Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26 abril 2007, INPE, p. 6437-6442.

Neural network based models for the retrieval of methane concentration vertical profiles from remote sensing data

Adenilson R. Carvalho[1]
João C. Carvalho[2]
Elcio H. Shiguemori[1]
José Demisio S. Da Silva[1]
Fernando M. Ramos[1]

[1] Brazilian National Institute for Space Research - INPE
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil
{adenilson, elcio, demisio, fernando}@lac.inpe.br,

[2] National Agency of Telecommunications – ANATEL
Caixa Postal 440 – 70.070-940 - Brasília - DF, Brasil
jcarvalho@anatel.gov.br

Abstract. Trace gas profile retrieval constitutes an ill-posed inverse problem. This paper explores the potential of the application of neural networks for the retrieval of methane vertical column densities from remote sensed radiances. Three types of neural networks have been tested using noiseless and noisy data. Here we focused on the reconstruction of vertical column densities using SCIAMACHY channel 8 nadir measurements (~ 2.3 mm). Overall, the use of neural networks was able to solve this difficult inverse problem even in the presence of noise in the data. A comparison among different network architectures was accomplished but it was not possible to detect great discrepancies in the performance of them.

Keywords: neural networks, greenhouse gas, inverse problem.

1. Introduction

Methane (CH_4) and carbon dioxide (CO_2) are the most important anthropogenic greenhouse gases. Recent studies have shown that atmospheric methane concentrations have increased by approximately 150% since pre-industrial times. This corresponds to a radiative forcing of $0.48Wm^{-2}$, which amounts to 20% of the total radiative forcing due to well-mixed greenhouse gases. More than half of the present-day methane emissions are of anthropogenic origin and the most important sources are fossil-fuel production, domestic ruminants, rice cultivation and waste handling. CH_4 absorbs the rising radiation from the earth-atmosphere system on the near infrared spectral range and plays an important role in the greenhouse effect and in the climatic change over the globe according to IPCC (2001).

Within the context above, it is becoming increasingly relevant to develop novel and better techniques to estimate concentration profiles of atmospheric constituents like methane. In this sense, a good alternative is the use of satellite remote sensing data. CH_4 retrieval feasibility and sensitivity studies have already been performed for the Atmospheric Infrared Sounder

1. Laboratory for Computing and Applied Mathematics – LAC/INPE -
Av. Astronautas, 1758 – CEP: 12227-010 - São José dos Campos, SP, Brazil
Fone: (12)3945.6934; e-mail: [adenilson, carlos , elcio, fernando, demisio]@lac.inpe.br

(AIRS), see Christi (2004) and Engelen (2004), and the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), see Buchwitz (2004) and Buchwitz (2005).

In this article we discuss a neurocomputing approach to deal with this kind of inverse problem. In a first step, we investigated the performance of the radiative transfer model (forward model) to be used in the inversion process. The forward model used here is the SCIATRAN code, developed by the University of Bremen (IFE/IUP) to simulate radiances from SCIAMACHY satellite spectrometer, as described by Rozanov (2002). Based on these studies we decided to exploit the Sciamachy data near-infrared channel 8 in order to do the CH_4 profile inversion as related above. In a second step, we developed a neural network for reconstructing methane vertical column densities from remote sensing data. To this end, we employed synthetic radiances simulated by SCIATRAN for the learning process of the neural network. Finally, we validated our approach using a set of more than a hundred test cases, with and without noise in the radiance data.

2. Forward model

The SCIATRAN, developed at University of Bremen (IFE/IUP), initially to simulate radiances from SCIAMACHY satellite spectrometer, was applied to generate the synthetic radiances used in this work. SCIATRAN, written in FORTRAN 95, has been developed to perform radiative transfer modeling in any observation geometry appropriate to measurements of the scattered solar radiation in the Earth's atmosphere, and has been used as a forward model in the retrieval of atmospheric constituents from measurements of scattered solar light by satellite, ground-based, or airborne instruments in UV–Vis–NIR spectral region.

SCIATRAN solves the radiative transfer equation using the Finite Difference Method for a plane-parallel vertically inhomogeneous atmosphere taking into account multiple scattering, a new version support additionally radiative transfer calculations in a spherical atmosphere. The wavelength range covered by the radiative transfer model is 175–2380 nm, including Schuman-Runge and Herzberg absorption bands of oxygen (see Rozanov (2002) for more details).

In our study, we used the SCIATRAN version 2.1; it was adjusted in such a way that all radiative transfer calculations were performed in a spherical atmosphere (i.e., all effects due to the sphericity of the Earth's atmosphere were considered). By default, a climatological data base obtained using a 2D chemo-dynamical model developed at MPI Mainz by Brühl and Crutzen (1993) is employed in SCIATRAN. This data base contains monthly and latitudinal dependent vertical distributions of atmospheric trace gas volume mixing ratios, pressure, and temperature between 0 and 60 km.

3. Neural network architecture

An artificial neural network (ANN) is an interconnected group of artificial neurons, elements of networks, that uses a mathematical or computational model for information processing based on a connectionist approach to computation. An input to a neuron consists of a number of values (x_1, x_2,..., x_n), while output is single value y. The neuron computes the weighted sum of its inputs, subtracts some threshold T, and passes the result to a non-linear function f (e.g., a sigmoid). In more practical terms, neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. In this paper, we used three types of neural networks: a feedforward backpropagation network, Elman backpropagation network, and a Radial basis Function network (RBF).

The first one is a multilayer perceptron (MLP) with backpropagation learning. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. MLP is especially useful for approximating a classification function that maps input vector (x_1, x_2, ..., x_n) to one or more classes (C_1, C_2, ,…, C_m). By optimizing weights and thresholds for all nodes, the network can represent a wide range of classification functions.

A simple recurrent network (SRN) is a variation on the multi-layer perceptron, sometimes called an "Elman network". A three-layer network is used, with the addition of a set of "context units" in the input layer. There are connections from the middle (hidden) layer to these context units fixed with a weight of one. At each time step, the input is propagated in a standard feed-forward fashion, and then a learning rule (usually back-propagation) is applied. The fixed back connections result in the context units always maintaining a copy of the previous values of the hidden units. Thus the network can maintain a sort of state, allowing it to perform such tasks as sequence-prediction that are beyond the power of a standard multi-layer perceptron.

The last type of neural network used in this work was the RBF network. Haykin (1994) and Bishop (1995) describe it as a type of neural network employing a hidden layer of radial units and an output layer of linear units, and characterized by a reasonably fast training and a reasonably compact architecture. The Radial Basis Function is embedded in a two layer neural network, where each hidden unit implements a radial activation function. The output units implement a weighted sum of hidden unit outputs. Their excellent approximation capabilities have been studied in Poggio & Girosi (1990). Due to their nonlinear approximation properties, RBF networks are able to model complex mappings, which perceptron neural networks can only model using multiple intermediary layers. A more detailed introduction on ANNs can be found in Haykin (1994) and Tsoukalas & Uhrig (1997).

Regardless their type or use, all neural networks have three stages in their application: the learning, the activation and the generalization steps. It is in the learning step that the weights and bias corresponding to each connection are adjusted to some reference examples (the input). In the activation phase, the output is obtained based on the weights and bias computed in the learning phase. The experimental data used here in the learning step were simulated adding a random perturbation to the exact solution for forward problem (SCIATRAN):

$$\widetilde{I} = I_{exact} - I_{exact}\sigma\mu \tag{1}$$

where σ is the noise standard deviation and μ is a random variable taken from a Gaussian distribution with zero mean and unitary variance. In all simulations we used $\sigma = 0.05$. Overall, more than one hundred pairs of concentration profiles and their corresponding radiances needed to inversion process, and that constitute what was called of radiance dataset or SBD. Similar data sets were utilized for the activation and generalization phases of the ANN.

4. Results

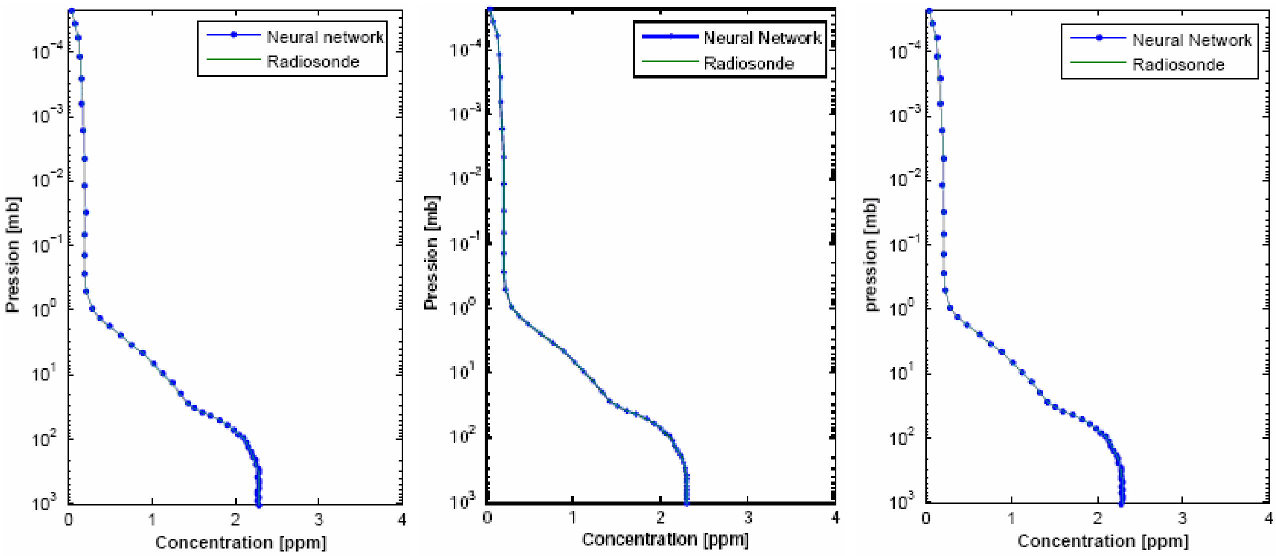
In order to analyze the performance of the ANNs in the retrieval of methane vertical column densities, two experiments were performed. In the first experiment noiseless data sets were used, and in the second one, 5% of white Gaussian noise ($\sigma = 0.05$) was added to the synthetic data, simulating the real experimental data. All ANNs were trained with only one hidden layer, varying the number of hidden neurons and the database training. The MLP neural network, the RBF and the Elman network were implemented using 20, 40 and 40 neurons, respectively. The training phase was carried out until a minimum value for the RMS error or a maximum number of iterations were reached. The results are available for five layers: Layer 1 [0.03 - 0.000027 mb]; Layer 2 [3.3 - 0.03 mb]; Layer 3 [43.7 - 3.3 mb]; Layer 4 [179 - 43.7 mb] and Layer 5 [1013 - 179 mb]. Some discrete points (pressure) are considered for each layer, 49 discrete points at all. It is important to notice that the first layer has more discrete points than other layers. It received more attention because the main interest for meteorological purposes concerning the trace gas retrieval issues are the layers below p =100 mb.

The mean errors of the simulation results for each atmospheric layer obtained with the neural network were calculated through:

$$Error = \frac{1}{N}\sqrt{\sum_{i=p_b}^{p_t}\left(C_i^{exact} - C_i^{neuralnetwork}\right)^2} \qquad (2)$$

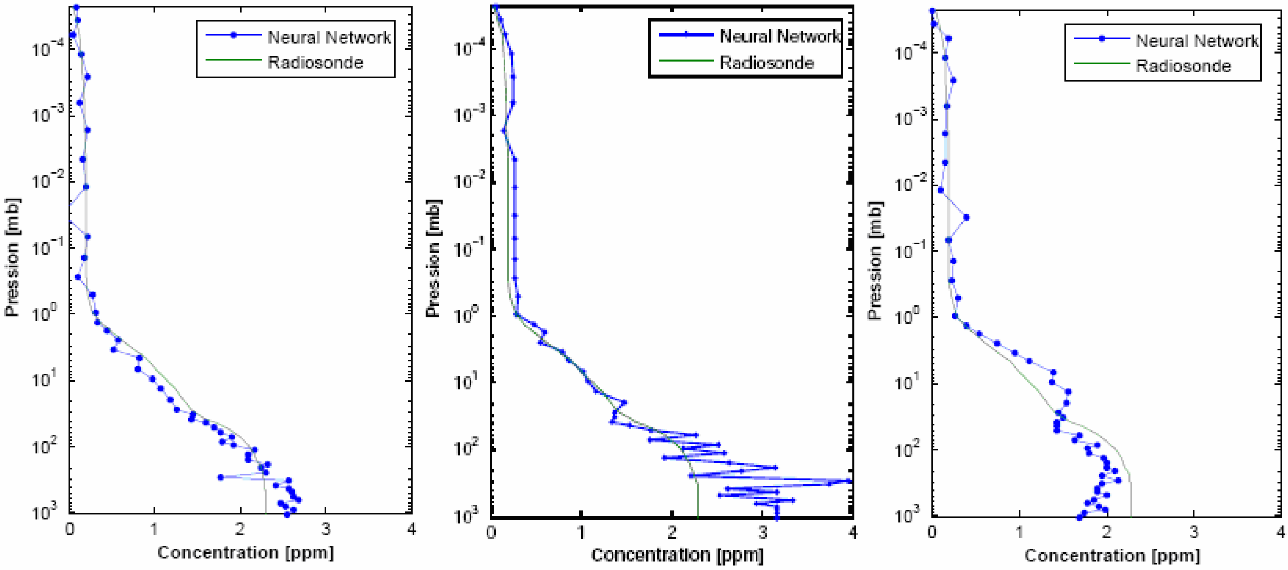
where N is the number of sample points (sub-layers) at each layer, p_b and p_t are, respectively, pressure (level) at bottom and top for each layer. It's also important observe that only the best training varying the number of hidden neurons for each one of ANN's are taken into account. **Figures 1** and **2** show the results of the generalization tests in comparison with the true model (here called radiosonde). **Table 1** summarizes the errors (in ppmv) for different neural network and each atmospheric layer considering a noiseless dataset. **Table 2** shows the same results as in **Table 1**, but for noisy data.

The results obtained with noiseless data are in excellent agreement with the true model. Reconstruction could be done up to a precision of 1% with reference to a synthetic database, for all layers. In this case, the RBF network presented the best global performance, although the MLP and the Elman networks also produced good profile reconstructions. Preliminary results considering noisy data show goods results for all layers, with RMS errors in order of 10% of reference data. The best results for noisy data were obtained for the RBF network, except for layer 5, where the Elman network gave better results. A despite of these discrepancies, we may say that all ANN analyzed here produced similar results.



(a) (b) (c)

Figura 1 - Generalization results of CH_4 vertical column densities retrieval (in ppm) for noiseless data using: (a) MLP Network, (b) RBF Network, and (c) Elman Network.



(a) (b) (c)

Figura 2 - Generalization results of CH_4 vertical column densities retrieval (in ppm) for noisy data using: (a) MLP Network, (b) RBF Network, and (c) Elman Network.

Table 1. Generalization noiseless results of CH_4 vertical column densities retrieval (in ppm).

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
MLP	2.01e-02	1.99e-02	5.39e-02	1.08e-02	4.60e-03
RBF	4.47e-05	6.11e-05	2.66e-04	5.90e-04	3.67e-04
Elman	1.54e-02	1.94e-02	3.98e-02	1.03e-02	4.40e-03

Table 2. Generalization noisy results of CH_4 vertical column densities retrieval (in ppm).

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
MLP	3.91e-02	5.78e-02	1.29e-01	2.65e-01	2.06e-01
RBF	2.11e-02	2.71e-02	3.82e-02	1.05e-01	2.43e-01
Elman	3.23e-02	4.93e-02	1.34e-01	2.02e-01	1.89e-01

5. Final remarks

In this paper, three architectures of neural networks were applied to the inverse problem of retrieval of methane vertical concentration profiles from remote sensing data. Here we focused on the reconstruction of vertical column densities using SCIAMACHY channel 8 nadir measurements (~ 2.3 mm).

Overall, the use of neural networks was able to solve this difficult inverse problem even when the data were contaminated with noise. A comparison among different ANN architectures was accomplished but it was not possible to detect great discrepancies in the performance of them. Some modifications could be implemented in order to improve the neural network performance. The use of noisy data in the training data is an obvious one. The incorporation of real satellite radiances is another. From this preliminary exercise, two advantages of the use of neural networks became clear: after the training phase, the reconstruction algorithm is much faster than the classical inversion methods; and it is an intrinsically parallel approach that can be very easily implemented in a parallel environment.

Acknowledgements

The first author thanks the finantial support given by CAPES-Brazil.

References

IPCC, **Climate change 2001: The scientific basis**, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., et al., Cambridge University Press, UK and New York, NY, USA, 2001.

Christi, M.J.; Stephens, G.L. Retrieving profiles of atmospheric CO2 in clear sky and in the presence of thin cloud using spectroscopy from the near and thermal infrared: A preliminary case study, **Journal of Geophysical Research-Atmospheres**, 109 (D4), D04316, 2004.

Engelen, R.J.; Stephens, G.L. Information content of infrared satellite sounding measurements with respect to CO2, **Journal of Applied Meteorology**, 43 (2), 373-378, 2004.

Buchwitz, M.; Noel, S.; Bramstedt, K.; Rozanov, V.V.; Eisinger, M.; Bovensmann, H.; Tsvetkova, S. ; Burrows, J.P. Retrieval of trace gas vertical columns from SCIAMACHY/ENVISAT near-infrared nadir spectra: first preliminary results, **Advances in Space Research**, 34, 809-814, 2004.

Buchwitz, M.; de Beek, R.; Burrows, J. P.; Bovensmann, H.; Warneke, T.; Notholt, J.; Meirink, J. F.; Goede, A. P. H.; Bergamaschi, P.; Körner, S.; Heimann, M.; Schulz, A. Atmospheric methane and carbon dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models, **Atmos. Chem. Phys**., 5, 941–962, 2005.

Rozanov, V.V.; Buchwitz, M.; Eichmann, K.U.; de Beek, R. and Burrows, J. P. SCIATRAN – a new radiative transfer model for geophysical applications in the 240– 2400 nm spectral region: The pseudo-spherical version. **Advances in Space Research**, 29, 11 (2002),1831–1835.

Brühl, C. and Crutzen, P. J. **MPIC two-dimensional model. In The atmospheric effects of stratospheric aircraft**, vol. 1292 of NASA Ref. Publ. 1993, pp. 103–104.

Haykin, S. **Neural Networks. A Comprehensive Foundation**. Macmillan, New York, NY, 1994.

Bishop, C. M. **Neural Networks for Pattern Recognition**. Oxford University Press, 1995.

Poggio, T.; Girosi, F. Networks for approximation and learning, **Proc. IEEE**, vol.78, no.9, pp.1481–1497, 1990.

Tsoukalas, L.H.; Uhrig, R.E. **Fuzzy and Neural Approaches in Engineering**. New York: John Wiley & Sons, 1997.