Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 11-15.

ARTIFICIAL NEURAL NETWORK TO ESTIMATE INTEGRATED WATER VAPOR USING SATELLITE DATA FROM HSB SENSOR.

Rosangela Saher Cintra*

José Demisio Simões Silva

Instituto Nacional de Pesquisas Espaciais – Brazil

1. INTRODUCTION

The steam of water, present in the atmosphere, performs a basic role to regulate the climate of our planet. The steam of water represents the connection between the land surface and the mantle of air that covers it. The content of the steam of water in the air depends on the magnitude of the processes of evaporation and precipitation (Randel et al., 1998). The transport of the steam in the atmosphere constitutes a very important component of the hydrological cycle, through which, great quantity of water is taken, in the form of steam, from a region to another. The Oceans and the great reservoirs of water in the surface of the Earth provide the atmosphere by means of evaporation process. This process reduces the moisture of the ground, until the precipitation comes back, and also affects the continents.

The knowledge of the vertical and horizontal distribution of the steam of water in global scale is applied for applications of numerical prediction, climatic modeling and climate global changes studies.

There are several forms to estimate the quantity of steam of water in the air. By using remote sensor data of the atmosphere, the meteorological satellites vertical soundings have been estimating the Total Precipitable Water (or atmospheric total steam of water).

This estimation corresponds to the height of water that would be formed on the surface if the whole steam of water in the vertical atmospheric column under sounding were coming to condense and to rain.

Due to its importance, several methods have been developed to measure and to monitor the behavior of the (IWV) "Integrated Water Vapor contents" in the atmosphere. They range from simple techniques, from earth surface temperature and pressure measurements, up to high precision methods, which employ sophisticated devices like radiometers and radiosondes.

The NASA EOS (Earth Observing System) program launched the AQUA satellite with a moisture sensor supplied by Brazil, the HSB (Humidity Sounder for Brazil). The Brazilian Institute for Space Research (INPE) tested the moisture sounder for obtaining information of the content of water vapor in the atmosphere. When used together instruments, both the AMSU-A (Advanced Microwave Sounding Unit-A) and the AIRS (Atmospheric Infrared Sounder), also aboard the AQUA, allow the inference of vertical profiles of atmospheric temperature and moisture under clear and cloudy sky conditions. NASA describes Aqua as focusing on the multi-disciplinary study of Earth's interrelated processes – atmosphere, oceans, and land surface– and their relationship to changes in the Earth system. The HSB is a sensor with the same characteristics of the sounder AMSU-B (on board the NOAA-KLM (National Oceanic and Atmospheric Administration) satellites series), however, with one channel less.

There were two sets of AQUA satellite data one was simulated with RTTOV-7 radiative model and other was measured by satellite sensor. They were combined and were used at the entry for Artificial Neural Network (ANN). Both were brightness temperatures data from HSB channel. These analyses data were developed from the RaCCI/LBA (Radiation, Cloud, and Climate Interactions/Large Scale Biosphere Atmospheric Experiment in Amazônia) experiment data. The data are collected during the period of September and October 2002, in Rondônia.

The LBA project is an international initiative of research released by Brazil. LBA objective is to produce new knowledge to understand the functioning of climate, ecology, biogeochemistry, hydrology, in the Amazon region, as well as the impact of the changes, of the use of the land and of the interactions between the Amazon region and the bio-geophysical global system of the planet. The campaign RaCCI (DRYTOWET) is inserted in the LBA and it took as objectives the collection of information on precipitation during the dry station and the rainy season in the state of Rondônia. Some data are specific to evaluate the HSB sensor. The use of ANN applied to estimate observations makes easy to establish linear relationship between them. This is the dynamic characteristic of almost all the meteorological phenomena. We propose the use of ANN to reach bigger performance and computational easiness to estimate Total Precipitable Water (TPW), like favorite method.

The aim of this work is to test the quality of the ANN estimated IWV by comparison with the radiosonde values and a linear regression method.

2. METHODOLOGY

2.1 HSB Data

The temperature brightness (Tb) is a measure of the intensity of thermal radiation given out by an object. This measure is in units of temperature,

**corresponding author address*: Rosângela S. Cintra, INPE, Laboratório de Computação e Matemática Aplicada, São José dos Campos, SP; e-mail: rosangela.cintra@lac.inpe.br

because it is correlated with the intensity of radiation and the physical temperature of the radioactive body.

The Aqua satellite was launched for global change research purposes. Aqua's scientific instruments are used to provide: atmospheric temperature and humidity profiles, clouds, precipitation and radiative balance; terrestrial snow and sea ice; sea surface temperature and ocean productivity; soil moisture; improvement of numerical weather prediction; monitoring of terrestrial and marine ecosystem dynamics.

Each of the six scientific instruments of the satellite is designed to monitor a different part of the global plumbing system. They are: the Atmospheric Infrared Sounder (AIRS); the Advanced Microwave Sounding Unit (AMSU-A); the Humidity Sounder for Brazil (HSB); the Advanced Microwave Scanning Radiometer for EOS (AMSR-E); the Moderate-Resolution Imaging Spectroradiometer (MODIS); and the Clouds and the Earth's Radiant Energy System (CERES).

The HSB is a 4-channel microwave sounder launched to obtain humidity profiles of the atmosphere. Its multi-channel passive radiometer for humidity profiling consists of four channels with frequencies 150, 183.31±1, 183.31±3, 183.31±7 GHz, and spatial resolution 13.5 km horizontal at nadir.

In this paper, we used the *Tb* from HSB sensor as in Table 1. These data were collected from the satellite passages at experimental sites of 0600 and 1800 GMT, from September 2 until October 30, 2002. These data are available at the Environment Satellites Division at INPE, in Cachoeira Paulista, Brazil.

Table 1- characteristics of HSB channels

Channel	**Frequency (GHz)**	**Weight Function (hpa)**
1	150 ± 0.9	Prox. surface
2	183,31 ± 1	400
3	183,31 ± 3	600
4	183,31 ± 7	750

The brightness temperatures were using the radiative model of transfer RTTOV-7, described by Matricardi et al (2001). This model uses information of the profiles of temperature, moisture and liquid water.

2.2 Radiosondes Data

The temperature and moisture profiles of radiosondes were also collected during the RaCCI/LBA experiment. The sites of the experiment were: Guajará Mirim (10,8ºS; 65,38ºW), Ouro Preto D'Oeste (10,75ºS; 62,36ºW), Rebio Jaru (10,18ºS; 62,9ºW) and Porto Velho (8,72ºS; 63,90ºW). In these sites the radiosondes were launched from four up to six times of each day. In Guajará Mirim a specific campaign was carried out, with the initiative of the scientific management of the HSB/INPE, for sensor validation. This campaign MAIN objective was to get atmospheric profiles of temperature and moisture to be compared to the profiles retrieved by using the HSB sensor, at the same instant of its passage. Additional radiosondes were launched minutes before the timetable of each passage of the satellite in Guajará Mirim.

Lima (2004) carried out comparisons between the measured *Tb* from the HSB channels with *Tb* simulated by the RTTOV-7 and with radiosondes in Guajará Mirim. Lima proposed a linear regression method whose results are compared to the obtained ANN results. The correlation coefficients among the simulated and measured brightness temperatures obtained in this work were high, with values around 0,993. The bias and rms (root mean square) error values were around (1,31 K) and (1,65 K), respectively.

The IWV were calculated using equation (1) that integrates the absolute moisture of the steam of water (ρ_w) from the surface (h_o) up to the altitude in which there is the steam of water (h) in a unit column of dry air (Vianello and Alves, 1991).

$$IWV = \int_{ho}^{h} \rho_w \, dh \qquad (1)$$

2.3 Linear Regression Method

A method to estimate IWV from ***Tb*** of satellites data (Lima, 2004) is based on the statistical correlation applied by Kakar and Lambrigtsen (1984). It considers the hypothesis that IWV, in a layer of the atmosphere, can be obtained through a linear combination of the HSB channels. Thus, the IWV in the layer can be estimated by:

$$IWV(n) = a(n) + \sum_{i=1}^{N} b_i(n) Tb(v_i) \qquad (2)$$

where n is the number of channels, $a(n)$ and $b_i(n)$ are regression coefficients of channels combinations, and $Tb(v_i)$ is the brightness temperature measured in frequency v_i. The layers were chosen so that each one represented the maximum of the weight function for the HSB channels.

The estimation of the regression coefficients in Equation (2) used three hundred and fifty radiosondes launched in the experimental sites for all timetables. The regression equations for each layer and the selected channels are shown in the Table 2.

This method was used as a reference in this work and the statistical results of this method will be used for comparison with the results of the ANN.

The ANN applications are still a challenge for meteorologist investigators. Though an ANN provides a way to establish a non-linear relation of the characteristics of the meteorological phenomena, the wrapped physics is not possible to explain with precision, as does the method presented by Lima (2004), what justifies the comparison of its results with the ANN method.

Table 2 - Linear regression solutions and multiple coefficients from correlation for each atmospheric layer.

Camada	Equação de Regressão	r
SFC_800	-58,81 + 0,266*TB1 + 0,645*T_com + 0,109*(TB4-TB2)	0,46
800_600	142,59 - 0,452*TB1 + 0,772*T_com - 0,318(TB4-TB3)	0,86
600_400	169,51 + 0,245*TB3 - 0,834*TB4	0,95
400_200	21,50 - 0,051*TB2 - 0,030*TB3	0,90

2.4 Artificial Neural Network

The application of a Multilayer Perceptron Neural Networks was designed to estimate precipitable water in the sites of the experiments.

Multilayer Perceptrons (MLPs) are feed-forward neural networks trained with the standard back-propagation algorithm which is a supervised learning algorithm, thus requiring a desired response to be trained. It learns how to map input data into a desired response. A Multi-layer Perceptron is made up of several layers of neurons. Each layer is fully connected to the next one.

Each neuron of a layer other than the input layer computes first a combination of the outputs of the neurons of the previous layer, plus a bias. The coefficients of the combinations and the biases are called the synaptic weights.

The supervised learning problem of the MLP can be solved with the *back-propagation algorithm* that consists of two steps, see Figure 2. In the *forward pass*, the predicted outputs corresponding to the given inputs are evaluated. In the *backward pass*, partial derivatives of the cost function with respect to the different parameters are propagated back through the network. The chain rule of differentiation provides the computational rules for the backward pass. The network weights can then be adapted using any gradient-based optimization algorithm. The whole process is iterated until the weights have converged.

The function a_j denotes the activation function of each node in the hidden layers. A sigmoid activation function is used:

$$a_j(t+1) = \frac{1}{1 + e^{-(\sum w_{ij} o_i(t) - b_j)}} \quad (3)$$

where $a_j(t+1)$ is the activation function of neuron *j* in layer *L* at time step (*t*+1); w_{ij} is the weight of connection between unit *i* in the previous layer (*L*-1) and unit *j* in layer *L*; o_i is the output of unit *i* in layer (*L*-1) and b_j is the bias of unit *j* in layer *L*. This ensures that the node acts like a threshold device. In the design proposed in this paper, the output layer is activated with a linear function:

$$y_j(t+1) = \sum_i w_{ji} a_i(t) \quad (4)$$

where $a_i(t)$ is the output of unit *i* of previous layer at time step *t*, and $y_j(t+1)$ is the output of neuron *j* in the output layer.

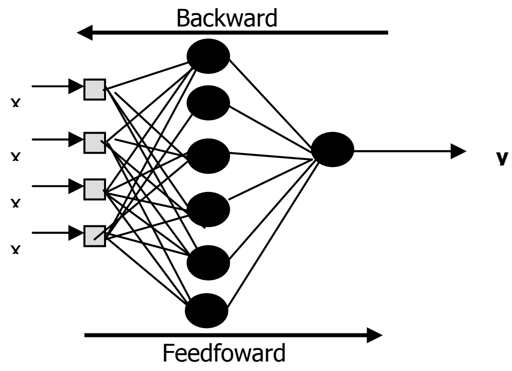

Figure 1 – The Multilayer Perceptron (MLP) topology used for application of precipitable water.

3. THE APLICATION

The training phase of the MLP required:

- Input data (***x***) of simulated and measured brightness temperatures in four channels of the HSB (Tb1, Tb2, Tb3, Tb4) at four sites of the experiment.
- Output Data (***y***) as the total IWV integrated from observations of moisture profiles of radiosondes measured in the sites of the experiment.

The training set was formed by the data from Guajará Mirim. The Tb data were simulated by HSB channels' data at all timetables. The IWV was based on all profiles measured in the experiment RaCCI/LBA.

The testing set was formed by the data from Rebio Jaru (RJ), Porto Velho (PV), Ouro Preto D'Oeste (OP) used like standards and radiosondes profiles at all timetables.

For the activation or generalization tests the Tb data of Guajará Mirim (GM) measured by the channels of the HSB and the calculated IWV of the radiosonde profiles at timetables of 0600 GMT and 18 GMT were used. These data were also used in the reproduction of the results obtained by the linear regression method proposed in Lima (2004).

Several tests were done with different architectures; the one that provided better results was a two-layer MLP with six neurons in the hidden layer. Considering a criterion of the least squared error and the best generalization did the choice of the number of neurons in the hidden layer. The ANN was trained with 20.000 "epochs", using the learning rate of 0.1. The training obtained the least error of 0.0024, thus achieving a convergence with a good computational performance.

4. RESULTS

The results are punctual and they are presented in the temporal variability. This variability is shown in days of the Julian calendar, initiate at 255 of 2002 equivalent to the day September 2nd. The linear regression (Rgr) method was applied for checking the results using the same data.

Figures (figures 2 to 6) present the estimated IWV by the radiosondes and ANN for all the sites in this study. From the graphs one may infer a high correlation between the estimated IWV by the ANN and the radiosondes data. The total value of precipitable water in Kg/m^2 is close to the integrated values of the radiosondes profiles of absolute moisture in the LBA/RaCCI experiment.

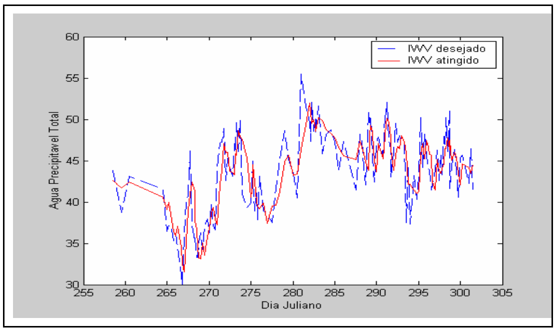

Figure 2 Results from the estimated IWV from Tb profiles. Radiosondes in Guajará Mirim (blue/diamond) and from MLP (red /asterisc).

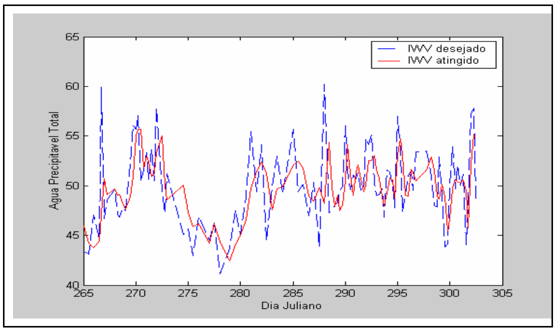

Figure 3 – Results from the estimated IWV from profiles. Radiosonde in Porto Velho (blue/diamond) and MLP (red /asterisc).

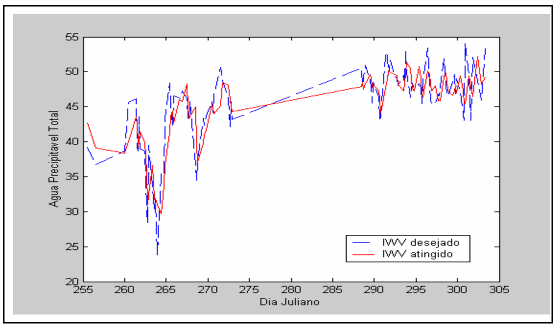

Figure 4 – Results from the estimated IWV from profiles. Radiosonde in Rebio Jaru (blue/diamond) and MLP (red /asterisc).

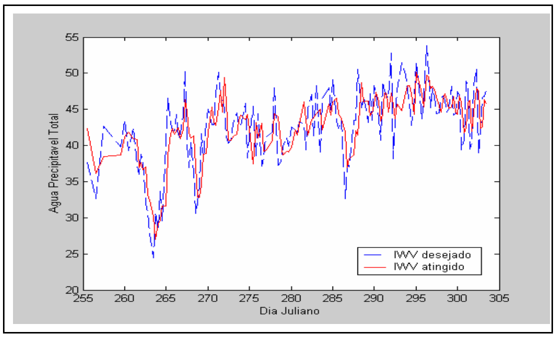

Figure 5 – Results from the estimated IWV from profiles. Radiosonde in Ouro Preto DÓeste (blue/diamond) and MLP (red /asterisc).

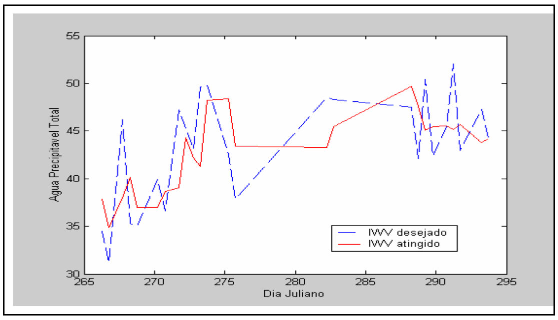

Figure 6 – Results from the estimated IWV from profiles. Radiosonde in Guajará Mirim (blue/diamond) and MLP (red /asterisc).

The HSB Brightness Temperatures were only measured at timetables of 0600 UTC and 18 UTC. These data were acquired on the same days of satellite coverage of this site. Straight lines correspond to non-available data, that is, days with no observations.

Figure 7 presents graphic line of the estimated IWV in Guajará Mirim with the brightness temperatures used by the regression model and the ANN method versus radiosondes. One may notice that the results of the ANN are closer to the radiosondes than the results of the linear regression method.

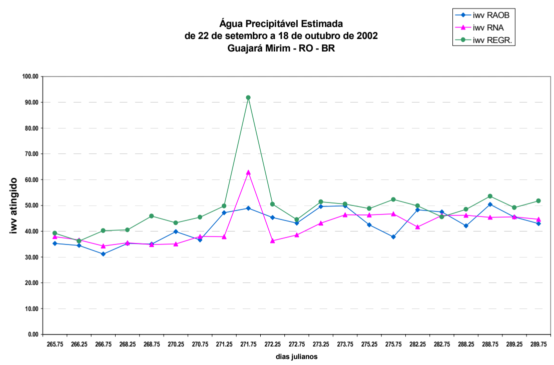

Figure 7 – The IWV measured by radiossondes in Guajará Mirim (blue/diamond), and the estimation provided by the MLP (red/triangle) and the linear regression method (green/dot). The Tb data are measured at timetables 06 ant 18 UTC.

Table 3 – Statistical parameters of the IWV comparison

TOTAL IWV					
			Statistics		
Sites	**Mean Radio.** (Kg/m^2)	**Mean RNA** (Kg/m^2)	**Bias** (kg/m^2)	**RMS**	**RMS** (Kg/m^2)
GM (Tr.)	43,57	43,55	0,02	0.06	2,95
PV	49,96	49,89	0,07	0.06	3,35
RJ	44,75	44,62	0,13	0.07	3,06
OP	42,73	42,69	0,04	0.08	3,51
GM HSB	42,33	41,91	0,42	0.13	5,57
GM Rgr	42,33	49,00(**)	-6,67	0.22	10,89
GM (*)	39,59	-	5,14	-	7,00

(*) Results from Lima, 2004. (**) Mean of IWVRgr.

The statistical results described in Table 3 present: the mean value of the radiosondes IWV values and output of the ANN. The *rms error* and the bias were in compared to the radiosondes and the ANN. It is to be noticed that the biggest error is 13% in case of Guajará Mirim (with the least number of observations). In comparison to the linear regression method, the rms (kg/m^2) was lesser than the one presented in Lima (2004). The statistics of the regression method were redone for comparison with the ANN. Considering the bias of less than 1, it shows that the results are not close to the ANN results.

5. CONCLUSIONS

This paper presents results of the analysis of using the HSB sensor microwave channels to estimate IWV in the atmosphere and the technique of Artificial Neural Network using real data from the RaCCI/LBA field experiment. The sensor AMSU-B channels of NOAA-16 satellite have the same characteristics of the HSB and the Tb measured data by these channels will be used in future work.

Based on the obtained results, one may conclude that the ANN method may satisfactorily estimate the IWV to the variability and distribution of the steam of water in the atmosphere. The estimation of the IWV allows the direct connection of the brightness temperature to the quantity of steam of water (for a determined vertical profile of temperature).

The observations of the radiosonde offer quite limited space coverage, especially in South America, and then it is of the great importance to develop a method to estimate vapor in the atmosphere, from satellite data. These estimated data will improve the limitations of meteorological observations of conventional stations (Cintra, 2004).

The use of the neural networks for the satellite data processing constitutes a useful tool with great computational simplicity.

Acknowlegdement

The authors thank MsC. Wagner Flauber Araújo for providing the data and the permission for comparison the results of the ANN with the method of linear regression implemented by him.

6. REFERENCES

Cintra, R. S. **Implementação do Sistema de Assimilação de Dados em Espaço Físico para o Modelo Global do CPTEC**. Dissertação de Mestrado em Computação Aplicada. Instituto Nacional de Pesquisas Espaciais, 2004. No prelo.

Grody, N.; Zhao, J.; Ferraro, R.; Weng, F.; Boers, R. Determination of precipitable water and cloud liquid water over oceans from the NOAA 15 advanced microwave sounding unit. ***Journal of Geophysical Research***, vol.106, no. D3, p. 2943-2953, February 16, 2001.

Haykin, S. ***Redes neurais: princípios e práticas***. 2 ed. Porto Alegre: Bookmann, 2001

Kakar, R. K.; Lambrigtsen, B, H. A statistical correlation method for the retrieval of atmospheric moisture profiles by microwae radiometry. ***J. Climate Appl. Meteor***., 23, 1110-1114, 1984.

Lambrigsen,B. H.; Calheiros, R. V. The Humidity Sounder for Brazil – An international partnership. ***IEEE Trans. Geosc. Remote Sensing***, n. 41, p. 352-361, 2003.

Lima, W. F. A.; Machado, L. A T. Análise do Sensor HSB na estimativa do conteúdo integrado de vapor D'água durante o experimento RaCCI/LBA. ***Revista Brasileira de Meteorologia*****.** No prelo.

Lima, W. F. A. *Um estudo sobre o uso do sensor HSB na estimativa da água precipitável e da precipitação*. São José dos Campos. Dissertação de Mestrado em Meteorologia. Instituto Nacional de Pesquisas Espaciais, 2004. No prelo.

Matricardi, M., Chevalier, F. and Tjemkes,S.: An Improved general fast radiative transfer model for the assimilation of radiance observations. ***ECMWF Research Dept. Tech. Memo. 345***. (available from the librarian at ECMWF), 2001.

National Aeronautics and Space Administration. NASA. Goddard Space Flight Center. ***Aqua-Project Science***. *http://aqua.nasa.gov/about/*

Randel, D. L.; Haar, V. T. H.; Ringerud, M.A. Observed variability of total column water vapor related to atmospheric temperature. In 9th Conference on Satellite Meteorology and Oceanography, UNESCO, Paris, France, 25-29 May, 1998. Pre-Prints. ***American Meteorological Society***, 1998 v.1, p. 15-17.

Teixeira, R. F. B. Índice de Água Precipitável da Atmosfera a partir dos canais 4 e 5 do AVHRR-NOAA. **Anais...** XIII Congresso Brasileiro de Meteorologia. Fortaleza, 2004.

Vianello, R. L.; ALVES, A.R. ***Meteorologia Básica e Aplicações***. Universidade Federal de Viçosa. Imprensa Universitária. 429p. 1991.

Wang, J. R.; CHANG, L. A. Retrieval of water vapor profiles from microwave radiometric measurement near 90 and 183 GHz. **J. Appl. Meteor**. 29 p. 1005-1013, Oct. 1990.