is measured with a sensor model 03102 (*Campbell Scientific Ltd.*), temperature and humidity with a probe model *50U-44212 (Vaisala*, *USA*), atmospheric pressure is measured with a sensor*SPA-900 (Druck Limited*, *USA*), incoming shortwave radiation and reflected solar radiation was measured with an Albedometer *CM14* (*Kipp* & *Zonen*, *Netherland*), upwelling and downwelling of longwave radiation was measured with a pyrgeometer (CG2) (*Kipp* & *Zonen*, *Netherland*). These sensors are installed at 2 m over the surface. In this case the dataloggers employed were *UA Geonica*, and the measurements were acquired in the same way that in data set 1. Fig. 3 show the radiometers used to measure the net radiation directly (left) or its components separately (right).

4. Results

In order to obtain the best neural network model, the models were trained using two hidden layers (by the Cybenko theorem it is known that two layers are necessary to establish the relationships between two data sets (Haykin, 2009)). The number of hidden neurons in each layer was varied from 2 to 20, these limits were imposed because, in all the tests, never reached the upper limits for the number of neurons using cross-validation. Moreover, since the learning algorithm is a local search algorithm on each architecture, there were a total of 100 different initializations of synaptic weights in order to avoid the problem of local minima (Bishop, 1995). The Levenger–Macquart algorithm was chosen using online learning because it is most appropriate in time series modeling problems. Furthermore, cross-validation was used in order to avoid the overfitting problem (Haykin, 2009). For this purpose the data set 1 and data set 2 were divided into two subsets, one for train the models (training set) and other for validate the models (validation set). The proportion selected for this division

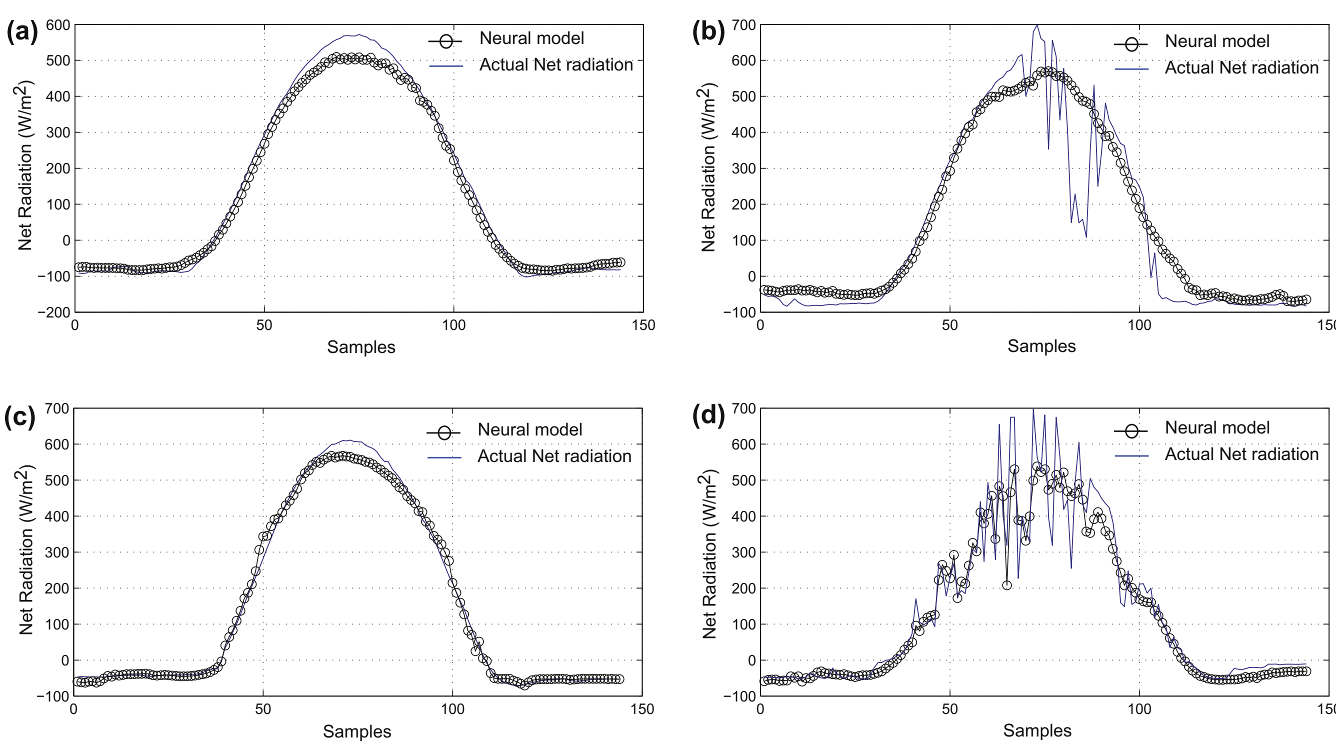

Fig. 5. Measured Rn values (···) and neural model prediction (-o-) for: (a) cloudy free-day, 2-7-2007, VAS data set; (b) cloudy day, 21-5-2007, VAS data set; (c) cloudy free-day, 15-8-2007, FESEBAV data set; (d) cloudy day, 26-6-2007, FESEBAV data set.

Table 4
Performance indices in sunny/cloudy days.

	MAE (W/m^2)	RMSE (W/m^2)	ME (W/m^2)
FESEBAV data set			
Cloudy days $N = 8784$	24.74	43.85	0.44
Sunny days $N = 4464$	11.41	17.21	−1.17
VAS data set			
Cloudy days $N = 17,712$	41.64	71.46	−0.34
Sunny days $N = 5904$	15.84	22.38	2.41

were 2/3 and 1/3 of each set. The input variables used to train the model were: month, day and hour of the measures (temporal information) and, wind speed, air temperature, air pressure and humidity of the air (meteorological variables). The output or target variable was net radiation. Table 1 shows the statistics of the input and output variables statistics in each analyzed data set.

The criterion used for choose the best neural model was that based on the lowest value of the mean absolute error, MAE, in the validation set, this index has been used because it is more robust to outliers than the RMSE (Root Mean Square Error). Fig. 4 shows the representation of net radiation predicted by the neural network model and the actual values of net radiation for the two stations (FESEBAV and VAS) and for training and validation sets.

Tables 2 and 3 show the statistical results obtained for the two sets using the best neural model. In these tables the values of MAE, RMSE and mean error (ME, which give an indication of possible bias on the model) can be seen. The indexes (slope and intercept) of linear regression between the measured net radiation values and the network output (the parameter a define the slope of the adjustment, and must be close to 1 and the parameter b also gives an idea of possible bias on model).

One possible source for the difference between MAE and RMSE obtained for the two sets, shown in Tables 2 and 3, is due to the quickly and high variation, in a short period of time, in net radiation that occur on specific days due to presence of clouds, mainly in the second set (VAS). In VAS dataset, the months of May and June 2008 were those in which net radiation values registered were higher than 800 W/m^2, reaching up to 1011.15 W/m^2, while in FESEBAV dataset, the maximum net radiation value was 741.30 W/m^2 (see 4). Despite this difference, the errors obtained show that neural networks can be used as substitute for radiometers, mainly in clear days, because the prediction errors committed by neural networks can be assumed in the net radiation estimation (Alados et al., 2003; Carrasco & Ortega-Farías, 2008).

Fig. 5 shows the values of net radiation for sunny days (in these cases are regular profiles of net radiation) and others with cloudy sky (irregular profiles) for the two data sets. The same figure shows the prediction of neural models. It is found that the diurnal cycle of net radiation, in cloudy days, have highly irregular shape, so the prediction is quite difficult in these days; in fact most of the errors committed by the model came from these days.

Table 4 shows the difference between the predictions in cloudy days and cloudy free-days.

After evaluate the performance of the neural models using different error rates, the next step was assess the model running through a sensitivity analysis (Gomez et al., 2006) to obtain the most important variables for the model. This kind of analysis has two objectives: (a) check the validity of the conclusions derived from the model; these conclusions should be consistent with the physical theory and (b) obtain new qualitative understanding of the problem. In order to do it the variation in the model output is determined when considering or not determined input variable. If the variation is small it means that the input variable is not very important to obtain the net radiation (the output variable). In another case, if there is much difference in the model output, take into account a specific input variable, or removing it, it means that this variables is very important to the problem. The sensitivity analysis of the model was determined in a experimental procedure. In the first step was choosing the better 25 neural models. After that the input variables, for both, FESEBAV and VAS data set, were arranged in each model by its importance. Finally a boxplot of the input variables position in each of the 25 models is made, Fig. 6.

Fig. 6 shows that the most important variables for the neuronal model are: atmospheric pressure, month, hour and air temperature. However the relevance of a variable sometimes can be masked by other correlated variable, and the atmospheric pressure

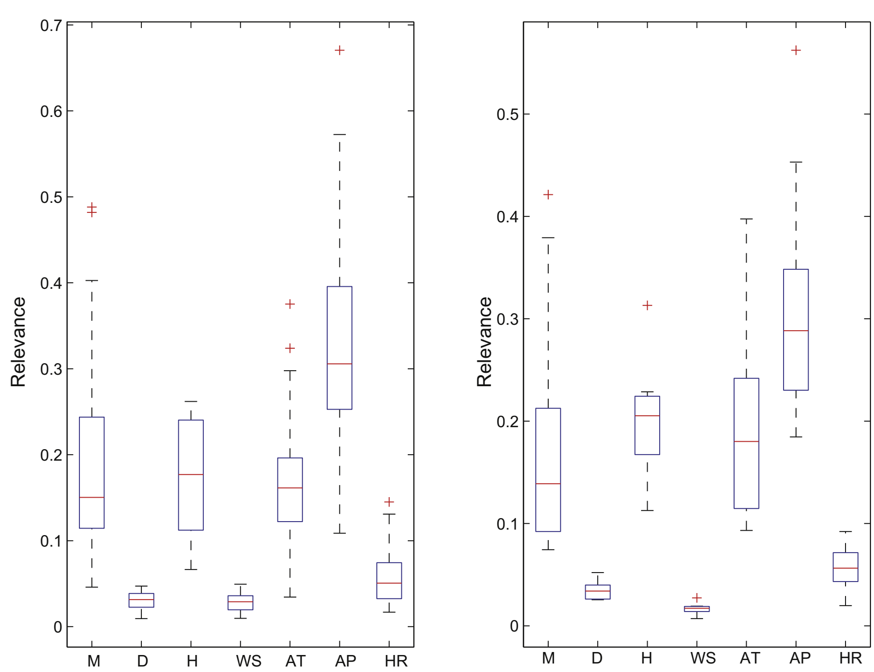

Fig. 6. Relevance of input variables. The inputs are: Month (M), Day (D), Hour (H), wind speed, air temperature, atmospheric pressure and relative humidity.

and air temperature are correlated variables. But it is important highlight that the net radiation at surface depends on surface atmospheric pressure rather than air temperature. Atmospheric pressure is a parameter that is strongly correlated with cloud cover (low pressure indicates more clouds and high pressure indicates less cloud). Thus, atmospheric pressure is an important variable in the radiation balance at the surface because the shortwave and longwave radiation at surface are strongly dependent on cloud cover fraction (Meetschen, van den Hurk, & Drusch, 2004).

5. Conclusions

Net radiation measure is important for the analysis and study of climate, but the devices used to do this are very expensive and difficult to manage requiring further constant care in the field. This paper demonstrates the ability of neural models to replace the use of radiometers for the measurement of surface net radiation. Using neural models and conventional weather variables can be estimated net radiation with an acceptable error without using expensive and costly radiometers. A sensitivity analysis has been carried out in order to obtain the importance of the variables and has been demonstrated that the neural model are valid from a qualitative point of view (its quantitative performance has been demonstrated using error measures). The conclusions drawn about the importance of the variables have physical meaning and agrees with the theory about that.

References

Alados, I., Foyo-Moreno, I., Olmo, F. J., & Alados-Arboledas, L. (2003). Relationship between net radiation and solar radiation for semi-arid shrub-land. *Agricultural and Forest Meteorology, 116*(3–4), 221–227.

Bennie, B. H. J., Wiltshire, A., Hill, M. O., & Baxter, R. (2008). Slope, aspect and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland. *Ecological Modelling, 216*(1), 47–59.

Bishop, C. (1995). *Neural networks for pattern recognition*. Clarendon Press.

Carrasco, M., & Ortega-Farías, S. (2008). Evaluation of a model to simulate net radiation over a vineyard cv. *Cabernet sauvugnon*. *Chilean Journal of Agricultural Research, 68*, 156–165.

Daughtry, C., Kustas, W., Moran, M., Pinter, P., Jackson, R., Brown, P., et al. (1990). Spectral estimates of net radiation and soil heat flux. *Remote Sensing of Environment, 32*, 111–124.

Gomez, J., Martin, J. D., Soria, E., Vila, J., Carrasco, J. L., & del Valle, S. (2006). Neural networks for analysing the relevance of input variables in the prediction of tropospheric ozone concentration. *Atmospheric Environment, 40*, 6173–6180.

Haykin, S. (2009). *Neural networks and learning machines*. Prentice-Hall.

Ji, X., Kang, E., Zhao, W., Zhang, Z., & Jin, B. (2009). Simulation of heat and water transfer in a surface irrigated, cropped sandy soil. *Agricultural Water Management, 96*(6), 1010–1020.

Kohsiek, W., Liebethal, C., Foken, T., Vogt, R., Oncley, S., Bernhofer, C., et al. (2007). The energy balance experiment ebex-2000. Part iii: Behaviour and quality of the radiation measurements. *Boundary-Layer Meteorology, 123*(1), 55–75.

Li, S., Tong, L., Li, F., Zhang, L., Zhang, B., & Kang, S. (2009). Variability in energy partitioning and resistance parameters for a vineyard in northwest china. *Agricultural Water Management, 96*(6), 955–962.

Meetschen, F. A. D., van den Hurk, B. J. J. M., & Drusch, M. (2004). Optimized surface radiation fields derived from meteosat imagery and a regional atmospheric model. *Journal of Hydrometeorology, 5*(6), 1091–1101.

Reed, R. D., & Marks, R. J. (1999). *Neural smithing: Supervised learning in feedforward artificial neural networks*. MIT Press.