Atmospheric Temperature Retrieval from Satellite Data: New Non-extensive Artificial Neural Network Approach

Elcio Hideiti Shiguemori[1,2]
[1]Institute for Advanced Studies (IEAv)
Brazilian General Command for Aerospace Technology
São José dos Campos, Brazil
+55 (0)12 3947-5356
elcio@ieav.cta.br

Haroldo Fraga de Campos Velho
[2] National Institute for Space Research
(INPE)
São José dos Campos, Brazil
+55 (0)12 3945-6547
haroldo@lac.inpe.br

José Demisio Simões da Silva
[2] National Institute for Space Research
(INPE)
São José dos Campos, Brazil
+55 (0)12 3945-6543
demisio@lac.inpe.br

ABSTRACT

In this paper, vertical temperature profiles are inferred by neural networks based inverse procedure from satellite data, non-linear function estimation. A new approach to classical Radial Basis Function neural network is trained using data provided by the direct model characterized by the Radiative Transfer Equation (RTE). The neural network results are compared to the ones obtained from classical neural networks Radial Basis Function and traditional method to solve inverse problems, the regularization. In addition, real radiation data from the HIRS/2 - High Resolution Infrared Radiation Sounder - is used as input for the neural networks to generate temperature profiles that are compared to measured temperature profiles from radiosonde. Analysis of the new approach results reveals the generated profiles closely approximate the results obtained with classical neural networks and regularized inversions, [5] [15], thus showing adequacy of neural network based models in solving the inverse problem of temperature retrieval from satellite data. The advantages of using neural network based systems are related to their intrinsic features of parallelism; after trained, the networks are much faster than regularized approaches, and hardware implementation possibilities that may imply in very fast processing systems.

Keywords

Inverse problems, neural networks, temperature retrieval

1. INTRODUCTION

The vertical structure of temperature and water vapor plays an important role in the meteorological process of the atmosphere. However due to logistics and economic problems, there is a lack of observation in several regions of the Earth. In this sense, the retrieval of temperature and humidity profiles from satellite radiance data became important for applications such as weather analyzes and data assimilation in numerical weather predictions models. Interpretation of satellite radiance in terms of meteorological parameters requires the inversion of the Radiative Transfer Equation (RTE) where measurements of radiation performed in different frequencies are related to the energy from different atmospheric regions. The degree of indetermination is associated with the spectral resolution and the number of spectral channels. Moreover, usually this solution is very unstable due to the presence of noise in the measuring process [17, 24]. Several methodologies and models have been developed to improve the satellite data processing. Due to the difficulty of obtaining correct RTE solutions, several approaches and methods were developed to extract information from satellite data [6, 21, 14].

In order to deal with the ill-posed characteristic of inverse problems, regularized solutions [22, 2, 4], and also regularized iterative solutions [1, 12, 8] have been proposed. Also, several research works have tried artificial neural networks to solve inverse problems [11, 2, 13, 25]. The atmospheric temperature estimation is a classical inverse problem that been approached by ANNs, employing an Multi Layer Perceptron [20] [18] and a Radial Basis Function [19].

In this paper a new approach for a Radial Basis Function ANN is used to retrieve vertical temperature profiles based on remote sensed data. The temperature retrievals of the new technique are compared to the ones obtained by [5, 15], who used Tikhonov and maximum entropy principle regularization techniques.

2. DIRECT MODEL

The radiative transfer process can be modeled using the linear integral-differential Boltzmann equation, considering absorption, scattering, and radiative emission [5]. However, depending on the range of satellite observation - infrared, in our case - the RTE can be simplified. The Schwarzchild's equation is an RTE version where the scattering phenomenon can be neglected, having a local thermodynamic equilibrium. This means that the atmosphere behaves like a black body, following the Planck's law, relating the radiances with the body temperature. This scenario represents our direct or forward problem. Equation (1) represents the mathematical formulation of the direct problem that permits the calcula-

SAC'08 March 16-20, 2008, Fortaleza, Ceará, Brazil

tion of radiance values from associated temperatures [14].

$$I_\lambda(0) = B_\lambda(T_s)\Im_\lambda(p_s) + \int_{p_s}^{0} B_\lambda[T(p)]\frac{\partial \Im_\lambda(p)}{\partial p}dp, \quad (1)$$

where, I_λ is the value of the spectral radiance, subscript s denotes surface; p is the considered pressure; $\Im$ is the space atmospheric layer transmittance function that is a function of the wavelength and the concentration of absorbent gas, which usually declines exponentially with the height. In pressure coordinate, the transmittance may be expressed by:

$$\Im_\lambda(p) = \exp\left[-\frac{1}{g}\int_{p_0}^{p} k_\nu(p)q(p)dp\right] \quad (2)$$

where, k_ν is the absorption coefficient, q is the ratio of gas mixture, g is the acceleration of the local gravity, and p_0 is the pressure in the top of atmosphere; B is Planck's function (Eq. 3), which is a function of the temperature T and wavelength λ:

$$B_\lambda(T) = \frac{2hc^2/\lambda^5}{[e^{hc/k_B\lambda T} - 1]} \quad (3)$$

h is Planck's constant; c is the light speed; and k_B is the Boltzmann's constant.

The pressure is used as a vertical coordinate, due to the linear relationship between these two quantities (hydrostatic equation): $p = -g\rho z$; being g the gravity, ρ the air density, and z the height above the ground.

Equation (1) is discretized, for practical purposes, using central finite differences leading to Equation 4:

$$I_i = B_{i,s}(T_s)\Im_{i,s} + \sum_{j=1}^{N_p}\left(\frac{B_{i,j} + B_{i,j-1}}{2}\right)[\Im_{i,j} - \Im_{i,j-1}] \quad (4)$$

where $i = 1, ..., N_\lambda$; $I_i \equiv I_{\lambda_i}$; N_λ is the number of satellite channels; $B_{i,j} = B_{\lambda_i}(T_j)$; $\Im_{i,j} = \Im_{\lambda_i}(p_j)$; and N_p is the number of atmospheric layers considered. It is assumed that each atmospheric layer has a characteristic temperature T_j to be computed.

3. NEURAL NETWORK ARCHITECTURE

RBF networks are feed-forward networks with one hidden layer, initially developed for function approximation in multidimensional space, but they may also be used to learn arbitrary input-output mappings [10] [3]. Each unit in the hidden layer has an associated receptive field centered at a particular input value at which it reaches a maximal output. The output of a unit tails off as the input moves away from its central point. The mostly used function in RBF networks is a Gaussian, though other types of function may be used [10]. The weights from the input layer to the hidden layer (the centers of the Gaussians) represent vector prototypes within the available data that may be determined by clustering methods or by random selection in data set. The weights between the hidden layer and the output layer are trained by using the delta rule [10], by presenting the available input-output pairs to the RBF through a number of iterations pursuing a specific performance index given by the minimization of the sum of square error over the entire training set. The classical RBF activation is obtained:

$$y = \sum_{i=1}^{N} w_i \phi\left(\|\bar{x} - \bar{t_i}\|\right) \quad (5)$$

where $\bar{x}$ is the input vector, $\bar{t_i}$ is the center of ith function and ϕ is the radial basis function.

This paper presents new results on the construction of a reformulated radial basis function neural network. A new family of radial basis functions based on non-extensive entropy. Are proposed. The non-extensive form proposed by Tsallis [23] is obtained by:

$$S_q(p(r)) = \frac{k}{q-1}\left(1 - \sum_{i=1}^{W} p(r)_i^q\right) \quad (6)$$

where p_i is a probability, W the total number of possibilities and q is a free parameter. In the thermodynamic theory, the k parameter is known as Boltzmann constant. In the limit, $q \to 1$, Tsallis entropy reduces to the usual form of Boltzmann-Gibbs-Shannon [9].

As presented in Tsallis [23], the distributions $p_q(r)$ to different values of q are obtained by:
if $q > 1$
$p_q(r) = \phi(r) =$

$$= \frac{1}{\sigma}\left[\frac{q-1}{\pi(3-q)}\right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{3-q}{2(q-1)}\right)} \frac{1}{\left(1 + \frac{q-1}{3-q}\frac{r^2}{\sigma^2}\right)^{\frac{1}{q-1}}}, \quad (7)$$

if $q = 1$

$$p_q(r) = \phi(r) = \frac{1}{\sigma}\left[\frac{1}{2\pi}\right]^{\frac{1}{2}} e^{-(r/\sigma)^2/2}, \quad (8)$$

if $q < 1$
$p_q(r) = \phi(r) =$

$$= \frac{1}{\sigma}\left[\frac{1-q}{\pi(3-q)}\right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{5-3q}{2(1-q)}\right)}{\Gamma\left(\frac{2-q}{1-q}\right)}\left[1 - \frac{(1-q)}{(3-q)}\frac{r^2}{\sigma^2}\right]^{\frac{1}{(1-q)}} \quad (9)$$

If $|r| < \sigma[(3-q)/(1-q)]^{1/2}$ and zero otherwise.

The family of solutions is presented in Figure 1. For $q = 1$ it is equivalent to a Gaussian function and for $q = 2$ it is equivalent to Cauchy distribution.

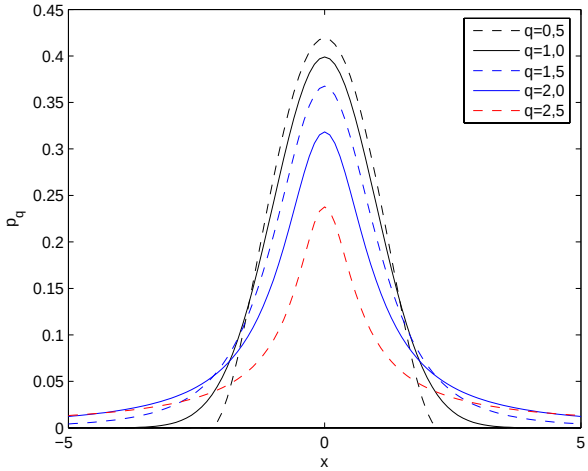

Figure 1: Distributions for different values of q

4. NEURAL NETWORK FOR ATMOSPHERIC PROFILE RETRIEVAL

Artificial neural networks have two stages in their application, the learning and activation steps. During the learning step, the weights and bias corresponding to each connection are adjusted to some reference examples. For activation, the output is obtained based on the weights and bias computed in the learning phase.

The experimental data, which intrinsically contains errors in the real world, is simulated by adding a random perturbation to the exact solution of the direct problem, such that

$$\tilde{I} = I_{\text{exact}} + I_{\text{exact}}\sigma\mu \tag{10}$$

where σ is the standard deviation of the noise and μ is a random variable taken from a Gaussian distribution, with zero mean and unitary variance. All numerical experiments results presented in this paper were carried out using σ=0.05.

The error of temperature profiles is computed in 5 layers, with Layer-1 [1000 hPa, 500 hpa]; Layer-2 [500 hPa, 250 hpa]; Layer-3 [250 hPa, 85 hpa]; Layer-4 [85 hPa, 20 hpa]. Some discrete points (pressure) are considered for each layer: **Layer-1**: 500, 570, 620, 670, 700, 780, 850, 920, 950, 1000 hPa; **Layer-2**: 250, 300, 350, 400, 430, 475 hPa; **Layer-3**: 85, 100, 115, 135, 150, 200 hPa; **Layer-4**: 20, 25, 30, 50, 60, 70 hPa.

This feature is important because the main interest for meteorological purposes are the layers below $p = 100$ hPa, where 1 hPa = 100 Pa.

The average errors of simulation results for each atmospheric layer obtained with the ANN are computed by:

$$\text{Error} = \frac{1}{N}\sqrt{\sum_{p_b}^{p_t}\left(T_i^{\text{Radiosonde}} - T_i^{\text{NeuralNetwork}}\right)^2} \tag{11}$$

where N is the number of sample points (sub-layers) at each layer, p_b and p_t are, respectively, pressure (level) at bottom and top for each layer.

4.1 Generalization

The activation test is an important procedure to indicate the performance of an ANN, the effective test is conducted using data that does not belong to the training set. This action is called the generalization test of the ANN. In this work, 3 different databases are employed in the ANN tests, the first with 101 synthetic profiles derived from some known atmospheric profiles in Brazil; the second using 324 profiles of TIGR data set; and the third using the union of the synthetic and the TIGR profiles. The generalization test errors presented in the Table 1 were obtained with the synthetic data, the Table 2 the TIGR database and Table 3 with the synthetic + TIGR database. From the Tsallis' formulation for the non-extensive thermodynamics, there is no theory to calculate the free parameter q. For some applications, it is possible to derive a criterion to compute the non-extensivity parameter [16]. For the present application, there is no theory for estimating such parameter. Therefore, numerical experimentation is the procedure that we have adopted to find out the best choise of parameter q. The Tables present the RMSE of the profiles obtained by the proposed Radial Basis Function Non-Extensive (RBF-NE) with $q = 0.5$, 1.0, 1.5, 2.0 and 2.5.

Table 1: Generalization with RFBR-NE - sintetic database

q	Layer_1	Layer_2	Layer_3	Layer_4	Layer_5
0,5	0,9080	0,9702	1,4511	**1,1228**	0,8088
1,0	**0,6867**	0,7992	1,0296	1,2185	0,7191
1,5	0,7728	**0,6911**	**0,9693**	1,3113	0,6978
2,0	0,7649	0,7269	1,0364	1,3539	0,6386
2,5	0,7603	0,7398	1,0873	1,3424	**0,6376**

Table 2: Generalization with RFBR-NE - TIGR database

q	Layer_1	Layer_2	Layer_3	Layer_4	Layer_5
0,5	1,1134	1,3122	1,9393	1,6463	1,9420
1,0	**1,1091**	1,3114	1,9560	1,5958	1,9553
1,5	1,1178	1,3121	1,9801	1,5814	1,9368
2,0	1,1158	1,3004	1,8942	1,5743	1,9314
2,5	1,1135	**1,2956**	**1,8940**	**1,5474**	**1,8928**

Table 3: Generalization with RFBR-NE - Synthetic + TIGR database

q	Layer_1	Layer_2	Layer_3	Layer_4	Layer_5
0,5	0,7023	1,5797	2,8594	2,0310	1,3913
1,0	**0,4736**	1,2221	**2,4897**	1,5770	1,1683
1,5	0,5029	1,1977	2,6024	**1,3525**	1,1594
2,0	0,5323	1,1907	2,6065	1,3822	1,1424
2,5	0,5628	**1,1810**	2,6619	1,4636	**1,1271**

As shown in the Tables 1, 2 and 3, different values of free parameter q can improve the results comparing with the classical RBF with Gaussian radial basis function, i.e. $q = 1.0$. To exemplify this improvement, Figure 2 shows the desired radiosonde profile, in black line, classical RBF with $q = 1$ in blue line and RBF-NE with $q = 2.5$ the red line.

4.2 Estimation Using Real Satellite Radiance Data

Simulations using real satellite radiance data, from the High Resolution Radiation Sounder (HIRS-2) of NOAA-14 satellite, have been performed to evaluate the accuracy of the RBF-NE. HIRS-2 is one of the three sounding instruments of the TIROS Operational Vertical Sounder (TOVS). ANN results are compared to in situ radiosonde measurements and results obtained by [5, 15], who used Tikhonov and Maximum Entropy Principle of second order regularization techniques.

The number of observations corresponds to a fraction of the number of temperatures to be estimated. For instance, in the example presented hereafter, 40 temperature values are estimated from 7 radiance measurements.

Figures 3, 4, 5, show the results of the generalization tests performed. It is possible to notice a reasonable agreement between RBF-NE's retrieval and the radiosonde measurements. Figure 3 shows the classical RBF with Gaussian radial basis function compared with radiosonde, Tikhonov and Maximum Entropy Principle of second order regularization techniques [5, 15]. Figures 4 and 5 show, respectively, the results of RFB-NE with $q = 0.5$ and $q = 1.5$.

The mathematical formulation of the problem of retrieving vertical temperature profiles from remote sensing data is given by the integral radiative transfer equation, and leads to the solution of a highly ill-conditioned Fredholm integral equation of the first kind. The results show the good agree-

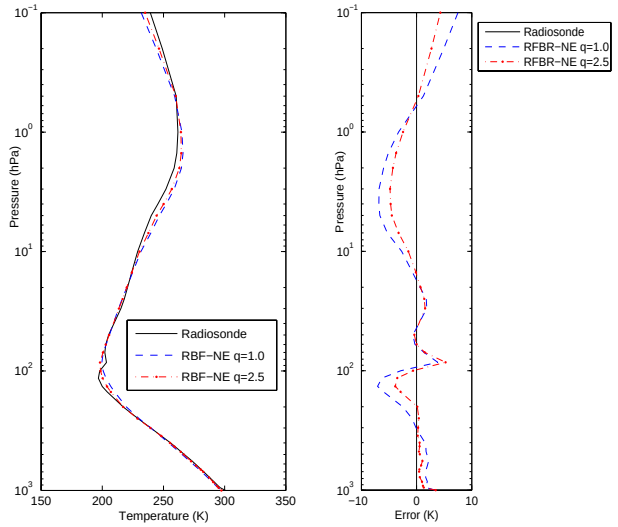

Figure 2: Generalization example obtained with RBF-NE and $q = 2.5$

ment between estimation from the ANNs and the observed temperature profile. It should be noted that the RBF-NE trained with other values of q presented the better performance in different layers of the atmosphere. This is an evidence of that different values of q in the RBF-NE neural network can improve the atmospheric temperature inversion.

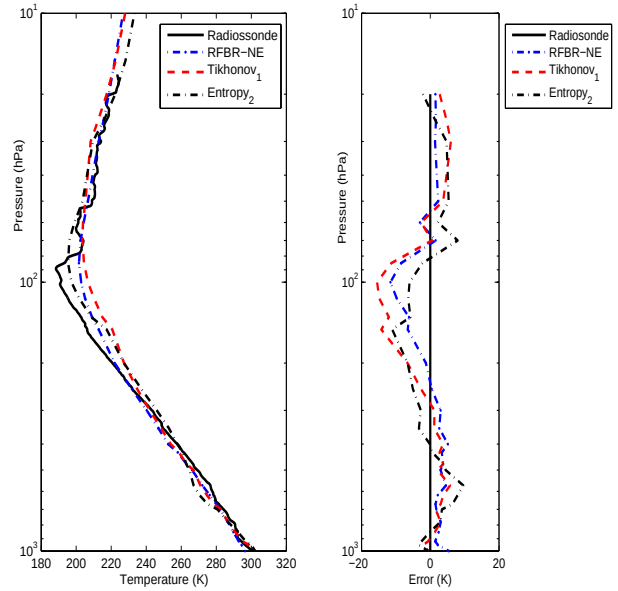

Figure 3: Retrievals achieved using radiance data from NOAA-14 satellite - RBF-NE with $q = 1.0$

5. CONCLUSION

In the present work ANNs are presented as effective tools for solving the inverse problem of retrieving atmospheric temperature profiles. The obtained reconstructions with the RBF-NE showed to be better than the ones obtained with

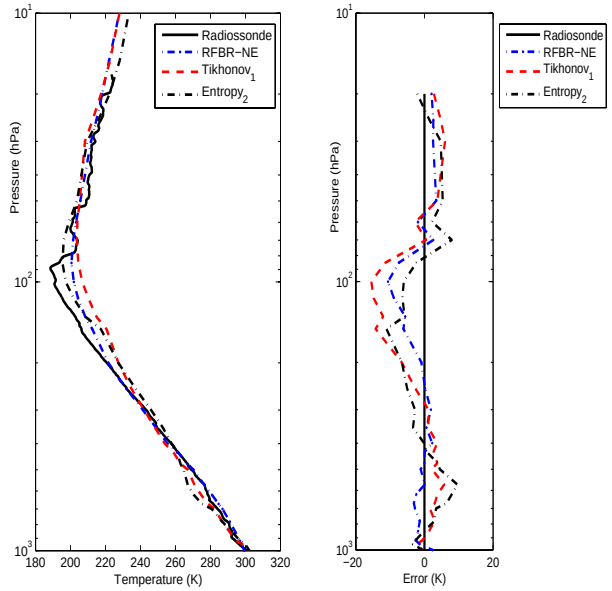

Figure 4: Retrievals achieved using radiance data from NOAA-14 satellite - RBF-NE with $q = 0.5$

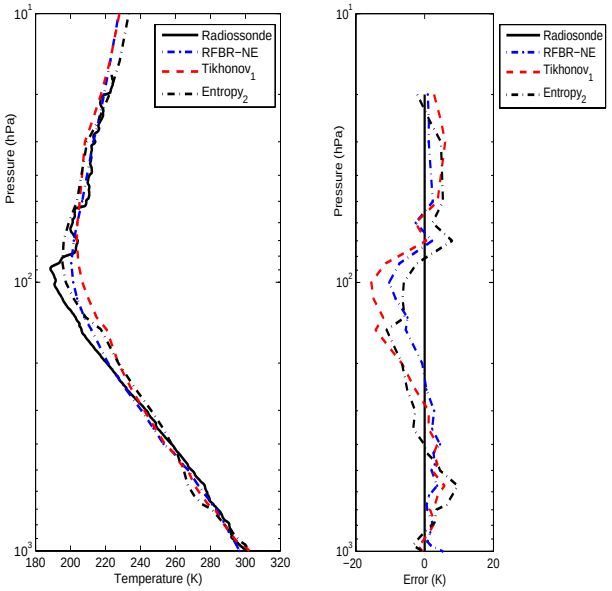

Figure 5: Retrievals achieved using radiance data from NOAA-14 satellite - RBF-NE with $q = 1.5$

classical RBF and regularization methods [5, 15], even for noisy data. It is to be noted that the most important layer for numerical weather prediction lies between the layers 1000 to 100 hPa. The results suggest that we need to combine different techniques, considering different atmospheric layers, in order to have a better inverse solution.

In practice, operational inversion algorithms reduce the risk of being trapped in local minima by starting the iterative search process from an initial guess solution that is sufficiently close to the true profile. However, the dependence of the final solution on a good choice of the initial guess represents a fundamental weakness of such algorithms, partic-

ularly in regions where less a priori information is available [7]. ANN approaches can relax this constraint incorporating more data in the dataset during the learning phase.

The ANNs can be inaccurate if they are used to extrapolate to cases outside the training domain. However the use of ANN techniques can provide good solutions when the training phase encompasses the domain of the potential solutions to the real problem. The TIGR database contains 2300 representative atmospheric situations collected around the world and selected by statistical methods from 80,000 radiosonde reports. Therefore, our training sets were chosen from professional criteria, minimizing the chance for producing unrealistic atmospheric temperature profile estimation. It can be seen – from generalization and real cases results worked in this paper – good inversions obtained. One problem related to the present formulation for this new RBF-NN is just because there is no theory to drive us to compute the non-extensive parameter q. It is important to mention that the present approach is under verification for application as an operational issue for the Center of the Weather Prediction and Climate Studies (CPTEC), a division of the National Institute for Space Research (INPE).

6. ACKNOWLEDGMENTS

The authors are gratefully indebted to CNPq and FAPESP.

7. REFERENCES

[1] O. Alifanov. Solution of an inverse problem of heat conduction by iteration methods. *Journal of Engineering Physics*, 26(11):471–476, 1974.

[2] M. Atalla and D. Inman. On model updating using neural networks. *Mechanical Systems and Signal Processing*, 12(1):135–161, 1998.

[3] D. S. Broomhead and D. Lowe. Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2:321–355, 1988.

[4] H. F. Campos Velho and F. M. Ramos. Numerical inversion of two-dimensional geoelectric conductivity distributions from eletromagnetic ground data. *Braz. J. Geophys.*, 15(2):133–144, 1997.

[5] J. Carvalho, F. Ramos, N. Ferreira, and H. de Campos Velho. Retrieval of vertical temperature profiles in the atmosphere. In *3rd International Conference on Inverse Problems in Engineering*, pages 235–238, 2000.

[6] M. T. Chahine. Determination of the temperature profile in an atmosphere from its outgoing radiance. *Journal of the Optical Society of America*, pages 1634–1637, 1968.

[7] A. Chedin, C. W. N.A. Scott, and P. Moulinier. The improved initialization inversion method : a high resolution physical method for temperature retrievals from satellites of the tiros-n series. *J. Climate Appl. Meteor.*, 24:128–143, 1985.

[8] L. Chiwiacowsky and H. CAMPOS VELHO. Different approaches for the solution of a backward heat conduction problem. *Inverse Problems in Engineering*, 11(6):417–494, 2006.

[9] H. de Campos Velho, F. Ramos, E. Shiguemori, and J. Carvalho. A unified regularization theory: The maximum non-extensive entropy principle. *Computational and Applied Mathematics*, 25:307–330, 2007.

[10] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Mcmillan, New York, 1994.

[11] H. Hidalgo and E. Gómez-Treviño. Application of constructive learning algorithms to the inverse problem. *IEEE T. Geosci. Remote*, 34(1):874–885, 1996.

[12] Y. Jarny, M. Özisik, and J. Bardon. A general optimization method using adjoint equation for solving multidimensional inverse heat conduction. *International Journal of Heat and Mass Transfer*, 34(11):2911–2919, 1991.

[13] J. Krejsa, K. Woodbury, J. Ratliff, and M. Raudensky. Assessment of strategies and potential for neural networks in the inverse heat conduction problem. *Inverse Problems in Engineering*, 7(3):197–213, 1999.

[14] K. N. Liou. *An introduction to atmospheric radiation*. Academic Press, Orlando, 1982.

[15] F. Ramos, H. Campos Velho, J. Carvalho, and N. Ferreira. Novel approaches on entropic regularization. *Inverse Problems*, 15(5):1139–1148, 1999.

[16] F. Ramos, R. Rosa, C. R. Neto, M. Bolzan, L. Sa, and H. de Campos Velho. Nonextensive statistics and three-dimensional fully developed turbulence. *Physica A*, 295(1/2):250–253, 2001.

[17] C. D. Rodgers. Retrieval of the atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.*, 14(4):609–624, 1976.

[18] E. H. Shiguemori, H. F. Campos Velho, J. S. Silva, and J. C. Carvalho. Neural network based models in the inversion of temperature vertical profiles from satellite data. *Inverse Problems in Engineering*, 14(5):543–556, 2006.

[19] E. H. Shiguemori, J. S. Silva, H. F. Campos Velho, and J. C. Carvalho. Radial basis function neural network in the inversion of temperature vertical profiles from satellite radiance data. In *Brazilian Symposium on Neural Networks*, 2004.

[20] E. H. Shiguemori, J. S. Silva, H. F. Campos Velho, and J. C. Carvalho. A multilayer perceptron approach for the retrieval of vertical temperature profiles from satellite radiation data. In *Proceedings of International Joint Conference on Neural Networks*, pages 2689–2693, 2005.

[21] W. Smith, H.M.Woolf, and W. Jacob. A regression method for obtaining realtime temperature and geopotential height profiles from satellite spectrometer measurements and its application to nimbus 3 sirs observations. *Monthly Weather Review*, 98(2):582–603, 1970.

[22] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. Winston & Sons, New York, 1977.

[23] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal Statistical Physics*, 52:479, 1988.

[24] S. Twomey. *Introduction to the mathematics of inversion in remote sensing and indirect measurements*. Elsevier Scientific, Amsterdam, 1977.

[25] K. Woodbury. Neural networks and genetic algorithms in the solution of inverse problems. *Bulletim of the Braz. Soc. for Comp. Appl. Math. (SBMAC)*, 2000.