Global NEST Journal, Vol 12, No 4, pp 384-389, 2010

ESTIMATION OF MICROCLIMATIC DATA IN REMOTE MOUNTAINOUS AREAS USING AN ARTIFICIAL NEURAL NETWORK MODEL-BASED APPROACH

K.I. CHRONOPOULOS[1,*]
I.X. TSIROS[2]
N. ALVERTOS[1]
I.F. DIMOPOULOS[3]

[1]Department of Chemical and Physical Sciences
Faculty of Sciences, Agricultural University of Athens
Iera Odos 75, 11855, Athens, Greece
[2]Department of Geologic Sciences and Atmospheric Environment, Faculty of Sciences, Agricultural University of Athens,, Iera Odos 75, 11855, Athens, Greece
[3]Technological Educational Institute
Antikalamos, 24100, Kalamata, Greece

Received: 24/11/09
Accepted: 05/02/10

*to whom all correspondence should be addressed:
e-mail: kchrono@aua.gr

ABSTRACT
An artificial neural network (ANN) model-based approach was developed and applied to estimate values of air temperature and relative humidity in remote mountainous areas. The application site was the mountainous area of the Samaria National Forest canyon (Greece). Seven meteorological stations were established in the area and ANNs were developed to predict air temperature and relative humidity for the five most remote stations of the area using data only from two stations located in the two more easily accessed sites. Measured and model-estimated data were compared in terms of the determination coefficient (R^2), the mean absolute error (MAE) and residuals normality. Results showed that R^2 values range from 0.7 to 0.9 for air temperature and from 0.7 to 0.8 for relative humidity whereas MAE values range from 0.9 to 1.8 ^oC and 5 to 9%, for air temperature and relative humidity, respectively. In conclusion, the study demonstrated that ANNs, when adequately trained, could have a high applicability in estimating meteorological data values in remote mountainous areas with sparse network of meteorological stations, based on a series of relatively limited number of data values from nearby and easily accessed meteorological stations.

KEYWORDS: Microclimate, Artificial neural networks, Estimation, Prediction, Mountain canyon, Environmental Management.

1. INTRODUCTION

In mountain regions, meteorological factors such as the solar radiation, the air temperature and the humidity in combination with the intense relief, different slopes, orientations and other topographic irregularities may result in a variety of local microclimates (Barry, 2001). On the other hand, in such areas the meteorological stations network is sparse due mainly to the difficulty of installing and maintaining the instrumentation. In other cases, meteorological data in the desired or required spatial resolution for climatic and bioclimatic assessments are not readily available. In such cases, there is thus a need to estimate data for meteorological parameters not recorded at several locations using observations of the same variable recorded at other sites. Therefore, researchers are often forced to evaluate these conditions using several methods: from data collected in nearby sites (Tang and Fang, 2006); using spatial interpolation techniques such as kriging, thin plate splines, inverse distance weighting (Tveito and Schöner, 2002); and using process-based techniques (e.g. Bolstad *et al*., 1998). Despite the fact that such techniques are commonly used, they suffer from limitations in areas with complex terrain. Recently ANN models have started to be applied in various aspects of the atmospheric sciences, (Benvenuto *et al*., 2000; Melas *et al*., 2000; Dimopoulos *et al*., 2004; Mihalakakou *et al*., 2004; Perez and Reyes, 2006; Tsiros *et al.*, 2009). ANN model applications to meteorological data values estimations are, in general, very few (Cheng *et al.*, 2002; Dimopoulos *et al*., 2005; Chronopoulos *et al*., 2008).

The purpose of the present work is to apply ANN models to estimate values for selected meteorological parameters in a number of sites as a function of the corresponding values of one or more reference stations located far away from the sites. The study area was the National Forest canyon of Samaria located in southern Greece. A previous study of the authors focused on estimating data for two stations located inside the canyon using data from two different stations also located inside the canyon (Dimopoulos *et al*., 2005). The examined period was the warm period of the year 2003. The present study, however, examines the application of ANN models to estimate air temperature and relative humidity for the five most remote stations of the area using data only from two stations located in the two more easily accessed sites. In addition, the present modelling effort uses data for 3 different years and from totally 7 stations located in the study area.

2. STUDY AREA AND DATA

The application site was the canyon of Samaria, a mountainous National Forest, located on the southwest Crete Island, Greece. The canyon extends from 35º18΄27΄΄N and 23º55΄06΄΄E to 35º14΄40΄΄N and 23º58΄01΄΄E , covering a total distance of about 18 km. The dataset used in the present work consists of measured mean hourly temperature and humidity data for 7 meteorological stations established in the canyon for the purposes of the present study and for the following time periods: 12/6/2003 – 4/8/2003 (total 1264 measurements), 6/8/2004-15/9/2004 (total 962 measurements) and 20/6/2005-27/10/2005 (total 3120 measurements). The meteorological stations were HOBO type of Onset Computer Corporation. The sensors were protected with radiation shields and were placed on trees about 1.5m above ground. All measurements were taken every 10 minutes and then were averaged to hourly values. Some statistics of the measured air temperature and relative humidity data are shown in Tables 1 and 2.

3. THE ARTIFICIAL NEURAL NETWORK (ANN) MODEL

The multilayer perceptron (MLP), the most commonly used artificial neural network model, was adopted for the present study. For the multilayer perceptron, the output with one hidden layer is given by (Rumelhart *et al*., 1986a):

$$f(x)=\phi^{s}\left(\sum_{i=1}^{I} w_{is}\phi^{i}\left(\sum_{e=1}^{n} w_{ei}x_{e}+w_{0}\right)+w_{s}\right) \quad (1)$$

where I is the number of hidden nodes, n is the number of input variables, w_{ei} and w_{is} are the weights of the input-to-hidden and hidden-to-output layer, w_0 and w_s are the corresponding thresholds (bias), ϕ^{i} and ϕ^{s} are the units' activation functions.

For model training, the back propagation algorithm was used, which is the most frequently used algorithm for training (Rumelhart *et al*., 1986b). The activation function for the hidden units as well as the output unit is the logistic sigmoid function $\phi(x)=\left(1+e^{-x}\right)^{-1}$. A major consideration in the use of MLP for model building is the determination of the optimal architecture of the network (number of inputs, number of layers and number of nodes per layer). Usually, a trial-and-error method is applied to test various alternative models with possibly different architectures, different values for the training parameters and initial training conditions and choose the one with the best performance. A network can fit the training data arbitrarily closely, but will not necessarily lead to its optimal generalisation ability, i.e. its ability to predict data other than those on which it has been trained. One commonly used method for model architecture selection and model testing is to use only part of the data for training the network. The remaining data is used to estimate the generalisation ability of the network (hold-out method). In the present work, the training set consisted of ½ of the data, the set for network architecture selection of ¼ of the data and the test set of the remaining ¼ of the data for the estimation of the generalisation ability, randomly assigned. The model network developed uses one hidden layer with 5 nodes since it was found that this is the number of layers that gives the best results on the selection set.

4. RESULTS

The ANN models were used to predict air temperature and humidity for the most remote stations of the area, S_2 – S_6, using data only from stations S_1 (entrance of the canyon) and S_7 (end of the canyon), located in the more easily accessed areas. Measured and estimated data of the test set of both air temperature and relative humidity were compared in terms of the determination coefficient (R^2) and the mean absolute error (MAE) (Table 3). R^2 values range from 0.7 to 0.9 for air temperature and from 0.7 to 0.8 for relative humidity whereas MAE values range from 0.9 to 1.8 ^oC and 5 to 9%, for air temperature and relative humidity, respectively.

Table 1. Statistics of the measured air temperature (^oC) data

Station	2003		2004		2005	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
S_1	22.0	3.0	18.8	4.5	18.1	4.8
S_2	28.1	3.6	24.4	5.3	23.2	6.1
S_3	26.8	3.9	24.9	4.4	23.4	5.4
S_4	26.5	3.8	24.9	4.0	23.7	4.3
S_5	26.8	4.1	25.3	4.1	23.8	5.7
S_6	26.3	3.3	25.4	3.2	24.2	4.6
S_7	27.2	2.8	25.9	3.0	25.5	4.5

Table 2. Statistics of the measured relative humidity (%) data

Station	2003		2004		2005	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
S_1	38.3	9.5	51.9	19.8	59.9	16.0
S_2	33.8	12.4	37.6	15.4	44.0	19.6
S_3	34.7	12.3	38.4	13.4	45.8	20.5
S_4	35.2	12.2	38.3	12.2	51.0	23.1
S_5	39.0	14.1	40.0	13.0	50.8	24.5
S_6	46.9	15.5	46.1	13.5	47.6	17.2
S_7	44.1	13.4	45.5	12.9	48.3	15.4

Table 3. Values of the determination coefficient (R^2) and the mean absolute error (MAE) of the ANN models predictions for air temperature and humidity at the stations along the canyon

Station	Air Temperature		Relative Humidity	
	R^2	MAE, (^oC)	R^2	MAE (^oC)
S_2	0.90	1.4	0.83	5.6
S_3	0.89	1.3	0.80	6.3
S_4	0.72	1.8	0.73	8.6
S_5	0.86	1.6	0.73	8.9
S_6	0.92	0.9	0.80	4.6

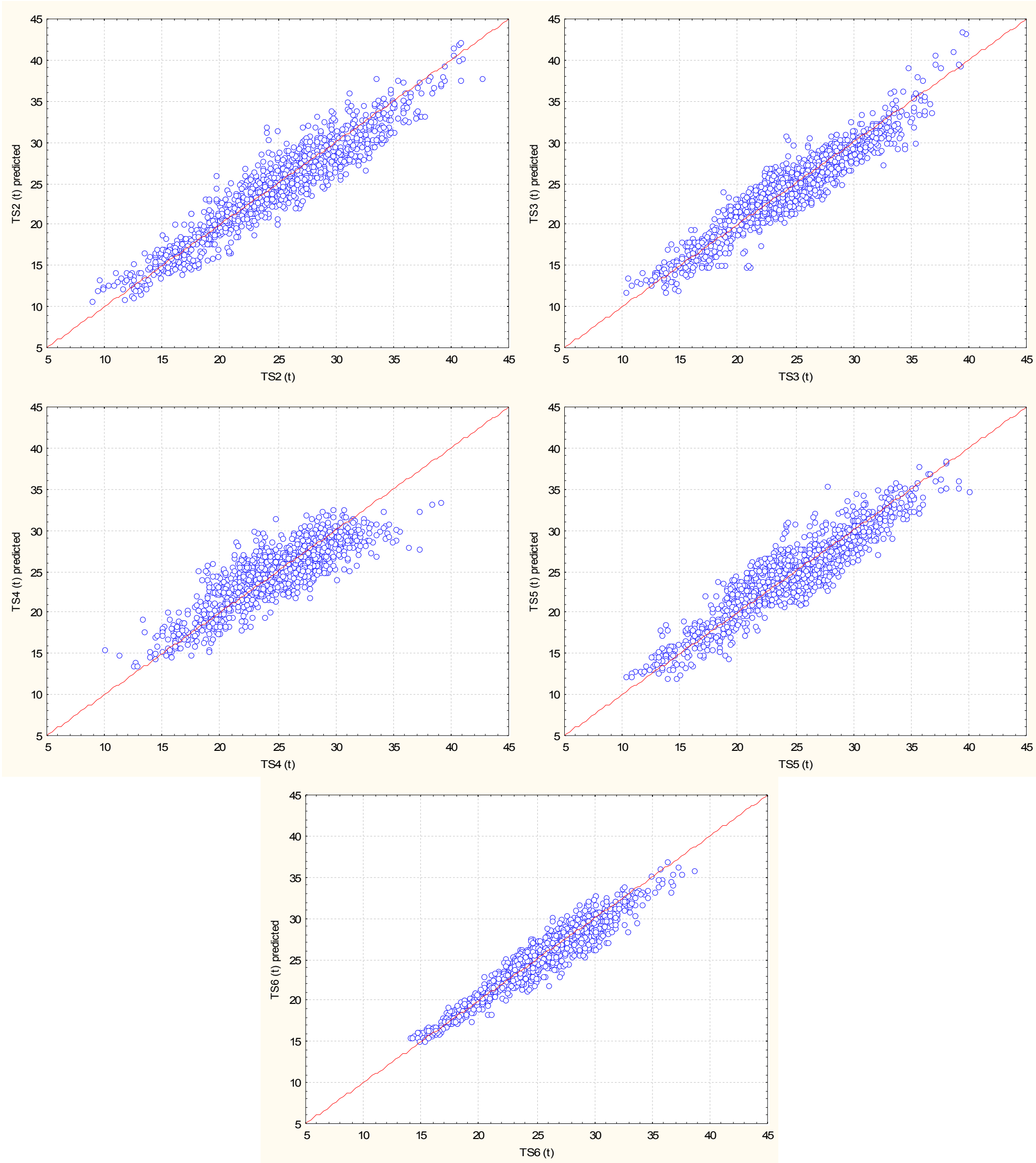

Figure 1. Scatter plots of observed versus predicted values of air temperature for the five stations inside the canyon

For the robustness of the model predictions, the distribution of the residuals was also examined. The normality of the residuals was also examined using the Shapiro-Wilk normality tests and it was found that residuals had a normal distribution. In addition, the relationship between residuals and the model-estimated values showed independence. Figures 1 and 2 show the scatter plots of the observed values versus the predicted ones, for air temperature and relative humidity estimations, respectively.

5. CONCLUSION

The results of the present study can be considered satisfactory since the artificial neural network model developed from a set of ‘training data’ was found able to predict microclimatic parameters for the remote stations of the study area using data from more easily accessed areas. The approach proposed in the present study may be used to any site, assuming that ANNs are adequately trained. ANN training, however, requires a relatively large number of microclimatic data times series and this may be an important limitation in some cases.

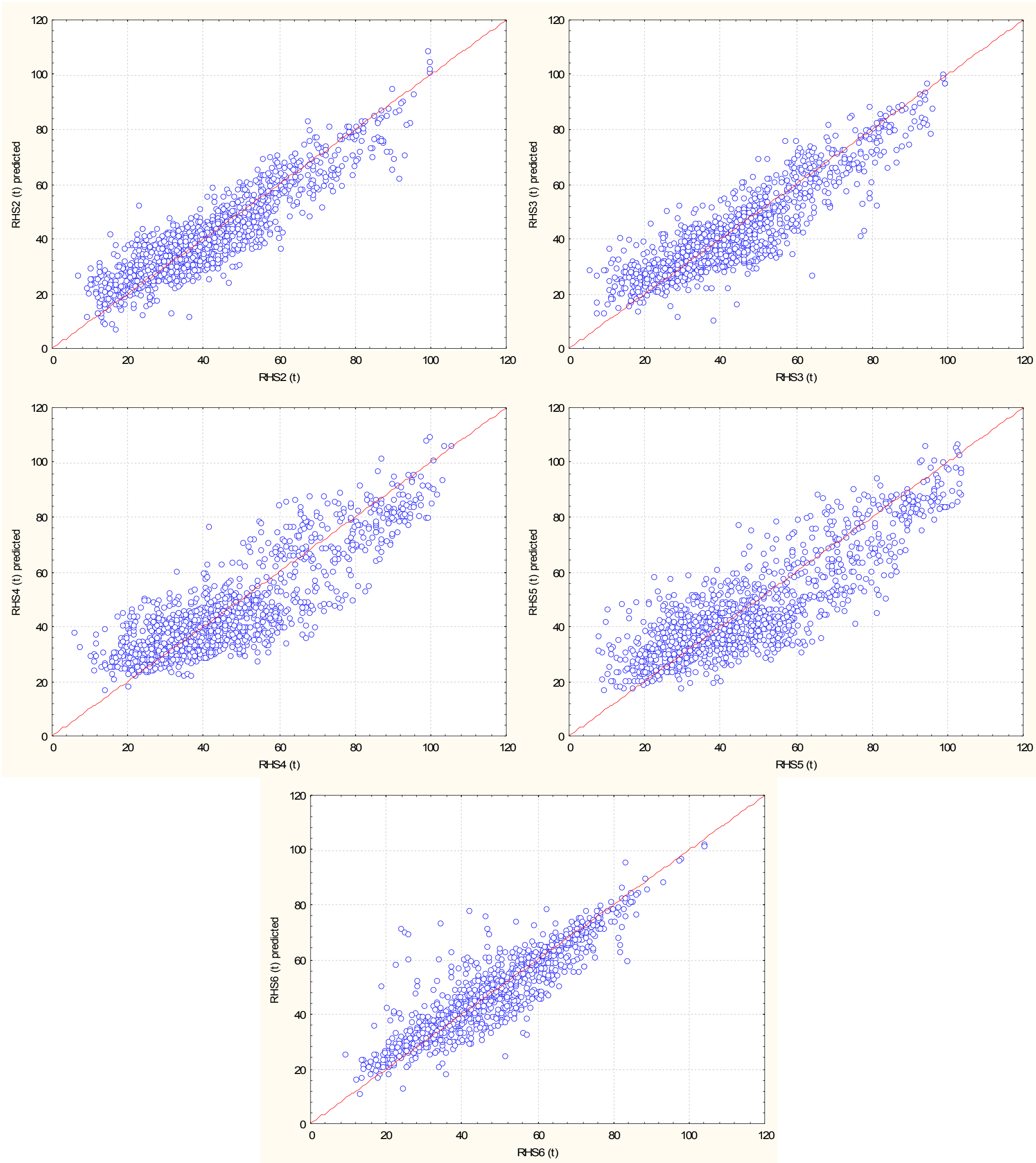

Figure 2. Scatter plots of observed versus predicted values of relative humidity for the five stations inside the canyon

REFERENCES

Barry L.G. (2001), Mountain weather and climate, Second Edition, Routledge, Taylor and Francis Group, New York, USA.

Benvenuto, F. and Marani, A. (2000), Neural networks for environmental problems: Data quality control and air pollution nowcasting, *Global Nest: the International Journal*, **2**, 281-292.

Bolstad P.V., Swift L., Collins F. and Régnière J. (1998), Measured and predicted air temperatures at basin to regional scales in the southern Appalachian mountains, *Agricultural and Forest Meteorology*, **91**, 161–176.

Cheng S.Y., Jin Y.Q., Liu L., Huang G.H., Hao R.X. and Jansson C.R.E. (2002), Estimation of atmospheric mixing heights over large areas using data from airport meteorological stations, *J. Environ. Sci. Health A*, **37**, 991 – 1007.

Chronopoulos K., Tsiros I., Dimopoulos I. and Alvertos N. (2008), An application of artificial neural network models to estimate air temperature data in areas with sparse network of meteorological stations, *J. Environ. Sci. Health A*, **43**, 1752-1757.

Dimopoulos I.F., Tsiros I.X., Serelis K. and Chronopoulou A. (2004), Combining neural network models to predict spatial patterns of airborne pollutant accumulation in soils around an industrial point emission source, *J. Air & Waste Manage. Assoc.*, **54**, 1506-1515.

Dimopoulos I., Chronopoulos K., Alvertos N. and Tsiros I. (2005). A Preliminary Study of Predicting Meteorological Parameters in a Canyon Using Statistical and Artificial Neural Network Methods. In: Proceedings of the 9th Conference on Environmental Science and Technology, September 1-3, Rhodos, Greece, Lekkas, Th. (Edit.), Vol. A, pp. 305-310. University of Aegean, Lesvos.

Melas D., Kioutsoukis I. and Ziomas I.C. (2004), Neural network model for predicting peak photochemical pollutant levels, *J. Air & Waste Manage. Assoc.*, **50**, 495-501.

Mihalakakou G., Santamouris M., Papanikolaou N., Cartalis C. and Tsangrassoulis A. (2004), Simulation of the Urban Heat Island Phenomenon in Mediterranean Climates, *Pure Appl. Geophys.*,**161,** 429-451.

Perez P. and Reyes J. (2006), An integrated neural network model for PM_{10} forecasting, *Atmos. Environ.*, **40**, 2845-2851.

Rumelhart D.E., Hinton G.E. and Williams R.J. (1986a), Parallel Distributed Processing Explorations in the Microstructure of Cognition, Rumelhart, D.E., McClelland, J.L., (Eds.), Vol. 1, MIT Press, Bradfords Books: Cambridge, MA, 318-362.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986b), Learning representations by back-propagating error, *Nature,* **323**, 533-536.

Tang Z. and Fang J. (2006), Temperature variation along the northern and southern slopes of Mt. Taibai, China, *Agricultural and Forest Meteorology*, **139**, 200-207.

Tsiros I., Dimopoulos I., Chronopoulos K. and Chronopoulos G. (2009), Estimating airborne pollutant concentrations in vegetated urban sites using statistical models with microclimate and urban geometry parameters as predictor variables: a case study in the city of Athens, *J. Environ. Sci. Health A*, **44**, 1496-1502

Tveito O.E. and Schöner W. (eds.) (2002) ‘Applications of spatial interpolation of climatological and meteorological elements by the use of geographical information systems (GIS)’, report no 28/02 KLIMA, NORWEGIAN METEOROLOGICAL INSTITUTE, ISSN 0805-9918.