7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.
Edited by M. Caetano and M. Painho.

Artificial Neural Networks applied in the determination of Soil Surface Temperature – SST

Maurício Roberto Veronez[1], Adriane Brill Thum[1], Anderson Silva Luz[2] and Deivis R. da Silva[2]

1 Graduate Program in Geology - Vale do Rio dos Sinos University
Avenida Unisinos, 950, São Leopoldo - RS, Brazil – CEP: 93022-000
Tel.: + 55 (51) 3591-1100 – Extension 1769 Fax: + 55 (51) 3590-8177
veronez@unisinos.br, adrianebt@unisinos.br

2 Students of Civil Engineering. - Vale do Rio dos Sinos University
Avenida Unisinos, 950, São Leopoldo - RS, Brazil – CEP: 93022-000
Tel.: + 55 (51) 3591-1100 – Ramal 1769; Fax: + 55 (51) 3590-8177
andersonluz@terra.com.br

Abstract

Artificial intelligence techniques are being used to facilitate modeling in many different research areas. One example of these techniques is the use of Artificial Neural Networks – ANNs. ANNs are not a new technique and have been studied since the 1940s. The technique was almost forgotten during the 1970s, but reappeared in the 1980s as a possible alternative to traditional computing. Today, ANNs are used in many projects, especially to forecast data, learn algorithms, optimize systems, recognize standards, and others. The surface temperature (ST) is a parameter influenced by changes in weather (temperature and air relative humidity, wind speed, precipitation, etc.) and indicates the hydric state of a plant. Thus, estimating the ST is very useful in monitoring projects that tend to hydric demands of cultures, which will contribute to irrigation programs. Another important application is in the use of the determination of evapotranspiration, where, together with other components of the hydrological cycle, it is important to evaluate the replenishing of underground water-bearing aquifers. Recently, a method used to estimate the ST uses the analysis of NOAA-AVHRR thermal images adapted to the split windows equation. This modeling relates emission amount of variables (generated by the images) and atmospheric data. It is a complex methodology, because besides difficult statistical modeling, it is necessary to digitally process the images to determine the emission amount. ANNs are indicated in this study due to their excellent capacity for generalization, classification, interpolation, extrapolation, tolerance to errors and noise, and because of the fact that they do not require the specification of explicit parameters to complete the modeling process. The purpose of this project is to verify the possibility of the use of ANNs to estimate soil temperatures. The test area selected is a part of an urban area located in Ivoti - RS. In different points of the area, values of ST were collected using a portable temperature sensor model. For each collecting point, the position was known by the UTM coordinates (E, N) and the elevation (H). Different network topologies were tested in a supervised way by the backpropagation algorithm, using the E, N and H coordinates, as data entry, and giving ST as results. The best topology found was that possible in the limitations of available time for the execution of the successive necessary refinements. Despite the fact that several tests have been done with two and three intermediate layers, a simple topology was adopted, with only one intermediate layer of 3 and 5 neutrons. Four learning algorithms were tested. The algorithms, Scaled Conjugate Gradient, Levenberg-Marquerdt, and Resilient, provided values for ST with mean errors below 0.9 ºC, during the simulations, while the Gradient Descendent showed a mean error below 1.9 ºC.

Keywords: Artificial Neural Networks, modeling, soil surface temperature

1 Introduction

The radioactive processes on surface are of paramount importance to the redistribution of humidity and heat along the soil and in the atmosphere. Heat and humidity exchanges affect the behavior of the Earth's biosphere, weather and climate (Bastiaanssen *et al*., 1998a; Roerink *et al*., 2000).

Changes in the energy and humidity balance in the Earth-ocean-atmosphere system, in the tropics, influence the global climate (Aguttes *et al.,* 2000). Energy exchanges in the vegetation-atmosphere interface, by means of the components of the radiation balance, sensible heat flux and latent heat flux are essential to climate modeling, as the magnitude of these flows and of their variants, in periods shorter than one day, are important in the parametrization and calibration of global circulation models. In longer intervals, these measures are also used in global climate impact models resulting from physiographic alterations on the surface (Sellers *et al*., 1995). The estimate of spatial variation in evaporation processes is fundamental in several applications related to water resources and climate modeling (Mohamed *et al*., 2004).

Owing to the shortage of meteorological data, the atmospheric and hydrologic models receive input from regional data whose resolution is inadequate to represent the atmospheric situations that are to be modeled (Paiva *et al*., 2005).

The continuous spatial monitoring of flows on the surface, which are very relevant to weather and climate forecasts in different space-time scales, is not yet available. The main reason for this gap is the complexity of the physical system involved; moreover, conventional or traditional evaluation methods over large areas require an extensive system of meteorological measures.

The energy balance method – Bowen Ratio, for example, uses measures of temperature and humidity gradients in the atmospheric layer that is next to the surface, coupled with measures of radiation and heat flux on the soil (Diak *et al*., 2004).

On the other hand, the aerodynamic method for turbulent transport takes the following into account: temperature and air humidity measures at two levels above the surface, wind speed, atmospheric stability conditions, and also the surface's aerodynamic properties (Rosenberg *et al*., 1983).

Because of data shortage, there is the crucial need for the use of alternative techniques so as to complement meteorological information on a given locality. Remote sensing techniques have been often used to obtain information on surface and atmosphere parameters, which are important for the monitoring of flows or of associated parameters in regional and global scales, whose level of detail is determined by the spatial resolution of the multispectral sensors used.

According to (Paiva, *et al*., 2005), several current algorithms used to obtain energy flows via remote sensing are still unsatisfactory due to the following problems:

- algorithms used to obtain flows via remote sensing require information on the surface that is available only during specific field experiments (Diak *et al*., 2004).
- use of empirical relations, which hinder the use of such algorithms for different categories of soil use, which are feasible only if supported by local calibrations (Blyth and Dolman, 1995).
- the hypothesis of equality of the heat source temperature and surface radioactive temperature, involved in turbulent processes of heat transport between the surface

and the atmosphere, which brings forth significant errors in the estimation of flows on the surface (Bastiaanssen, 1995; Diak *et al*., 2004).

Within multispectral methods, Split-Window (SW) is the most widespread.

The panorama depicted above discloses the great complexity in the estimation process of Soil Surface Temperature. Through the use of Artificial Neural Networks (ANNs), there appears a new method of estimating ST in a simpler manner, as fewer variables are involved.

2 Objectives

This study aimed at estimating Soil Surface Temperature concerning the urban area, based on ANNs. Through the use of the backpropagation algorithm or one of its variants, different topologies as well as several sets and training parameters were tested in an attempt to define the most appropriate neural network for the generalization of the issue addressed. In order to spatialize the area under study, training sessions and tests were carried out with the use of supervised learning algorithms, receiving the following input data: surface temperature, average air temperature and position (UTM coordinates and altitude).

3 Soil Surface Temperature estimate by means of the Remote Sensing technique

For an estimate of the ST, there is the need for determining the bright temperatures of channels 4 and 5 of *AVHRR* sensor of the *NOAA-14* satellite (*National Oceanic & Atmospheric Administration*), which correspond to the following wavelength ranges: 10.5-11.5μm and 11.5-12.5μm, respectively. In its turn, the bright temperature is a function of the radiance registered by the satellite. The radiometric correction, or radiometric calibration, is the procedure through which digital information registered by the satellite sensor is turned into radiance. *AVHRR*'s channels 4 and 5 are calibrated by orbital parameters which are determined for each scene, based on two standard radiation sources: space, whose radiation is nearly zero, and a blackbody on board of the satellite, whose temperature is kept at nearly 288K.

The equation which turns the digital number registered by the sensor into radiance, for a given j channel centered in a given ν wave number, is given by equation 1:

$$B_{j(\nu)} = S_{j(\nu)} \cdot DN + I_{j(\nu)} \tag{1}$$

where:

Bj(ν) represents the radiance (mW/sr m^2 cm^{-1}),

Sj(ν) represents the angular coefficient of the j channel calibration equation (mW/m^2 sr cm^{-1} count),

DN represents the image's digital number,

Ij(ν) represents the linear coefficient of the j channel calibration equation (mW/m^2 sr cm^{-1}).

The calibration equation coefficients hold the information concerning the sensor's response function in a given channel. Further details on these coefficients can be found in the *NOAA* polar orbiter data user's guide (Kidwell, 1991).

Due to the lack of linearity in the response from the *AVHRR* sensor, the radiances obtained through equation 2 have to be corrected as follows:

$$B_{j(\nu)corr} = A_j \,.\, B_{j(\nu)} + B_j \,.\, B_{j(\nu)}{}^2 + D_j \tag{2}$$

where:

$B_{j(\nu)corr}$ represents the corrected radiance (mW/sr m^2 cm^{-1}),

A_j, B_j and D_j represent the correction coefficients, for a given j channel, because the AVHRR sensor lacks linearity.

The A_j, B_j and D_j coefficients, as far as satellite *NOAA-14* is concerned, assume the following values: 0.92378; 0.0003822 and 3.72, respectively, for *AVHRR*'s channel 4.

For channel 5, these values are 0.96194; 0.0001742 and 2.00, respectively. The conversion from radiance into bright temperature for a given temperature range (265 to 320K) is expressed by:

$$T_{bj} = \frac{1.438833 \cdot v_j}{\ln\left(\dfrac{1 + 1.1910659 \text{ x } 10^{-5} \cdot v_j{}^3}{B_{j(\nu)corr}}\right)} \tag{3}$$

where:

Tbj represents channel j's bright temperature,

νj represents channel j's wave number,

Bj(ν)corr represents the radiance that was corrected as per equation (2).

Further details on radiometric calibration and bright temperature setting, and on the tables that have the coefficients used in these expressions, can be found in section 1.4.10 of NOAA Polar Orbiter Data User's Guide (http://www2.ncdc.noaa.gov/docs/podug/html/c1/sec-1.410.html) and in (Kidwell, 1991).

4 Artificial neural networks applied in the determination of geoidal undulations

Basically, an Artificial Neural Network is a non-algorithmic technique based on systems of equations that are usually nonlinear and linked, in which the output value (result) of an equation is the input for other several equations of the network. The Neural Network is formed by artificial neurons. These ones have been projected to have an analog behavior in relation to the Biological Neurons.

Formally, according to Haykin (1999), an artificial neural network is a distributed parallel processor, consisting of simple units of processing with which knowledge can be stored and used for consecutive assessments. Its behavior reminds the human brain by two aspects: knowledge is acquired through a process of learning and the connections among the neurons, known as synaptic weights, are used to store the acquired knowledge.

4.1 Artificial Neuron

The Neural Network is formed by processing units called "Artificial Neurons". Each neuron has the following behavior: input data is multiplied by the synaptic weights (w_{ki}), added and subjected to an activation function that provides the output as Figure 1 illustrates.

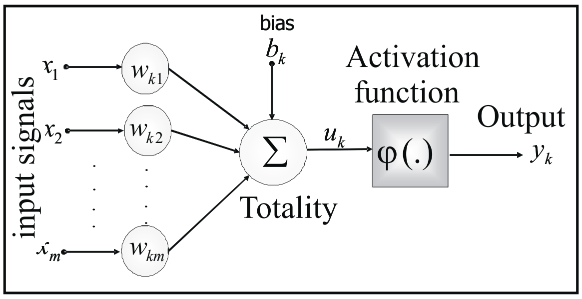

Figure 1 Artificial Neuron. Adaptation from Haykin (1999).

Among the several kinds of existing neural networks, the Multilayer Perceptron Model (MLP) was chosen because its implementation is easy and it is relatively simple.

The MLP networks show a great computing power due to the insertion of intermediate layers that differ from the Perceptron model originated by Rosemblatt (1958) which only had one level of neurons connected directly to the output layer.

The solution for separable nonlinear problems can be worked out by the use of networks with one or more intermediate layers. The network is then formed by at least three layers: the input one, the intermediate or hidden one, and the output one.

According to Cybenko (1989), a network with an intermediate layer can implement any continuous function and with two intermediate layers, it is possible to approximate to any mathematical function.

4.2 Generalized Delta Rule and Backpropagation Algorithm

The generalized delta rule is the most used learning rule for training a MLP network. It consists on the application of the gradient descent by the use of the Backpropagation algorithm.

When a determined pattern is supplied to the network for the first time (first learning cycle), it produces a random output. The difference between this output and the desired one is the error. The Backpropagation algorithm is responsible for the calculation of the error functions. The aim of the training phase is to constantly reduce its value. For this, the weights must be updated every new iteration. Equation 4 shows the error function MSE – Minimum Squared Error.

It is important to mention that the ANN's input and output data go through a process of normalization where they are usually grouped with intervals between [0 and 1} or [-1 and 1].

$$MSE_p = \frac{1}{2}\sum_j e_{pj}^2 \text{ , with } e_{pj} = d_{pj} - y_{pj} \tag{4}$$

where:

d_{pj} = desired output value,

y_{pj} = obtained output value,

j = a neuron of the output layer,
p = pattern of the neural network.

5 Materials and method

The locality of the study corresponds to a section of the urban area of the Ivoti municipality (29º30'21", south latitude and 51º15'43" longitude, west of Greenwich) located in Rio Grande do Sul State – Brazil, as Figure 2 shows. The figure also shows the spatialization of the points sampled for the neural network training (points identified as triangles) as well as those established in the process of validating the model (points identified as circles).

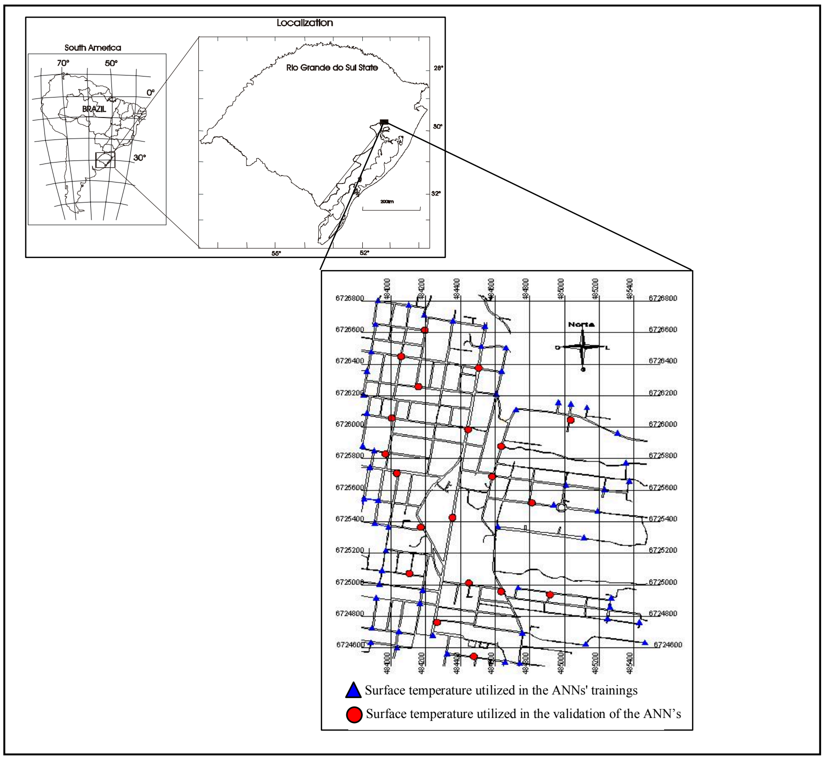

Figure 2 Illustration of the study area.

The data used in all the investigation procedures refer to:

- UTM Coordinates (East, North) and Orthometric Height obtained through the aerophotogrammetric map of the Ivoti municipality, and which corresponded to the neural network inputs.
- Soil Surface Temperatures colleted with *a laser* thermometer in the intersections of the transport system of the municipality and which corresponded to the output data of the neural network.

For the development of the method proposed, one has attempted to use only easily accessible data and information, so that it can be referred to later.

As the nomenclature itself suggests, the training and validation sets are used during the training process and the test set is used for obtaining the estimates.

The definition of *Neural Network Topology* is one of the greatest challenges in neural modeling, but it is also an important step towards conceiving the latter, as it narrows down the type of problems that can be addressed by the network. Thus, in order for a good generalization to exist, the network must receive as much information as possible. However, that also implies building a topology with a great amount of neurons, which brings forth another issue: the computational complexity which restricts the use in relation to processing time.

In this study, the following algorithms were used: Scaled Conjugate Gradient Backpropagation, Levenberg-Marquardt, Resilient Backpropagation and Gradient Descendent Backpropagation. For these algorithms, the following were used: activation function of the intermediate layers equals to the hyperbolic tangent, 0.3 learning rate, and 0.9 momentum term.

For each algorithm aforementioned, the results refer to 4 different network topologies:

- Topology 3-3-1 with 900 training cycles.
- Topology 3-3-1 with 1800 training cycles.
- Topology 3-5-1 with 900 training cycles.
- Topology 3-5-1 with 900 training cycles.

The efficiency of each algorithm tested was based on 20 points (Figure 2) through the obtention of the Mean Error as per equation 5:

$$\overline{E} = \frac{\left(ST_{(Known)} - ST_{(ANN)}\right)}{20} \tag{5}$$

Where:

$\overline{E}$ = mean error of soil surface temperature in °C,

$ST_{(Known)}$ = known soil surface temperatures in °C,

$ST_{(ANN)}$ = soil surface temperature obtained through artificial neural networks, respectively in °C.

6 Results and Discussion

During a few months, dozens, or rather, hundreds of different topologies were certainly tested. This way of defining the topology takes a considerable amount of time, and it is nevertheless quite likely that an untested combination might have a better response to the expected generalization and convergence time than the one selected.

Although several tests were carried out using two and three intermediate layers, a simple topology was selected – having only one intermediate layer with 3 and 5 neurons. This was a convenient decision so as to speed up the method, since time spent on network processing grows exponentially as the number of layers and the number of neurons in these layers also grows.

The simulations showed that an increase in the number of neurons in the intermediate layer has brought nearly no significant improvement to the surface temperature forecast. For the 3-3-1 and 3-5-1 configurations, the Scaled Conjugate Gradient, Levenberg-Marquardt, and Resilient algorithms have shown mean errors below 0.9 ºC. The worst results were from the Gradient Descendent Backpropagation algorithm whose mean error was below 1.9ºC.

Different numbers of cycles were also tested; one could verify that an increase in this value (from 900 to 1800), in addition to bringing forth a longer processing time, has also made obtaining ST slightly worse. This was verified in the 4 models studied. Nevertheless, the error rates were again always below 1.85ºC. These tests were conducted in order to verify whether an increase in this value could lower down the error rate when ST is obtained. However, no significant gain was verified for this issue.

Table 1 shows all the results obtained in the tests conducted:

Table 1 Final results of the mean errors when obtaining ST with the tests conducted.

Simulation	Topology	Algorithm	Training Cycles	Mean Error (°C)
01	3-3-1	Scaled Conjugate Gradient	900	0.41
02	3-3-1	Scaled Conjugate Gradient	1800	0.43
03	3-5-1	Scaled Conjugate Gradient	900	0.48
04	3-5-1	Scaled Conjugate Gradient	1800	0.42
05	3-3-1	Levenberg-Marquardt	900	0.44
06	3-3-1	Levenberg-Marquardt	1800	0.37
07	3-5-1	Levenberg-Marquardt	900	0.34
08	3-5-1	Levenberg-Marquardt	1800	0.63
09	3-3-1	Resilient	900	0.49
10	3-3-1	Resilient	1800	0.40
11	3-5-1	Resilient	900	0.50
12	3-5-1	Resilient	1800	0.88
13	3-3-1	Gradient Descendent	900	1.01
14	3-3-1	Gradient Descendent	1800	1.17
15	3-5-1	Gradient Descendent	900	1.84
16	3-5-1	Gradient Descendent	1800	1.59

It is then clear that the Gradient Descendent model had the worst results. Increasing the number of cycles did not yield a significant difference in the rate of mean errors, but the neural network training time became longer.

It was found that the method proposed may be an efficient alternative when determining Soil Surface Temperature, since there are few, easily-accessible variables involved in the modeling.

7 Conclusions and Recommendations

The main conclusions below are based on the findings of this study:

- As regards the four algorithms tested, the Scaled Conjugate Gradient, the Levenberg-Marquardt and the Resilient yielded, during the simulations, a mean error lower than 0.7ºC when obtaining ST, whereas for the Gradient Descendent the mean error was lower than 1.9ºC. Even though a greater number of tests still have to be run, the findings seem to signal that neural networks can be a very efficient way to obtain ST from the soil. It should also be noted that, according to (Schirmbeck, 2004), the model for obtaining the ST based on Remote Sensing offers a rate of uncertainty of ± 1.6ºC.
- Using a large number of neural network training cycles may increase the mean error. The tests conducted showed that the best results were the ones obtained with cycles ranging from 900 to 1800.
- An increase in the number of intermediate layers of the neural network may also reduce the obtention of ST. Results have shown that, for the tests carried out, the 3-3-1 and the 3-5-1 topologies were the ones with the best results. Certainly, several different untested configurations are likely to be analyzed and may provide compatible results with the ones from this study.

Other suggestions for further research are the following:

- Testing topologies other than the ones used in this study.
- Conducting experiments in non-urban areas with the same configurations used in this study.
- Adopting the coupled use of the Remote Sensing technique and neural networks in order to obtain ST in large areas.
- Conducting experiments which will involve not only the surface temperature but also evapotranspiration in the neural network output.

References

Aguttes, J. P. *et al*., 2000, MEGHA – TROPIQUES: A satellite for studying the water cycle and energy exchanges in the tropics. In *Proceedings of the* 51 st International Astronautical Congress, Rio de Janeiro, Brazil.

Bastiaanssen, W. G. M., 1995. *Regionalization of surface flux densites and moisture indicators in composite terrain*. Ph.D. Thesis. Wageningen Agricultural University, Wageningen, The Netherlands, pp. 273.

Bastiaanssen, W. G. M., 1998. *Remote Sensing in Water Resources Management: The State of the Art*. International Water Management Institute, Colombo, Sri Lanka, pp. 118.

Blyth, E. M., Dolman, A. J., 1995, The roughness length for heat of sparse vegetation. *Journal of Applied Meteorology*, 34 (1), pp. 583 – 585.

Cybenko, G., 1989, Approximation by superposition of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2 (1), pp. 303-314.

Diak, G. R., *et al*., 2004, Estimating land surface energy budgets from space: Review and current efforts at the University of Wisconsin – Madison and USDA – ARS. *Bulletin of American Meteorological Society*, January, pp. 65 – 78.

Haykin, S., 1999, *Neural networks: a compreensive foundation*. 2ª Ed. Prentice-Hall, pp. 842.

Kidwell, K. B., 1991, NOAA polar orbiter data user guide (TIROS-N, NOAA-6, NOAA-, NOAA-8, NOAA-9, NOAA-10, NOAA-11, and NOAA-12). National Oceanic and Atmospheric Administration, Washington, D.C.

Mohamed, Y. A., *et al*., 2004, Spatial variability of evaporation and moisture storage in the swamps of the upper Nile studied by remote sensing techniques. *Journal of Hydrology*, 289 (1), pp. 145-164.

Paiva, C. M., *et al*., 2004, Estimativa das componentes do balanço de energia via satélite através do modelo SEBAL. In *Proceedings of the* XIII Congresso Brasileiro de Meteorologia, Fortaleza, Sociedade Brasileira de Meteorologia.

Rosemblatt, F., 1958, The Perceptron: A Probabilistic Model Of Information Storage And Organization In The Brain. *Cornell Aeronautical Laboratory Phychological Review*, 65 (6), pp. 386-408.

Rosenberg, N. J., *et al*., (1983). *Microclimate: The Biological Environment*. 2nd Edition, John Wiley & Sons, Lincon-Nebrasca, pp. 495.

Schirmbeck, J., 2005, *Estimativa da evapotranspiração regional em área de ocorrência do Sistema Aqüífero Guarani – Vale do Rio dos Sinos – RS*. São Leopoldo, Tese de Mestrado. Universidade do Vale do Rio dos Sinos UNISINOS, Programa de Pós-Graduação em Geologia, pp. 127.

Sellers, P. J. *et al*., 1995, Remote sensing of the land surface for studies of global change: model-algorithm-experiments. *Remote Sensing of Environment*, 51 (1), pp. 3-26.