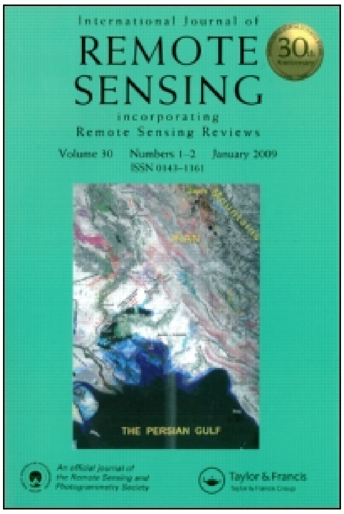

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

An algorithm to estimate soil moisture over vegetated areas based on *in situ* and remote sensing information

N. D. Ramírez-Beltran[a]; C. Calderón-Arteaga[b]; E. Harmsen[c]; R. Vasquez[d]; J. Gonzalez[e]
[a] Department of Industrial Engeering, University of Puerto Rico, Mayagüez, PR [b] Department of Mechanical Engineering, University of Puerto Rico, Mayagüez, PR [c] Department of Agricultural Engineering, University of Puerto Rico, Mayagüez, PR [d] Department of Electrical and Computer Engineering, University of Puerto Rico, Mayagüez, PR [e] Department of Mechanical Engineering, Santa Clara University, Santa Clara, CA, USA

Online publication date: 21 May 2010

To cite this Article Ramírez-Beltran, N. D. , Calderón-Arteaga, C. , Harmsen, E. , Vasquez, R. and Gonzalez, J.(2010) 'An algorithm to estimate soil moisture over vegetated areas based on *in situ* and remote sensing information', International Journal of Remote Sensing, 31: 10, 2655 — 2679

To link to this Article: DOI: 10.1080/01431160903085628
URL: http://dx.doi.org/10.1080/01431160903085628

International Journal of Remote Sensing
Vol. 31, No. 10, 20 May 2010, 2655–2679

Taylor & Francis
Taylor & Francis Group

An algorithm to estimate soil moisture over vegetated areas based on *in situ* and remote sensing information

N. D. RAMÍREZ-BELTRAN*†, C. CALDERÓN-ARTEAGA‡, E. HARMSEN§, R. VASQUEZ¶ and J. GONZALEZ|

†Department of Industrial Engeering, University of Puerto Rico, Mayagüez, PR 00680
‡Department of Mechanical Engineering, University of Puerto Rico, Mayagüez, PR 00680
§Department of Agricultural Engineering, University of Puerto Rico, Mayagüez, PR 00680
¶Department of Electrical and Computer Engineering, University of Puerto Rico, Mayagüez, PR 00680
|Department of Mechanical Engineering, Santa Clara University, Santa Clara, CA, USA

(*Received 4 September 2007; in final form 5 June 2008*)

An algorithm is proposed for estimating soil moisture over vegetated areas. The algorithm uses *in situ* and remote sensing information and statistical tools to estimate soil moisture at 1 km spatial resolution and at 20 cm depth over Puerto Rico. Soil moisture within the study region is characterized by spatial and temporal variability. The temporal variability for a given area exhibits long- and short-term variations that can be expressed by two empirical models. The average monthly soil moisture exhibits the long-term variability and is modelled by an artificial neural network (ANN), whereas the short-term variability is determined by hourly variation and is represented by a nonlinear stochastic transfer function model. Monthly vegetation index, land surface temperature, accumulated rainfall and soil texture are the major drivers of the ANN to estimate the monthly soil moisture. Radar, satellite and *in situ* observations are the major sources of information of the soil moisture empirical models. A self-organized ANN was also used to identify spatial variability to be able to determine a similar transfer function that best resembles the properties of a particular grid point and estimate the hourly soil moisture across the island. Validation techniques reveal an average absolute error of 3.34% of volumetric water content and this result shows that the proposed algorithm is a potential tool for estimating soil moisture over vegetated areas.

1. Introduction

Accurate initial soil moisture content distribution is an important requirement for mesoscale atmospheric simulation models (Fast and McCorkle 1991, Clark and Arritt 1995, Gallus and Segal 2000, Golaz *et al.* 2001). Soil moisture affects land surface–atmosphere interactions by influencing the partition of incoming radiation into sensible and latent heat fluxes, and the separation of precipitation into infiltration and surface runoff (Pan *et al.* 2003). The regional atmospheric modelling system (RAMS) has being used to simulate the climate dynamics in the Caribbean basin and

*Corresponding author. Email: nazario@ece.uprm.edu

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2010 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903085628

it has been found that RAMS is highly sensitive to soil moisture initial conditions (Comarazamy 2001). Therefore, incorrect soil moisture initial conditions can generate misleading modelling results. For instance, Balsamo *et al.* (2004) reported that erroneous estimation of total soil moisture affects the quality of the forecast for several days when using a numerical weather prediction tool.

Due to the importance of estimating the soil moisture, intensive research efforts have been devoted to the subject during the last three decades. Jackson (1982) applied passive microwave remote sensing to estimate soil moisture for fields covered with moderate vegetation. Jackson derived a theoretical and empirical model to estimate soil moisture. Parameters of the model were estimated by using a regression approach based on observations collected over experimental plots. Njoku and Entekhabi (1996) used a spaceborne microwave remote sensing system to estimate soil moisture at a spatial resolution of 10 to 20 km. Their preliminary results show that applications under different soil moisture conditions need to be investigated. Wetzel and Woodward (1987) studied the statistical relationships between soil moisture and infrared surface temperature observations taken from the visible Infrared Spin Scan Radiometer (VISSR) at a Geostationary Operational Environmental Satellite (GOES). They used regression techniques to relate soil moisture to surface temperature, wind speed, vegetation coverage, and low-level temperature advection. Pan *et al.* (2003) proposed a diagnostic soil moisture equation derived from the linear stochastic partial differential equation and rainfall observations. They used data from NEXRAD and assumed that the soil moisture is a scalar field; this assumption was introduced by means of a water balance equation in which the rainfall is the fundamental variable to represent the soil moisture. Huang *et al.* (1996) used the water balance principle to model the soil moisture. They used the surface air temperature and total precipitation to derive a soil moisture model. In May 2002 NASA's Aqua satellite was launched with a passive microwave radiometer known as the Advanced Microwave Scanning Radiometer (AMSR-E). This scanning instrument senses microwave radiation at six frequencies, ranging from 6.9 to 89.0 GHz. It has been shown that the C and X bands at 6.9 and 10.7 GHz are strongly related to the dielectric constant of the surface soil and this relation is used to estimate the soil moisture with global coverage and coarse resolution (Njoku and Li 1999, Njoku *et al.* 2003). Daily products are available at the National Snow and Ice Data Center (NSIDC) dating from February 4004. Wang and Schmugge (1980) proposed an empirical model to measure the dielectric constant as a function of soil moisture content. They observed that the dielectric constant increases slowly with moisture content up to a transition point and beyond the transition region the dielectric constant increases rapidly with moisture content.

It has been shown that precipitation is highly correlated with soil moisture (Jiang and Cotton 2004), and land surface temperature also has a significant correlation with the soil moisture (Sun and Pinker 2004). Jiang and Cotton (2004) implemented an artificial neural network (ANN) algorithm to estimate soil moisture. They used daily precipitation, vegetation index, cloud-mask infrared skin temperature and soil moisture profile. They claim that an ANN algorithm is capable of estimating soil moisture from remotely sensed infrared data with high spatial and temporal resolutions. They reported that the application of ANN exhibits some difficulties during the training process due to the need for high quality training data.

It should be noted that the AMSR-E signal cannot be used to estimate soil moisture over densely vegetated areas, since the satellite signal is contaminated with vegetation effect. Thus, in this work we are proposing a methodology to estimate soil moisture over vegetated areas and an application is developed for the Puerto Rico (PR) environmental conditions.

2. Data

In situ and remote sensing information used to develop the proposed methodology are organized into three different sources of information: field sampling, remote sensing and soil texture, and topographic information.

2.1 *Field sampling data*

Observations of soil moisture were obtained from 15 soil stations located in Puerto Rico and figure 1 shows the location of these stations. Twelve stations are owned and operated by our research project and three are owned and operated by the Natural Resources Conservation Services (NRCS). Our stations include three ECHO-20 soil moisture sensors that collect information at 20 cm depth, an ECHO temperature sensor located at 20 cm height, an ECRN-50 rain gauge, and a data logger that collects the information from the five sensors every hour. Detailed descriptions of the sensors used can be found at: http://www.decagon.com/Ech2o/sensors1. The NRCS stations include soil moisture sensors at 20.32 cm depth (and other depths); rainfall, soil and air temperature, wind and other parameters are also measured. Each station collects soil moisture, rainfall and air temperature data on an hourly basis. Undisturbed soil sample cores were obtained from each station with the purpose of determining soil texture and calibrating the soil moisture sensors for PR climatological conditions.

2.2 *Remote sensing data*

Land surface temperature and vegetation index values were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument. MODIS instruments are installed onboard the Terra and Aqua satellites. MODIS information is provided by the Goddard Distributed Active Archive Center (DAAC). Land

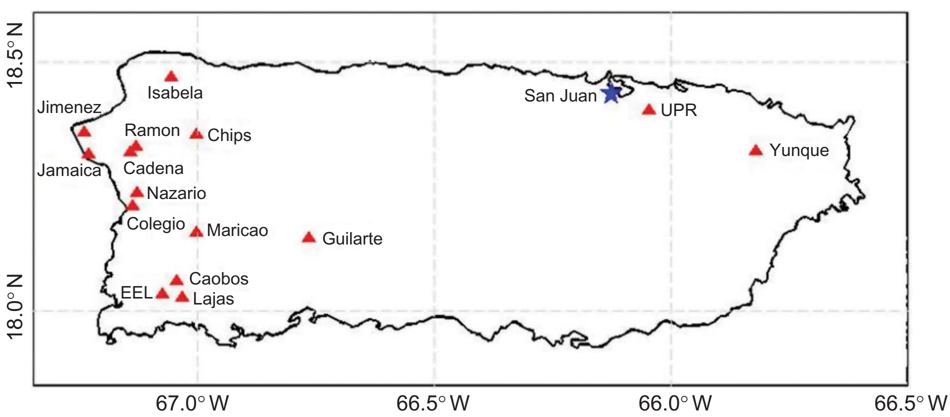

Figure 1. Soil moisture network stations. Red triangles indicate the location of stations. Guilarte, Isabela and Maricao are owned and operated by Natural Resources Conservation Services (NRCS) and the remaining stations are owned and operated by our project.

surface temperature data were provided at 1 km spatial resolution. The amplitude of temperature was obtained by computing the difference between the monthly average diurnal temperatures and the monthly average nocturnal temperatures. Eight-day temperature files were organized to obtain the monthly average surface temperature and the daily surface temperatures were used to estimate the hourly temperature. The monthly normalized difference vegetation index (NDVI) was also extracted from MODIS at 1 km spatial resolution. Rainfall data were obtained from the Next Generation Radar (NEXRAD), which is managed by the National Weather Service (NWS). The NWS recently implemented the multisensor precipitation estimation (MPE) algorithm that integrates gauge- and radar-derived precipitation estimates (Breidenbach and Bradberry 2001). MPE generates an hourly rainfall product in the XRMG format, which is a binary rainfall file. NEXRAD rainfall data is given at a 1 km spatial resolution. Several convective events were extracted in the XRMG format and were compared with a dense rain gauge network (PR has 125 rain gauge networks and collects rainfall observation every 15 min). An arbitrarily selected event is given in figure 2, which shows an example of comparing the observed and estimated rainfall values during 17 April 2003. The top panel of figure 2 presents the individual comparison between observed (rain gauges) and estimated (radar pixel) rainfall values. The bottom panel presents the accumulated rainfall between the observed and estimated rainfall values. This particular event provides a mean absolute error of 0.99 mm and an average bias ratio of 1.01. The NEXRAD validation project for the case of PR is undergone and for the purpose of this research we assumed that the NEXRAD provides reliable estimates of rain rates.

2.3 *Soil texture and topography data*

Soil samples were obtained from each station to determine soil texture. A soil texture map at 20 cm depth and at 1 km spatial resolution was developed based on the irregular distribution of soil texture provided by the United States Department of Agriculture (USDA) of the Natural Resources Conservation Service (NRCS). The soil textures at 20 cm depth from 118 samples were used to develop the PR soil texture maps, assuming that organic matter is negligible.

Digital elevation for PR was provided at 30 m spatial resolution by the United States Geological Survey (USGS). This map was upscaled to 1 km spatial resolution to obtain all the variables in the same spatial distribution. The upscaling was performed by aggregating the smaller pixels into a larger pixel and the average elevation from the smaller pixels was assigned to the elevation of the larger pixel.

3. Soil moisture modelling

One of the purposes of this work is to develop a soil moisture estimation model for tropical areas such as PR, which has complex topography and vegetation. The soil moisture exhibits long- and short-term variability and it will be modelled by long- and short-term memory models, respectively. Figure 3 shows the scheme of the soil moisture model in which the long- and short-term memory models are integrated. This figure also shows the major sources of information to derive the soil moisture estimates. The f function represents an artificial neural network model and the g function is a nonlinear stochastic transfer function model. The artificial neural network model estimates the monthly soil moisture variation and the transfer function model estimates the hourly soil moisture variation.

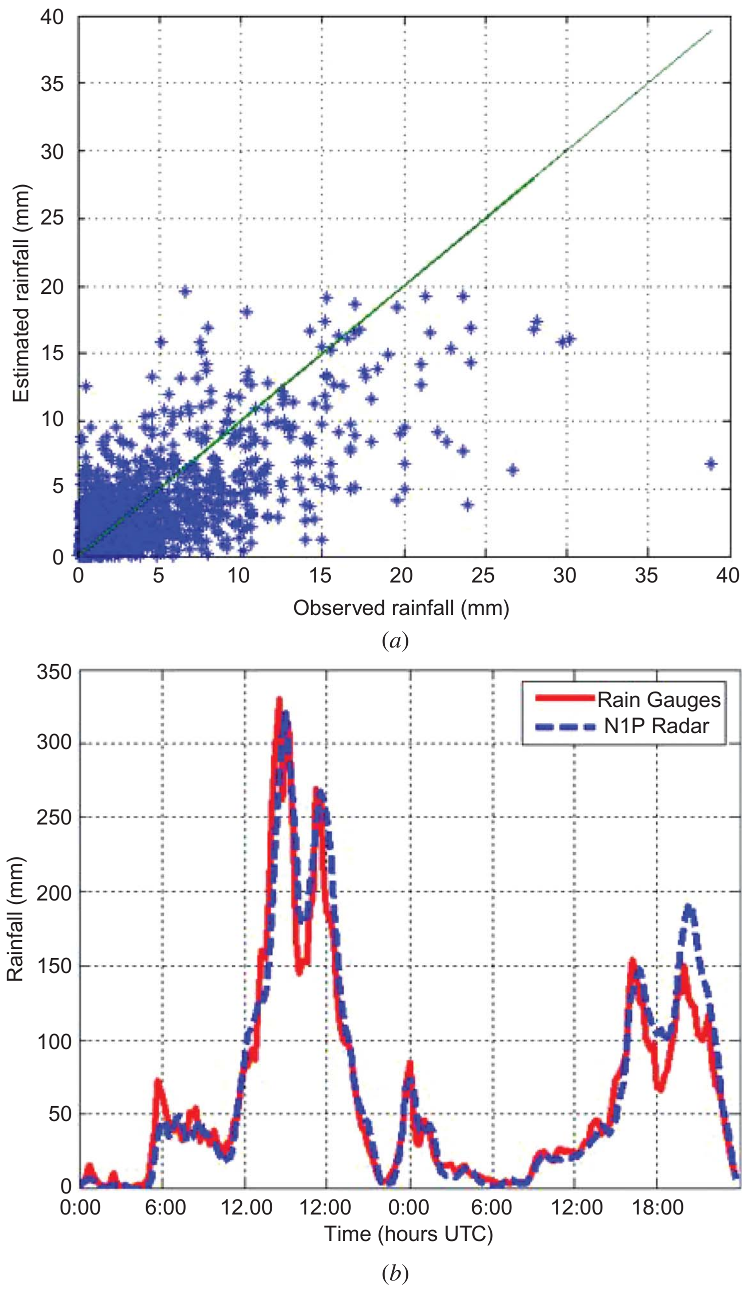

Figure 2. (*a*) Individual comparison between observed and estimated rainfall, (*b*) comparison between rain gauge and radar using accumulated rainfall data.

3.1 *Long term memory model*

The formulation of the model is based on careful analysis of the soil moisture process. Historical records show that there are strong correlations between monthly soil moisture cumulative rainfall, soil texture, vegetation index and surface temperature. Thus, an ANN was used to estimate the long-term soil moisture response. The selected ANN was

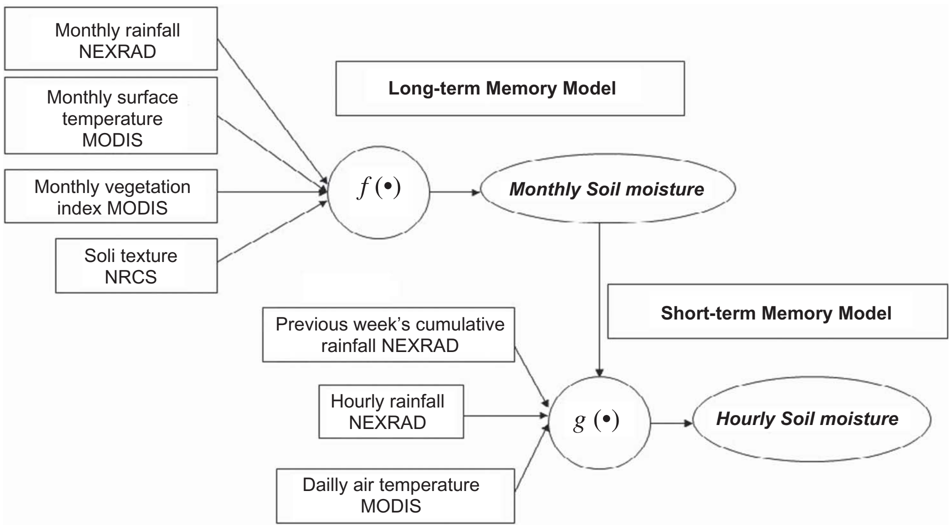

Figure 3. Soil moisture algorithm.

the general multi-layered perceptron with the Levenberg–Marquardt Backpropagation algorithm because the dynamic water content of soil behaves as a nonlinear process and the ANN has been proven to be an efficient approach when the variables of a system are nonlinearly related (Hagan *et al.* 1996). An ANN determines an empirical relationship between the inputs and outputs of a given system. Thus, it is important that the user has a good understanding of the science behind the underlying system to provide the appropriate inputs, and consequently to support the identified relationship. The key to obtaining a successful ANN application is to select the appropriate training variables and to identify a suitable ANN structure. The structure of the ANN consists of determining the number of layers, the number of neurons in the hidden layers, and selecting the best activation function for each layer. In this case we used an input vector with a two-layer ANN because this size happens to be a computational efficient structure (Hagan *et al.* 1996). The selected structure includes five input variables and one output variable. Three activation functions were explored (linear, sigmoidal and hyperbolic tangent functions), and a sigmoidal function for the hidden layer and a hyperbolic tangent function for the output layer performed best. The number of hidden layers ranged from one to five. The Levenberg–Marquardt Backpropagation algorithm was selected because this algorithm automatically computes the optimal learning rate (Hagan and Menhaj 1994). For each run, discrete searching techniques were used to determine the optimal number of hidden layers. We published a detailed procedure to identify the appropriate structure of a neural network (Ramirez-Beltran and Montes 2002) and this procedure was adopted in this research.

The training patterns of the neural networks are formed by the input and output vectors of monthly data, which can be written as follows:

$$\boldsymbol{p}_{m,j} = [C_j, \quad S_j, \quad \alpha_j, \quad \nu_{m,j}, \quad \Delta_{m,j}] \tag{1}$$

$$\boldsymbol{q}_{m,j} = [h_{m,j}] \tag{2}$$

where $\boldsymbol{p}$ is the input vector, and $\boldsymbol{q}$ is the output vector, which is also known as the target vector; the variable h is the monthly average of soil moisture measured at 20 cm depth; S and C are the percentage of sand and clay at a given location, respectively; the variable α is the monthly accumulated rainfall, ν is the monthly vegetation index; Δ is the range of monthly surface temperatures, (maximum – minimum surface temperatures). The subscripts m and j have been omitted from these definitions and represent the m^{th} month and the j^{th} location, respectively.

An ensemble procedure was implemented to develop a stable and robust estimation. Five members of the ensemble were generated with the best initial points that were obtained by a random search that minimizes the difference between the output of the neural network and the observed soil moisture. Thus, the monthly soil moisture was obtained by computing the median of the optimal soil moisture estimates, i.e. the estimated value is the optimal value of the central tendency. The final estimates from the long-term memory model will be soil moisture on a monthly basis, at 1 km spatial resolution, and this information will be used to create the initial conditions at every grid for the hourly soil moisture model.

3.2 *Short-term memory model*

One of the major purposes of this work is to develop a soil moisture hourly estimation model for tropical areas such as PR, which has complex vegetation and topography. The formulation of the model is mainly based on careful study of soil moisture observations and how the soil moisture behaviour is related to other environmental variables. This model considers hourly precipitation and temperature as the fundamental input variables, since it has been shown that these variables exhibit high correlation with soil moisture (Jiang and Cotton 2004). The hydrologic cycle suggests that the soil moisture dynamics can be expressed by a first order partial differential equation with constant coefficients. The variation of soil moisture is intrinsically related to the moisture state as follows (Thornthwaite 1948, Mather 1978, Huang *et al.* 1996):

$$\frac{\partial h(t)}{\partial t} = k_1 R(t) + k_2 T(t) + k_3 h(t) \tag{3}$$

where $\frac{\partial h(t)}{\partial t}$ is the soil moisture variation with respect to time t; $R(t)$, $T(t)$ and $h(t)$ are functions of rainfall, air temperature and soil moisture at time t, respectively. The functions $R(t)$, $T(t)$ and $h(t)$ are continuous in time; however, in practice they are measured at discrete point times and they are expressed by R_t, T_t and h_t, respectively. In addition to k_1, k_2, and k_3 are the constants for model calibration. Pandit and Wu (1983) showed that a first order differential equation can be written as a first order difference equation, and this result can be easily generalized by the first order partial differential equation. Thus, the discrete form of equation (3) can be written as a partial difference equation with constant coefficients as follows (Box and Jenkins 1976, Pandit and Wu 1983, Wei 1990, Hamilton 1994, Brockwell and Davis 2002):

$$h_t = \gamma h_{t-1} + \varpi R_t + \varphi T_t \tag{4}$$

where γ, ϖ and φ are constants that are estimated from data and the remaining variables are defined in equation (3). The performance of the last equation and the analyses of the collected data suggest that the stochastic and discrete transfer function can be used as the best representation of the soil moisture behaviour. The relationship between the continuous and discrete transfer functions was also discussed by Pandit and Wu (1983) and Box and Jenkins (1976). The stochastic and discrete generalized transfer function for ***n*** inputs $x_1, x_2, \ldots, x_n$ and a single output y_t, can be written as follows:

$$y_t = \frac{\omega_1(B)}{\delta_1(B)} B^{d_1}(x_{1,t}) + \frac{\omega_2(B)}{\delta_2(B)} B^{d_2}(x_{2,t}) + \cdots + \frac{\omega_n(B)}{\delta_n(B)} B^{d_n}(x_{n,t}) + \varepsilon_t \qquad (5)$$

where $\omega_j(B) = \omega_{j,0} - \omega_{j,1}B - \omega_{j,2}B^2 - \cdots - \omega_{j,s}B^s$, and $\delta_j(B) = \delta_{j,0} - \delta_{j,1}B - \delta_{j,2}B^2 - \cdots - \delta_{j,r}B^r$. B is the back shift operator, and the power on B means the time delay on variable x_t, for instance $B^{d_j}(x_t) = x_{t-d_j}$, and d_j is the time delay between the jth input and the output variable; the s and r are the order of the series and they are estimated for a specific process based on data and model identification techniques. The ratio of two polynomials $\omega_j(B)$ and $\delta_j(B)$ creates an infinite time series that is known as the impulse response function between the input and the output variables of the system; ε_t is a white noise process at time t with a mean zero and constant variance. It should be noted that the infinite series of the impulse response function converges when the system is asymptotically stable (Pandit and Wu 1983). The rational representation of the transfer function is a parsimonious model and the evaluation of the transfer function is given by long division of two polynomials, in which only the most significant coefficients are considered. The asymptotic stability of the transfer function is measured by computing the roots of the polynomial ($\delta_j(B) = 0$). The system is asymptotically stable if the roots of ($\delta_j(B) = 0$) fall outside of the unitary circle (Pandit and Wu 1983).

From experience gained in the collection of soil moisture data and experimentation, rainfall and temperature processes are the major drivers of a short-term soil moisture response. During wet episodes the soil moisture level is mainly controlled by the current and historical rainfall events (see figure 4 from hour 201 to the end). During dry episodes the soil moisture behaviour is controlled mainly by the temperature; this effect can be noted in figure 4 during the period from hour 50 to hour 200, while instantaneous changes are controlled by the rainfall process. In a sequence of consecutive rain episodes, the soil moisture can arrive to a maximum level, called the saturation point, from which it quickly descends to the field capacity level; this behaviour can be observed in figure 4 from hour 250 to the last hour.

Rainfall and soil moisture patterns exhibited in figure 4 also suggest that rainfall effects on the soil moisture response can be represented by an impulse response function in which the soil moisture response is modulated by an exponential term associated with the accumulated rainfall, r_L. Thus, if no rainfall events occur during the previous days, the r_L will be zero and the instantaneous soil moisture response will be large. On the other hand, if several rainfall events occur in a short time interval then the r_L will be greater than zero and consequently the soil moisture response will be attenuated by the exponential term. Therefore, the proposed nonlinear stochastic transfer function to model the hourly soil moisture is:

$$h_t = \left(\frac{\omega_{0,1} - \omega_{0,2}B}{1 - \delta_{0,1}B}\right) h_{t-1} + \left(\frac{\omega_{1,1} - \omega_{1,2}B}{1 - \delta_{1,1}B}\right) R_t \exp(-\tau r_L) + \left(\frac{\omega_{2,1} - \omega_{2,2}B}{1 - \delta_{2,1}B}\right) T_t + \varepsilon_t \quad (6)$$

where R_t is the instantaneous rainfall at time t, r_L is the accumulated rainfall during the last L number of hours, the integer variable L is obtained during the parameter estimation process, the ωs and δs are the parameters of the impulse response functions. The coefficient τ is the attenuation parameter that modulates the impulse

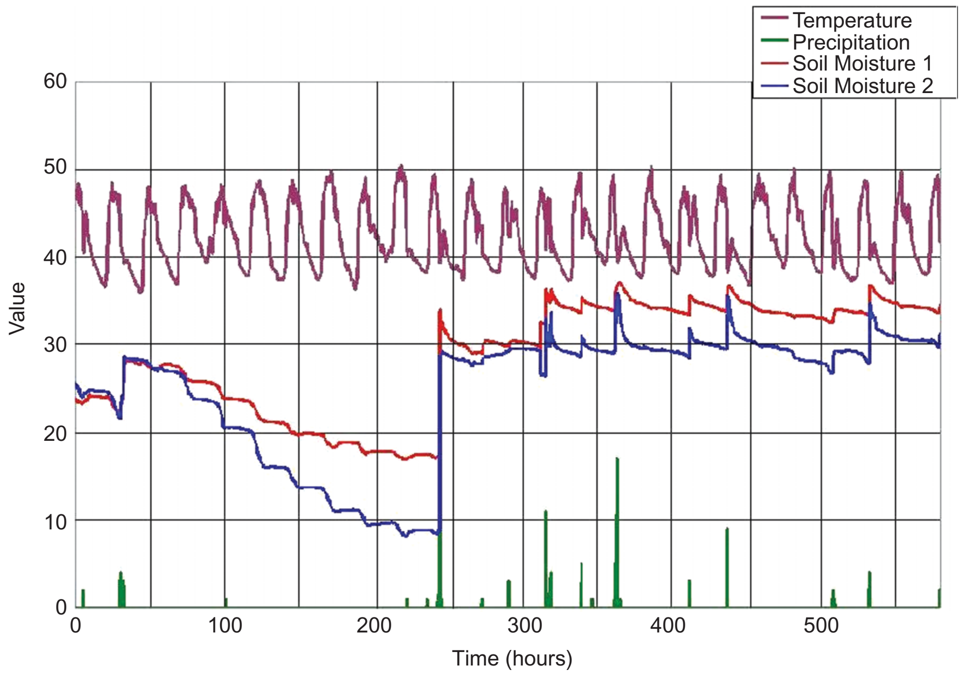

Figure 4. Temporal variation of soil moisture, rainfall, and air temperature. The units on the vertical axis are the percentage of volumetric water content for soil moisture, mm h^{-1} for rainfall, and °C for temperature. 15°C was added to temperature values, to avoid line crossing and to allow all four quantities to be plotted on the same graph.

response function associated with the instantaneous rainfall response. T_t is the air temperature at time t; h_t and h_{t-1} are the soil moisture at the time t and t-*1*, respectively; ε_t is a white noise process with zero mean and constant variance. It should be noted that the model was built using data from stations; however, the extrapolation of soil moisture was developed by using R_t from radar estimates, T_t from satellite estimates and h_{t-1} was obtained from an ANN.

The estimation of hourly soil moisture is required to perform the following additional tasks: (1) estimation of parameters of the transfer function models; (2) estimation of soil moisture trend; (3) estimation of the trend and seasonal components of the air temperature; (4) identification of similarities in spatial variability, and (5) estimation of the hourly soil moisture across the island.

3.2.1 Parameter estimation. A transfer function model was identified for each station, assuming that the coefficients of the impulse response function characterize the climatological properties for a given location. Thus, it is assumed that ωs, δs, τ and r_L coefficients that exhibited the inherent climatological characteristics of a specific location, and consequently the spatial variability are indirectly expressed by the coefficients of the impulse response function. Thus, evaluating the transfer function with data from another location that exhibits similar climatological characteristics of a known station will estimate the dynamics of the soil moisture response. It should be noted that the self-organized neural network was used to identify the spatial similarities, and this task is described in §3.2.2.

The estimation of the transfer-function parameters is not a trivial task and requires a well planned procedure as shown in Ramirez-Beltran *et al.* (2008). The estimation procedure can be summarized as follows:

(*a*) The first step consists of applying the autocorrelation function to determine whether or not the soil moisture is a stationary process. Typically, the soil moisture is a nonstationary process due to significant changes in the mean that have occurred over time, due to rainfall long-term impacts over the soil moisture response. The process becomes stationary after the trend is removed. Usually, this is accomplished by taking the first, $(1 - B)h_t$, or the second, $(1 - B)^2h_t$, difference of the process or by removing a parametric function. In this case an ANN was used to take into account the nonlinearities between the soil moisture and environmental variables. The ANN is applied to identify and remove the soil moisture trend. It has been postulated that the soil moisture trend can be expressed as a function of the following variables: the elevation, the sequence of observations (time dependency), the accumulated rainfall from the previous month, and the soil texture. The training patterns can be written as follows:

$$\boldsymbol{p}_i = \left[\frac{e_i}{1,200}, \frac{o_i}{1,000}, C_i, S_i, \alpha_i\right] \tag{7}$$

$$\boldsymbol{q}_i = [\psi_i] \tag{8}$$

where $\boldsymbol{p}$ is the input vector; e is the elevation; o is the number of the soil moisture observation, and takes the following values: $o = 1, 2, 3, \ldots, n$ and n represents the total number of hours of available observations; α is the accumulated rainfall in the previous month; C and S are the percentage of clay and sand; $\boldsymbol{q}$ is the output vector, and ψ is the trend of soil moisture after the mean is removed. The subscript i was removed to simplify notation and represents the location of the i^{th} grid at a given point in time. The elevation and the observation number were divided by a constant to facilitate the convergence of the Levenberg–Marquardt Backpropagation algorithm, which was used to train the ANN (Hagan *et al.* 1996). The trend was subtracted from the original observations, and the transformed series is known as the stochastic soil moisture component.

(*b*) The second step consists of performing a random search to determine the initial point for a nonlinear optimization routine. A set of random numbers with a uniform probability distribution was used to generate 100 points over a specific range and the mean square error (MSE) was used to identify a suitable initial point, i.e. the one that exhibits the smallest MSE from the selected sample was chosen, where an error is the difference between the observed and the estimated soil moisture for a particular point in time. The random search was restricted to a given range to speed up the convergence process. The selected empirical ranges that provide satisfactory results are given by two vectors:

$$\mathbf{min} = [-2, -2, -0.9, -2, -2, -0.9, -2, -2, -0.9, 0, 12] \tag{9}$$

$$\mathbf{max} = [2, 2, 0.2, 2, 2, 0.9, 2, 2, 0.9, 1, 96] \tag{10}$$

where **min** and **max** represent the minimum and maximum values, respectively, that can be assigned to the parameters that are organized as in the following vector:

$$\boldsymbol{p}_a = [\omega_{0,1},\quad \omega_{0,2},\quad \delta_{0,1},\quad \omega_{1,1},\quad \omega_{1,2},\quad \delta_{1,1},\quad \omega_{2,1},\quad \omega_{2,1},\quad \delta_{2,1},\quad \tau,\quad L] \tag{11}$$

(*c*) The third step consists of using the sequential quadratic programming (SQP) algorithm to estimate the parameters of the impulse response functions, while the L parameter is maintained fixed at a given constant value, and at the beginning it was selected by inspection (Reklaitis *et al.* 1983, MathWorks 2000).

(*d*) The fourth step consists of fixing the parameters of the impulse response function and the Hooke and Jeeves (HJ) algorithm was used to estimate the L parameter (Reklaitis *et al.* 1983). The HJ algorithm is a function evaluation technique, i.e. a direct searching integer procedure was used to determine the L value that minimizes the MSE.

(*e*) The last step consists of using the previous parameter values as the initial point and simultaneously implementing the SQP and HJ algorithms to determine the complete set of parameters. This task was successfully accomplished by using the Matlab software (MathWorks 2000). It should be noted that the order of the polynomial $\delta_0(B)$ is one (i.e. $\delta_0(B) = 1 - \delta_{0,1}(B)$), the decay function of the soil moisture response can be controlled by the δ coefficient, the subscripts are eliminated to simplify the description. If δ is close to zero the rainfall and temperature effect on soil moisture will disappear very fast; on the other hand if the δ coefficient is close to one the rainfall and temperature effects on soil moisture will last during several units of times. If δ is larger than one, the process becomes unstable and the estimates of soil moisture increase without control (Brockwell and Davis 2002). Therefore, the values of δ must be limited to the following range: $-1 < \delta < 1$.

Table 1 shows the parameter estimation results for the fitted Transfer Function (TF) models. The TF model parameters depend on the soil type, vegetation, topography and the atmospheric conditions of a given area. Thus, the value of the δ parameter of equation (6) controls the length of the soil moisture response, the $\omega_{0,t}$ coefficients control the instantaneous soil moisture response due to the previous soil moisture value, the $\omega_{1,i}$ coefficients control the soil moisture response due to instantaneous rainfall, and the $\omega_{2,t}$ coefficients control the soil moisture response due to the current air temperature.

3.2.2 Spatial variability. A self-organized neural network (SONN) (Hagan *et al.* 1996) was used to capture the spatial variability and to identify grid points that show similar properties to a place where a transfer function model is known, and the identified station is called the similar station. The variables used to identify similarities of climatological characteristics were: diurnal and nocturnal monthly mean temperature, elevation, vegetation index, and accumulated rainfall of the corresponding month. Once a grid was assigned to a similar station the hourly air temperature and the hourly soil moisture were computed for every grid, and therefore estimations across the island were obtained.

Twenty climatological regions were requested to the SONN; however, only 11 regions were identified. This indicates that 11 transfer function models are required to perform the extrapolation across the island and each transfer function must be located in a different region. Figure 5 shows the histogram that exhibits the SONN results, and a

Table 1. Soil moisture parameter estimation results for 15 stations.

	$\omega_{0,1}$	$\omega_{0,2}$	$\delta_{0,1}$	$\omega_{1,1}$	$\omega_{1,2}$	$\delta_{1,1}$	$\omega_{2,1}$	$\omega_{2,2}$	$\delta_{2,1}$	τ	L
Cadena	–0.0304	0.0424	0.3207	0.7382	0.7527	–0.3164	1.0711	0.3312	–0.4192	0.1329	40
Caobos	–0.0653	0.0421	–0.6045	1.1781	1.2170	–0.4822	1.1714	0.1363	–0.3139	0.1388	68
Chips	–0.2299	0.2015	0.2034	0.5752	0.5309	–0.5029	0.7561	–0.0895	0.3265	0.0370	71
Colegio	–0.0355	0.0333	–0.1960	0.2456	0.3117	–0.2324	0.9549	0.6753	–0.6438	0.0177	12
EEL	0.0398	–0.0325	0.4496	1.3382	–0.9019	0.3316	1.4310	–0.0855	–0.3538	0.0095	12
Guilarte	–0.0023	–0.0041	–0.2168	1.0824	0.6826	–0.6241	0.9684	0.4386	–0.4183	2.6946	15
Isabela	1.0738	1.2207	0.6428	0.4839	0.7223	0.2639	0.5179	–0.2920	0.7676	0.0496	76
Jamaica	–0.0244	0.0066	–0.0007	0.1760	0.3066	–0.4228	1.1537	0.5988	–0.7589	0.1083	39
Jimenez	–0.1538	0.1476	–0.2574	0.2160	0.2517	–0.0806	0.8537	0.1221	0.0115	0.0000	12
Lajas	0.0061	0.0005	0.8729	1.2440	1.3600	0.0741	0.5646	0.2352	0.1911	0.1038	70
Maricao	–0.0007	–0.0007	–0.7245	0.0068	0.0596	0.0342	1.2271	–0.0018	–0.2323	0.0000	12
Nazario	–0.0491	0.0479	–0.1906	0.1376	0.1130	–0.2296	1.0564	0.2118	–0.2766	0.0190	85
Ramon	–0.0600	0.0566	0.4033	0.4285	0.5246	–0.4255	0.7810	–0.6792	0.8994	0.0349	113
UPR	–0.1930	0.1468	0.0000	0.0000	0.4834	–0.2172	1.0177	0.4967	–0.5216	0.0300	65
Yunque	–0.0251	0.0708	–0.6699	0.3887	0.4911	–0.4594	1.2030	0.2096	–0.4274	0.0693	46

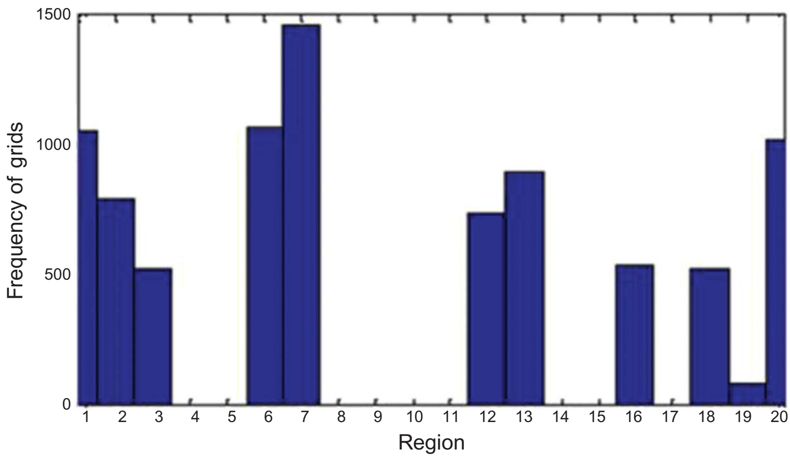

Figure 5. Histogram of climatological similarities. Regions with frequency zero do not exist, i.e. there are only 11 regions.

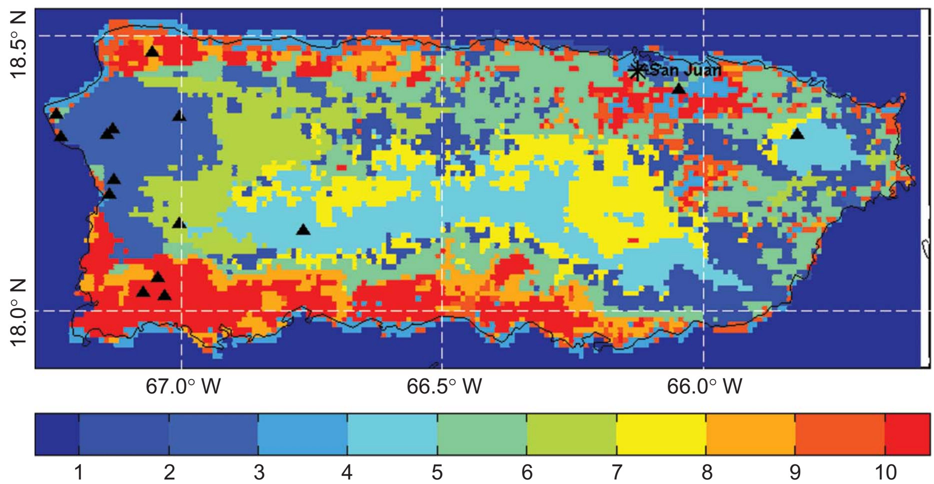

Figure 6. Areas with similar climatological properties. Eleven climatological regions were found by the self-organized neural network (SONN) and each colour shows the location of a region with similar climatic conditions. A black triangle indicates the location of a station.

region with no frequency indicates that this particular region does not exist. For instance class number 4 has no frequency and this implies there is no region assigned to this particular class. Data for the year 2006 were used to identify similarities and figure 6 shows the PR map that describes the 11 similar climatic regions.

3.2.3 Trend and seasonal components of hourly air temperature. The hourly air temperature is required to be applied as an input of the TF model. The hourly air temperature has three major components: the trend, the seasonality, and the stochastic components. The trend and seasonality are deterministic components and they will be estimated as follows: the trend of hourly air temperature is estimated by using an

ANN, the periodicity component by using a regression equation, and the stochastic component was not estimated since for our calculus it is a negligible component.

The training patterns used to estimate the trend of air temperature are the following:

$$\tilde{\boldsymbol{p}}_i = \left[\frac{e_i}{1200}, \frac{o_i}{1000}\right] \tag{12}$$

$$\tilde{\boldsymbol{q}}_i = [\lambda_i] \tag{13}$$

where $\tilde{\boldsymbol{p}}$ is the input vector; e and o were defined in equation (7); $\tilde{\boldsymbol{q}}$ is the output vector, and λ is the trend of the air temperature. Again the subscript i was omitted to simplify notation and represents the i th location.

The seasonal component was estimated by using a regression equation. A regression method has also successfully been used to estimate air temperature based on MODIS surface temperature (Mostovoy *et al.* 2005). In this paper we developed regression equations using air temperature from stations, MODIS surface temperature, and station elevation. An empirical model was developed and has the following sinusoidal form (Calderon-Arteaga 2007):

$$\Omega_{t,j} = \sum_{i=1}^{4}\left[a_{i,j}A_j \sin\left(\frac{2\pi t}{\rho_{i,j}} + \phi_j\right) + b_{i,j}A_j \cos\left(\frac{2\pi t}{\rho_{i,j}} + \phi_j\right)\right] + c_j \quad \rho_{ij} = 6, 8, 12, \text{ and } 24 \tag{14}$$

where

$$A_j = 0.5(T_{\max,j} - T_{\min,j}) \tag{15}$$

$$T_{\max,j} = 1.0068\ T_{\mathrm{d},j} - 0.0082e_j \tag{16}$$

and

$$T_{\min,j} = 1.0121\ T_{\mathrm{n},j} - 0.0801e_j \tag{17}$$

where Ω is the estimated periodic component of the air temperature, the subscript t refers to time and j refers to the location; a, b, and c are the regression coefficients; A is the semi range of the daily air temperature for each station; ϕ is also a regression coefficient and represents the phase angle, which is inherent to each station; and ρ_{ij} is the size of the ith period at the j th station, which was identified by using a periodogram. $T_{\max,\ j}$ and $T_{\min,\ j}$ are maximum and minimum air temperature for the j^{th} station, respectively. $T_{d,j}$ and $T_{n,j}$ are the MODIS daily and nightly surface temperature for the j^{th} pixel, respectively; e_j is the elevation of the j^{th} station. A total of 15 multiple linear regression equations were developed (one for each station). The phase angle ϕ was determined in such a way that the mean squared error of the regression model was minimized. Parameter estimation for the regression equations are given in Calderon-Arteaga (2007). To extrapolate the hourly air temperature to other places where no equation is available it is necessary to use the SONN to identify and apply the corresponding equation. The models were evaluated across the island using the MODIS daily land surface temperature and the digital elevations of PR.

3.2.4 Estimation of hourly soil moisture. The performed calculations are organized to obtain the hourly soil moisture, which is the final product, i.e. the ANN is used to compute the trends, and the transfer function models are used to estimate the stochastic component of the soil moisture. Thus, the trend and the stochastic components are added to create the hourly soil moisture for every grid. It should be mentioned that spatial similarities were used to extrapolate the air temperature and the soil moisture across the island. The final product from the empirical models is soil moisture on an hourly basis at 1 km of horizontal resolution and results are presented in the next section.

4. Validation and results

The described methodology was implemented for PR climatic conditions. Model validation and results are described in this section. The model validation was performed for both the long- and short-term memory models, using different strategies to test whether or not the proposed scheme has the capability of estimating the spatial and temporal variability of soil moisture and to measure the accuracy of the estimation scheme.

4.1 *Validation and results of long-term memory model*

The long-term memory model was used to express the monthly soil moisture and was developed with 2005 data and validated with 2006 observations. Since the available number of stations is small, the validation strategy is the leave-4-out method and consists of removing four stations from the dataset at a time and training the ANN with the remaining stations. The estimates of the monthly soil moisture for the four removed stations are compared with the observed values. For instance, figure 7 shows the comparison between the observed and the estimated monthly soil moisture during March 2006. The validation exercise was repeated five times with different initial weights to derive the central tendency of the estimation. The mean absolute error (MAE) and the mean squared error (MSE) are the preferred measurements of

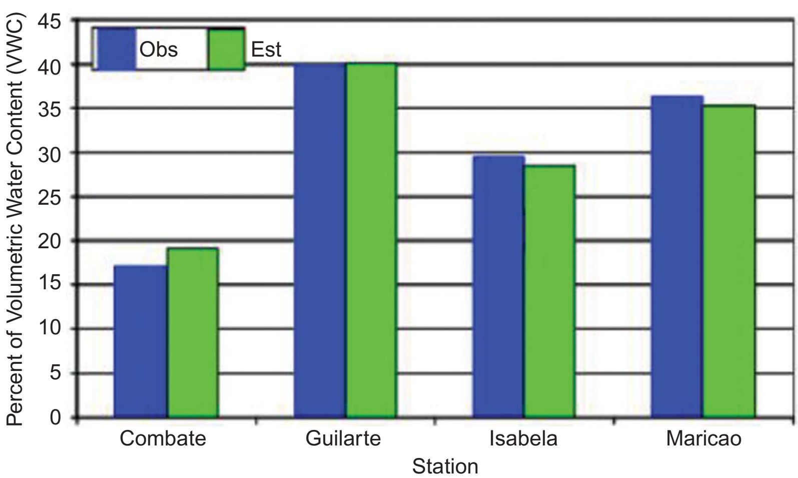

Figure 7. Comparison between the observed (Obs) and estimated (Est) soil moisture for March 2006.

estimation accuracy. The validation process was implemented for each month during 2006. Table 2 shows the MAE and the MSE during the validation period and the corresponding averages are 2.72% and 12.32, respectively.

The hourly rainfall data from NEXRAD were used to obtain the monthly accumulated rainfall, the NDVI was extracted from satellite data and both were used to estimate the monthly average of soil moisture across PR. Figure 8 shows an example of the soil moisture average for August 2006. This figure shows that on the average the largest values of the soil moisture are mostly located in the north and western parts of the island and the largest values are around 40% of the volumetric

Table 2. Monthly averages of soil moisture accuracy. Data are given as percentages of volumetric water content. The validation exercise was repeated five times with different initial weights to derive the central tendency of the estimation.

	MAE	MSE		MAE	MSE
January	2.34	11.95	July	4.15	28.43
	2.90	10.67		3.88	24.18
	2.00	5.12		4.59	31.77
	1.67	3.55		4.62	32.48
	1.80	4.51		2.66	8.57
Average	**2.14**	**7.16**	Average	**3.98**	**25.08**
February	3.53	19.96	August	3.03	13.27
	3.09	26.21		3.45	14.91
	1.89	7.74		2.60	10.45
	2.34	8.55		3.49	12.65
	3.33	18.40		3.08	12.10
Average	**2.83**	**16.17**	Average	**3.13**	**12.68**
March	1.38	2.39	September	3.65	16.03
	2.42	10.60		3.00	13.53
	2.77	8.93		2.57	7.73
	3. 19	14.75		3.49	20.33
	2.74	10.09		2.57	7.56
Average	**2.50**	**9.35**	Average	**3.05**	**13.04**
April	2.67	8.26	October	2.34	11.95
	1.09	1.60		3.53	19.96
	1.55	2.56		4.62	32.48
	2.18	7.33		1.70	4.17
	2.50	13.42		3.33	18.40
Average	**2.00**	**6.63**	Average	**3.10**	**17.39**
May	2.41	8.44	November	3.45	14.91
	1.27	1.94		3.16	18.33
	2.40	8.33		2.58	10.35
	1.39	2.33		2.63	9.14
	2.31	9.87		2.93	11.51
Average	**1.96**	**6.28**	Average	**2.95**	**12.85**
June	3.12	12.11	December	2.10	7.70
	2.47	10.54		2.99	16.56
	1.70	4.17		2.55	14.56
	3.12	11.95		1.24	2.15
	3.45	17.12		2.18	9.54
Average	**2.77**	**11.18**	Average	**2.21**	**10.10**

MAE, mean absolute error; MSE, mean squared error.

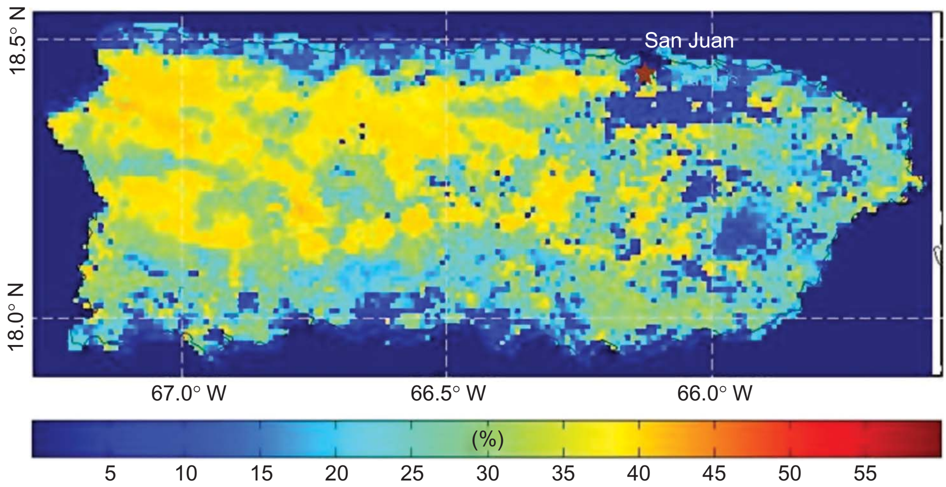

Figure 8. Average estimates of soil moisture for August 2006.

water content. It should be mentioned that these results can be used to initialize the short-term memory model of soil moisture for any day and at a grid during August 2006.

4.2 *Validation and results of the short-term memory model*

The short-term memory model is validated by applying two different strategies. Using data from the same station and using information from a similar station.

(a) The first strategy consists of dividing the available data for a given station into two parts. The first part of the time series is used to build the model and the second part is used to estimate the soil moisture using the identified TF model and to compare the observed with the estimated values. An example of model validation is given in figure 9, which shows the observed and estimated soil moisture from the transfer function at Ramon station from 1 August to 30 September 2006. This figure exhibits the performance of the transfer function model during the model fitting and validation processes. The blue line shows the observed values of the soil moisture, the green line shows the estimated values during the fitting process, the red line shows the estimated values during the validation period, and the asterisk shows the time at which a rainfall event has occurred.

Rainfall and air temperature data observed on an hourly basis were used to evaluate the TF model. Model validation results at 20 cm depth are presented in table 3. The first column shows the name of the soil moisture stations. The second column shows the MAE as a percentage of volumetric water content during the model fitting process. The third column shows the MSE during the model fitting process. The fourth and the fifth columns show the MAE and MSE during the validation process, respectively.

(b) The second validation strategy consists of leaving a set of variables out and estimating the soil moisture across PR. The soil moisture of a removed station is estimated by using the TF model of a similar station and data from a place where the TF model is unknown. For example suppose we want to estimate the soil moisture at

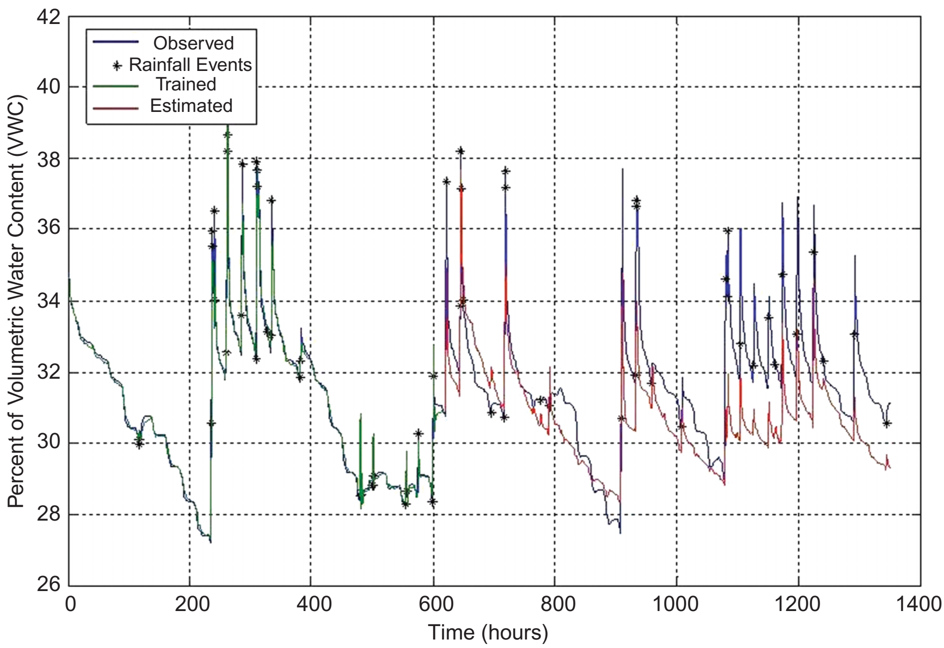

Figure 9. Fitting and validation process of the transfer function (TF) in Ramon station.

Table 3. Hourly soil moisture validation using data from the same station. Data are given as percentages of volumetric water content.

	Model fitting		Model validation	
Station	MAE	MSE	MAE	MSE
Cadena	0.11	0.08	1.21	2.17
Caobos	0.09	0.05	2.08	7.12
Chips	0.21	0.19	1.70	3.96
Colegio	0.15	0.14	2.31	7.60
EEL	0.12	0.09	3.43	13.05
Guilarte	0.04	0.00	0.28	0.16
Isabela	0.16	0.11	1.08	2.09
Jamaica	0.07	0.02	1.29	3.19
Jimenez	0.17	0.16	2.32	7.46
Lajas	0.11	0.06	1.82	4.82
Maricao	0.05	0.01	0.78	1.21
Nazario	0.09	0.02	0.78	0.97
Ramon	0.15	0.15	1.13	2.00
UPR	0.23	0.36	3.08	13.44
Yunque	0.13	0.06	1.93	5.29
Average	0.13	0.10	1.68	4.97

MAE, mean absolute error; MSE, mean squared error.

the pixel where Jamaica station is located. The SONN was used to identify a similar station (Chips) since it exhibits similar climatic characteristics to Jamaica station. Thus the TF from Chips was evaluated with data from Jamaica station and the results are shown in figure 10. Estimates were compared with the actual soil moisture observations recorded at Jamaica station. Table 4 shows the accuracy of the soil moisture estimation based on a similar station. The errors are the difference between the estimates and the observed soil moisture values from the removed station. The last column of table 4 shows the similar station used. The average of the MAE and the MSE for the removed stations were 3.34%, and 16.96(%)^2 of volumetric water content, respectively.

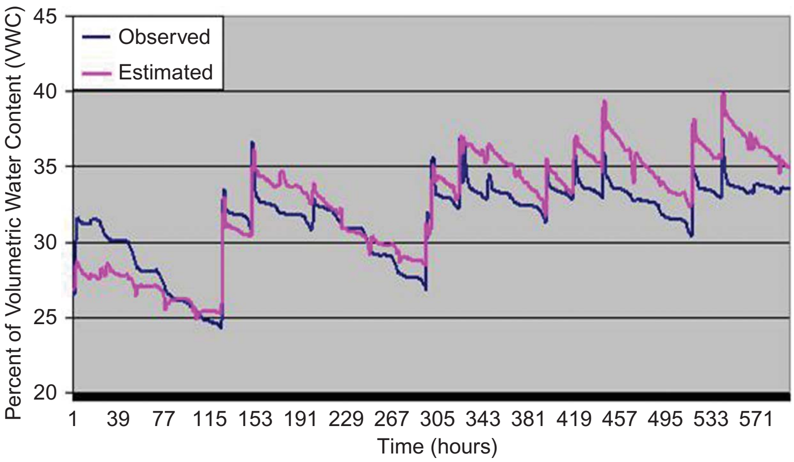

Figure 10. Validation of the transfer function (TF) model of the Chips station with data from Jamaica station (September 2006). This figure shows the second validation technique, which consists of using the TF model from the similar station (Chips) and data from the other station (Jamaica) located in the same region.

Table 4. Soil moisture validation using a similar station. Data are given as percentages of volumetric water content.

Station	MAE	MSE	Region	Similar station
Cadena	4.51	22.29	2	Chips
Colegio	1.15	1.95		
Jimenez	2.59	12.45		
Ramon	3.45	14.17		
Jamaica	1.49	3.47		
Guilarte	2.91	10.32	6	Caobos
Yunque	2.36	7.44		
EEL	5.54	35.61	20	Nazario
Maricao	6.10	44.95		
Average	3.34	16.96		

MAE, mean absolute error; MSE, mean squared error.

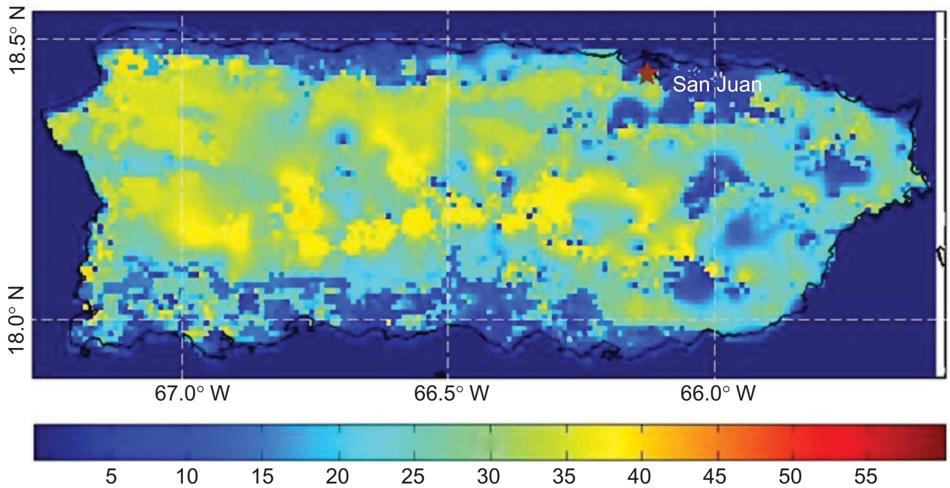

Figure 11. Soil moisture estimates at 10:00 am on 5 August 2006.

The complete methodology was implemented to estimate the hourly soil moisture across PR. A month and a day during 2006 were arbitrarily selected to perform an estimation exercise. For instance, figure 11 shows the soil moisture estimation at 10:00 am, for 5 August 2006. This figure reveals the spatial variation of the soil moisture for the selected point in time. It should be mentioned that there are a few regions where no transfer functions are available. Therefore, in those regions where similar TF models were not available, the estimation of the soil moisture was derived using a kriging algorithm (Wackernagel 2003).

Twenty four hours of spatial and temporal variation of soil moisture are given in figures 12 and 13. Figure 12 shows the soil moisture variation from 0:00 am to 11:00 am and figure 13 from 12:00 to 23:00 hours. On this particular day, these figures show that the soil moisture increases during the night-time due to rainfall events that occurred on this particular day and decreases during the daytime due to the temperature effect.

5. Conclusions

A new methodology for estimating soil moisture over densely vegetated areas is proposed. The estimation algorithm includes a short-term memory model and a long-term memory model. The long-term memory model is a neural network algorithm that estimates the monthly variation of soil moisture based on rainfall from radar, vegetation index, surface temperature from satellite, soil texture and topography. The short-term memory model is a nonlinear stochastic transfer function model that estimates the hourly soil moisture variations based on short-term rainfall from radar and surface temperature from satellite data. A transfer function model was developed at each soil moisture station and a self-organized neural network was used to identify 'a similar station'. Finally, the transfer function model from a similar station is used to extrapolate over grids where no transfer function models are available. There are few places where no transfer function can

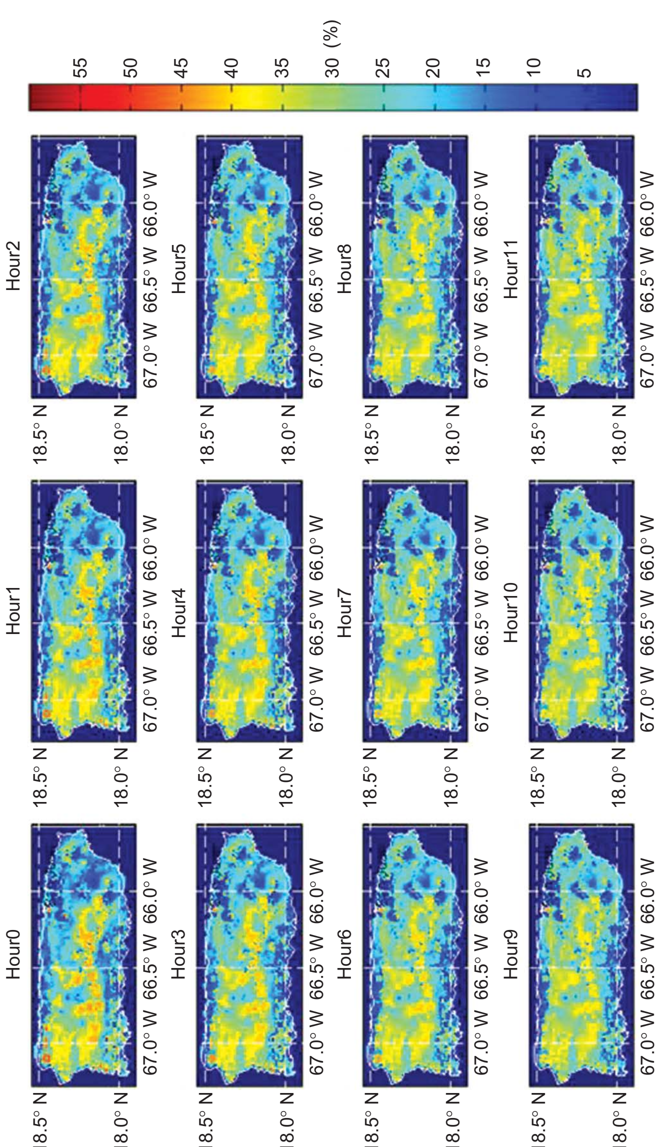

Figure 12. Hourly estimation of soil moisture, from 0:00 to 11:00 on 5 August 2006. The kriging algorithm was used to interpolate the grids where a similar transfer function (TF) model was not available.

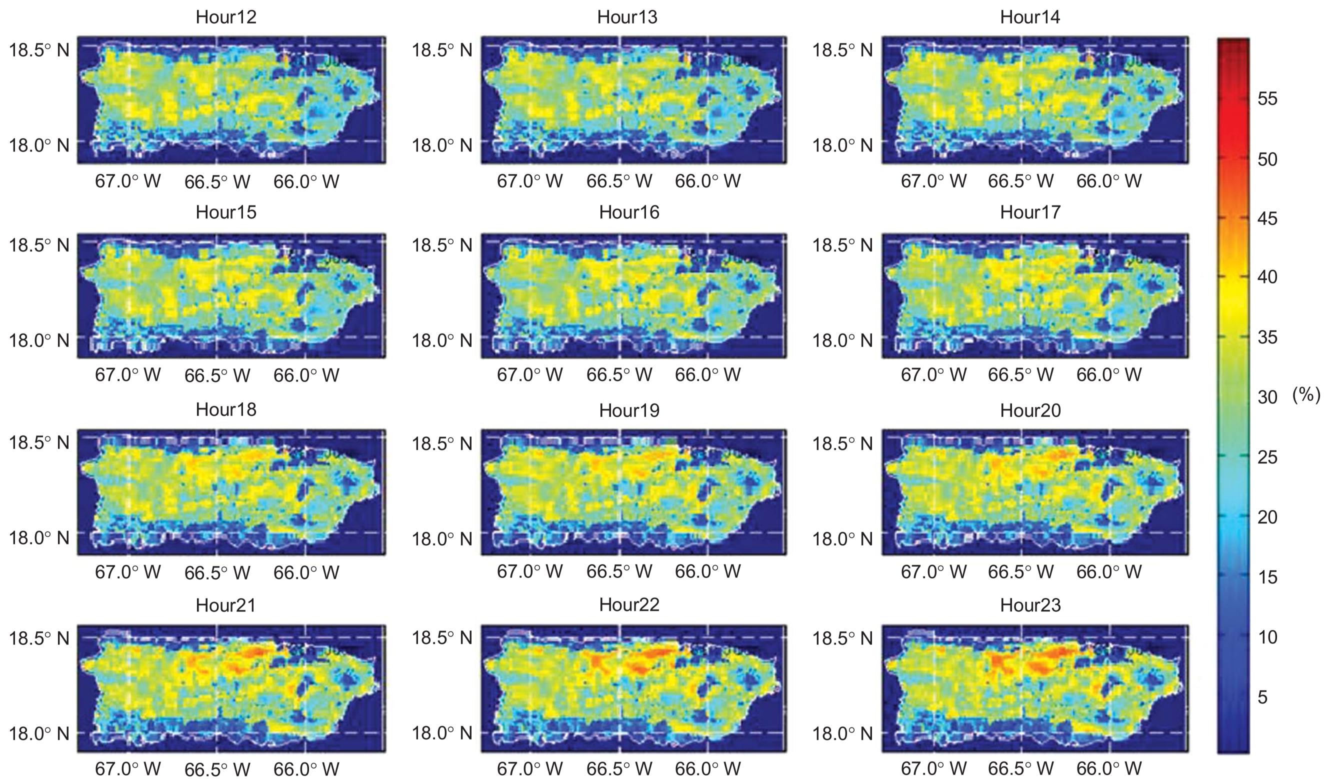

Figure 13. Hourly estimation of soil moisture, from 12:00 to 23:00 on 5 August 2006. The kriging algorithm was used to interpolate the grids where a similar transfer function (TF) model was not available.

be applied to estimate soil moisture, and those places are estimated by using a kriging algorithm.

Validation techniques were implemented to measure the estimation accuracy of the developed models. The long-term memory model shows that the mean absolute error and the mean square error during the validation period were 2.72% and 12.32%, respectively. The validation of the short-term memory model was performed using two strategies. The first strategy consisted of dividing the time series of a given station into two parts. The first part of the time series was used to build a stochastic transfer function model and the second part to estimate the soil moisture. The second validation strategy consists of leaving a set of stations out and selecting similar transfer function models to estimate the soil moisture for the removed stations. The estimated soil moisture values were compared with the observed values. The mean absolute error and the mean square error when using data from the same station to perform the estimation were 1.64% and 4.96%, respectively. The mean absolute error and the mean square error when using a similar station to perform the estimation were 3.34% and 16.96%, respectively. Validation results show that the proposed methodology provides a reasonably small error over vegetated areas, which are located across the island. Thus, this methodology can potentially be implemented over tropical areas, characterized by densely vegetated areas with complex topography. The proposed methodology may be a valuable tool for government agencies and private companies located in PR concerned with weather and climate monitoring, runoff and flood control.

Artificial neural networks were used in this research mainly with two purposes. A self-organized ANN was used to identify spatial variability and a feedforward ANN was used for nonlinear modelling. The spatial variability was used to select stations that exhibit similar climatic characteristics, and then to properly apply the transfer function model for estimating the short-term soil moisture variability across the island. The self-organized ANN uses the Kohonen learning rule to perform a robust classification of the spatial variability whereas the feedforward ANN uses the Levenberg–Marquardt Backpropagation algorithm to estimate the long-term soil moisture variation.

The proposed algorithm can be used to generate the soil moisture initial conditions to run a regional atmosphere model. For instance, the RAMS requires accurate soil moisture initial conditions to perform atmospheric dynamic simulations. There is also a possibility of introducing the soil moisture in the numerical model to compare soil moisture observation with estimates at different points in time and consequently improve the performance of the atmospheric dynamic simulation.

Acknowledgments

This research has been supported by NASA-EPSCoR grant NCC5-595, by the NOAA-CREST grant NA17AE1625 and also by the University of Puerto Rico at Mayagüez. The authors appreciate and recognize the funding support from these institutions.

References

Balsamo, G., Bouyssel, F. and Noilhan, J., 2004, A simplified bi-dimensional variational analysis of soil moisture from screen-level observations in a mesoscale numerical weather-prediction model. *Quarterly Journal of the Royal Meteorological Society*, **130**, pp. 895–915.

BOX, G.E.P. and JENKINS, G.M., 1976, *Time Series Analysis Forecasting and Control* (Oakland, CA, USA: Holden-Day).

BREIDENBACH, J.P. and BRADBERRY, J.S., 2001, Multisensor precipitation estimated and produced by National Weather Service River Forecast Centers for Hydrological Applications. In *Proceedings 2001 Georgia Water Resources Conference*, J. Hatcher (ed.), Institute of Ecology, University of Georgia, Athens, pp. 179–182.

BROCKWELL, P.J. and DAVIS, R.A., 2002, *Introduction to Time Series and Forecasting*, 2nd edn (New York: Springer-Verlag).

CALDERON-ARTEAGA, C.H., 2007, Redes Neuronales Artificiales y Funciones de Transferencia aplicadas a la Estimación de Humedad de Suelo y Temperatura del Aire. MSc thesis, Department of Mechanical Engineering, University of Puerto Rico, Mayaguez, Puerto Rico.

CLARK, C.A. and ARRITT, R.W., 1995, Numerical simulations of the effect of soil moisture and vegetation cover on the development of deep convection. *Journal of Applied Meteorology*, **34**, pp. 2029–2045.

COMARAZAMY, D.E., 2001, Atmospheric Modeling of the Caribbean Region: Precipitation and Wind Analysis in Puerto Rico for April 1998. MSc thesis, Department of Mechanical Engineering, University of Puerto Rico, Mayaguez, Puerto Rico.

FAST, J.D. and MCCORKLE, M.D., 1991, The effects of heterogeneous soil moisture on a summer baroclinic circulation in the central United States. *Monthly Weather Review*, **119**, pp. 2140–2167.

GALLUS JR, W.A., and SEGAL, M., 2000, Sensitivity of forecast rainfall in a Texas convective system to soil moisture and convective parameterization. *Weather Forecasting*, **14**, pp. 509–525.

GOLAZ, J.-C., JIANG, H. and COTTON, W.R., 2001, A large-eddy simulation study of cumulus clouds over land and sensitivity to soil moisture. *Atmospheric Research*, **59–60**, pp. 373–392.

HAGAN, M.T., DEMUTH, H.B. and BEAL, M., 1996, *Neural Network Design* (Boston, MA: PWS).

HAGAN, M.T. and MENHAJ, M., 1994, Training feedforward networks with the Marquardt algorithm. *IEEE Transactions and Neural Networks*, **5**, pp. 989–993.

HAMILTON, J.D., 1994, *Time Series Analysis* (Princeton, NJ: Princeton University Press).

HUANG, J., VAN DEN DOOL, H. and GEORGAKAKOS, C.P., 1996, Analysis of model-calculated soil moisture over the United States (1931–93) and application to long range temperature forecasts. *Journal of Climate*, **9**, pp. 1350–1362.

JACKSON, T.J., 1982, Passive microwave sensing of soil moisture under vegetation canopies. *Water Resources Research*, **16**, pp. 1137–1142.

JIANG, H. and COTTON, W.R., 2004, Soil moisture estimation using an artificial neural network: a feasibility study. *Canadian Journal of Remote Sensing*, **30**, pp. 827–839.

MATHER, J.R., 1978, *The Climatic Water Balance in Environmental Analysis* (Lexington, MA: D.C. Heath).

MATHWORKS, 2000, *Optimization Toolbox: User Guide* (Natick, MA: MathWorks).

MOSTOVOY, G.V., KING, R., REDDY, K.R. and KAKANI, V.G., 2005, Using MODIS LST data for high-resolution estimates of daily air temperature over Mississippi. In *IEEE Proceedings of 3rd International Workshop of Multi-temporal Remote Sensing Images*, 16–18 May 2005, Biloxi, Mississippi, pp. 76–80.

NJOKU, E.G. and ENTEKHABI, 1996, Passive microwave remote sensing of soil moisture. *Journal of Hydrology*, **184**, pp. 101–129.

NJOKU, E.G. and LI, L., 1999, Retrieval of land surface parameters using passive microwave measurements at 6–18 GHz. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, pp. 79–93.

NJOKU, E.G., JACKSON, T.J., LAKSHMI, V., CHAN, T.K. and NGHIEM, S.V., 2003, Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 215–229.

PAN, F., PETERS-LIDARD, C.D. and SALE, M.J., 2003, Analytical method for predicting surface soil moisture from rainfall observations. *Water Resources Research*, **39**, pp. 3–1–3–12, doi: 10. 1029/2003 WR002142.

PANDIT, S.M. and WU, S.M., 1983, *Time Series and Systems Analysis with Applications* (New York: John Wiley and Sons).

RAMÍREZ-BELTRAN, N. D. and MONTES, J. A., 2002., Neural Networks to model dynamic systems with time delays, *IIE Transactions,* **34**, pp. 313–327.

RAMÍREZ-BELTRAN, N.D., CASTRO, J.M., HARMSEN, E. and VÁSQUEZ, R., 2008, Stochastic transfer function model and neural networks to estimate soil moisture. *Journal of the American Water Resources Association*, **44**, pp. 847–865.

REKLAITIS, G.V., RAVINDRAN, A. and RAGSDELL, K.M., 1983, *Engineering Optimization: Methods and applications* (New York: John Wesley and Sons).

SUN, D. and PINKER, R.T., 2004, Case study of soil moisture effect on land surface temperature retrieval. *IEEE Geoscience and Remote Sensing Letters*, **1**, pp 127–130.

THORNTHWAITE, C.W., 1948, An approach toward a rational classification of climate. *Geographical Review*, **38**, pp. 55–94.

WACKERNAGEL, H., 2003, *Multivariate Geostatistics: An Introduction with Applications*, 3rd edn (Berlin: Springer).

WANG, J. and SCHMUGGE, T., 1980, An empirical model for the complex dielectric permittivity of soils as a function of water content. *IEEE Transactions on Geoscience and Remote Sensing*, **18**, pp. 288–295.

WEI, W.S., 1990, *Time Series Analysis: Univariate and multivariate methods* (Redwood City, CA: Addison-Wesley).

WETZEL, P.J. and WOODWARD, R.H., 1987, Soil moisture estimation using GOES-VISSR infrared data: a case study with a simple statistical method. *Journal of Climate and Applied Meteorology*, **26**, pp. 107–117.