GLOBAL TEMPERATURE ASSIMILATION BY ARTIFICIAL NEURAL NETWORKS FOR AN ATMOSPHERIC GENERAL CIRCULATION MODEL

Rosângela S. Cintra and Haroldo F. de Campos Velho
Brazilian National Institute of Space Research - INPE
{rosangela.cintra,haroldo}@lac.inpe.br

Abstract – An Artificial Neural Network (ANN) is designed to investigate a application for data assimilation. This procedure provides an appropriated initial condition to the atmosphere to weather forecasting. Data assimilation is a method to insert observational information into a physical-mathematical model. The use of observations from the earth-orbiting satellites in operational numerical prediction models provides large data volumes and increases the computational effort. The goal here is to simulate the process for assimilating temperature data computed from satellite radiances. The numerical experiment is carried out with global model: the "Simplified Parameterizations, primitivE-Equation DYnamics"(SPEEDY). For the data assimilation scheme was applied an *Multilayer Perceptron*(MLP) with supervised training. The MLP-ANN is able to emulate the analysis from the *Local Ensemble Transform Kalman Filter*(LETKF). The ANN was trained with first three months for years 1982, 1983, and 1984 from LETKF. A hindcasting experiment for data assimilation cycle was for January 1985, with a MLP-NN performed with the SPEEDY model. The results for analysis with ANN are very close with the results obtained from LETKF. The simulations show that the major advantage of using MLP-NN is the better computational performance, with similar quality of analysis.

Keywords – Artificial neural networks, multilayer perceptron, data assimilation, numerical weather forecasting.

1. INTRODUCTION

The procedure which takes atmospheric observed data and creates meteorological fields over some spatial or temporal domain is usually called *analysis* or *data assimilation*. The data are distributed in time and the procedure uses an explicit dynamical model for the time evolution of the atmospheric flow. The fields produced by an analysis or an assimilation must satisfy two basic requirements: they must be close to the observations and, they must verify dynamical and/or statistical relationships which are known to be satisfied by the real atmospheric fields.

In atmospheric modeling the scientist is generally faced with a set of observations of parameters, for instance, wind, temperature, water, ozone, etc., as well as either the knowledge or expectation of correlated behavior between the different parameters. The use of numerical techniques to represent the partial differential equations that represent the model physics is a straightforward way to develop a model. There are many approaches to discretization of the dynamical equations that govern geophysical processes [1].

The numerical weather prediction (NWP) was confronted with having to solve an initial-value problem. Data assimilation is the objective melding of observed information with model-predicted information. Data assimilation rigorously combines statistical modeling with physical modeling; thus, formally connecting the two approaches.

Atmospheric Data Assimilation is an important task in NWP Centers. Several methods of data assimilation have been applied in models of atmospheric and oceanic dynamics. Methods using Artificial Neural Networks (ANN) have been proposed showing consistent results regarding implementation in simple models. This paper presents an ANN approach to emulate an Ensemble Kalman Filter (EnKF) as a method of data assimilation with an Atmospheric General Circulation Model(AGCM), with dynamical nonlinear, using synthetic temperature data simulating satellite radiances.

In meteorological data assimilation the conventional data are very important to the quality of the analysis and the forecast. The satellite data assures high quality global analyses. It is very clear that assimilation of satellite observations will make a key contribution to that improvement in forecast skills, given the future growth (five orders of magnitude increase in satellite data over ten years) and improvement of the global observing system expected in the area of space-borne observing systems. As a result there is a need an assimilation method able to get the initial field for the numerical model in time to make a prediction. At present most NWP centers cannot assimilate all the data due to computational costs and limitations in storing the data.Operational satellite data are taken and processed in real-time and distributed around the world.

Data assimilation adds an additional forcing to the representative equations of the physical model; namely, information from the observations. From mathematical point, the assimilating process can be represented by

$$x^a = x^f + W[y^o - H(x^f)], \tag{1}$$
$$W = (HP^fH^T + R) \tag{2}$$

the in the equation are as follows: x^a is the analysis filed with innovation that represents the observation-based correction to the model; y^o are observations of the constituent, x^f is a model forecast, simulated, estimates of the constituent often called the first guess; H is the observation operator andthe superscript T is the matrix transform operation.. W is the weighting matrix, generally computed from the covariance matrix of the prediction errors from forecasting and observation. P^f is the error covariance function of the forecast; R is the error covariance function of the observations. This explicitly shows that data assimilation is the error-weighted combination of information from two primary sources.

The analysis is the best estimate of the state of the system based on the optimization criteria and error estimates.The observation operator, H, is a function that maps the parameter to be assimilated onto the spatial and temporal structure of the observations. In the case of satellite observations the measurements are radiances and the observation operator might include a forward radiative transfer calculation from the model"s geophysical parameters to radiance space.

The computational challenge to the traditional techniques of data assimilation lies in the size of matrices involved in operational NWP models, currently running at a million equations (equivalent to full matrix elements of the order of 10^{12}!). In this scenario the applications of ANN in data assimilation were suggested by [2, 3], [4] and [5]. But the first implementation of the ANN as a approach for data assimilation was employed by [6, 7], improved by [8, 9]. [6] used an ANN over the entire domain space, the [9] strategy is generate the analysis at each grid point, it had large gain in computacional efforts. Continuing these investigations, [10] evaluated the performance of an ANN to emulate Kalman filter, the Particulate Filter and Variational data assimilation method applied to Lorenz chaotic system. And em [11], this approach was applied with a atmospheric circulation model with sybthetic conventional data.

The ANN technique uses neural networks to implement the function:

$$x^a = F_{RNA}(y^o, x^f) \tag{3}$$

where F_RNA is the data assimilation process.

Methods using Artificial Neural Networks (ANN) have been proposed showing consistent results regarding implementation in simple models. This paper presents a experiment using an Atmospheric General Circulation DYnamics (AGCM) the SPEEDY model, see [12], which is a 3D dynamic model, with simplified physics parameterization by [13]. Here is employed a set of Multilayer Perceptron (MLP)(see [14]), which were trained to emulate the LETKF (Local Ensemble Transform Kalman Filter) by [15, 16].

2. Methodology

2.1 The SPEEDY Model

The model *Simplified Parameterizations PrimitivE-Equation Dynamics* (SPEEDY) is an atmospheric general circulation model (AGCM) developing to study global-scale dynamics and numerical weather prediction. The dynamic variables on the primitive meteorological equations are integrated by spectral method in the horizontal at each vertical level, see [17, 18]. The model has a simplified set of physical parameterization schemes that are similar to realistic weather forecasting numerical models. The goal of this model is to achieve computational efficiency while maintaining characteristics similar to the state-of-the-art AGCMs with complex physics.

The model is global with spectral resolution T30L7 (horizontal truncation of 30 numbers of waves and seven vertical levels), corresponding to regular grid with 96 zonal points (longitude), 48 meridian points (latitude) and 7 vertical pressure levels (100, 200 , 300, 500, 700, 850, 925 hPa). The computational cost is one order of magnitude less than that of state-of-the-art AGCMs at similar horizontal resolution. According to [13], the SPEEDY model simulates the general structure of global atmospheric circulation fairly well, and some aspects of the systematic errors are similar to many AGCMs. The package is based on same physical parameterizations adopted in more complex schemes of AGCM like convection (simplified diagram of mass flow), large-scale condensation, clouds, short-wave radiation (two spectral bands), long-wave radiation (four spectral waves), surface fluxes of momentum and energy (aerodynamic formula), and vertical diffusion. Details of the simplified physical parameterization scheme can be found in [13].

The boundary conditions of the SPEEDY model includes topographic height and land-sea mask, which are constant, and sea surface temperature (SST), sea ice fraction, surface temperature in the top soil layer, moisture in the top soil layer and the root-zone layer, and snow depth, all of which are specified by monthly means, and bare-surface albedo and fraction of land-surface covered by vegetation, which are specified by annual-mean fields. The lower boundary conditions such as SST are obtained by ECMWF"s reanalysis in the period 1981-90. The incoming solar radiation flux and the boundary conditions (SST etc.), except bare-surface albedo and vegetation fraction, are updated daily. The SPEEDY model is a hydrostatic model in sigma coordinates, and the transformed vorticity-divergence scheme is described by [17]. The prognostic variables of input and output model are the absolute temperature (T), surface pressure (ps) component of zonal wind (u), component of meridian wind (v) and an additional variable and specific humidity (q).

2.2 Brief Description on Local Ensemble Transform Kalman Filter

The analysis is the best estimate of the state of the system based on the optimization criteria and error estimates. The probabilistic state space formulation and the requirement for the updating of information when new observations are encountered are ideally suited for the Bayesian approach. The Bayesian approach and in ensemble Kalman Filter (EnKF) or particle filtering (PF) methods are a set of efficient and flexible Monte-Carlo methods to solve the optimal filtering problem. Here one attempts to construct the posterior probability density function (PDF) of the state based on all available information, including the set of received observations. Since this PDF embodies all available statistical information, it may be considered to be a complete solution to the estimation problem. In the field of data assimilation, there are only few contributions in sequential estimation (EnKF or PF filters).

The ensemble Kalman filter (EnKF) was first proposed by [19] and developed by [20] and [21]. It is related to particle filters in the context that a particle is identical to an ensemble member. EnKF is a sequential filter method, which means that the model is integrated forward in time and, whenever observations are available; these are used to reinitialize the model before the integration continues. The EnKF originated as a version of the Extended Kalman Filter (EKF) by [22], for large problems. The classical KF [23] method is optimal in the sense of minimizing the variance only for linear systems and Gaussian statistics. Similar to the particle filter method, the EnKF systems from a Monte Carlo integration of the Fokker-Planck equation governing the evolution of the PDF that describes the prior, forecast, and error statistics. In the analysis step, each ensemble member is updated according to the KF scheme and replaces the covariance matrix by the sample covariance computed from the ensemble.

The first application of EnKF to an atmospheric system by [24, 25]. It applies an ensemble of model states to represent the model statistical error of the estimate. The scheme of analysis acts directly on the ensemble of model states when observations are assimilated. The ensemble of analysis is obtained by assimilation of perturbed observations for each member of the set of the reference model. Several ways of perturbed observations used to represent the covariance matrix of the analysis, have been derived many schemes from the EnKF approach: the Local Ensemble Transform Kalman Filter (LETKF) is one of them.

LETKF was proposed by [15] as an efficient upgrade of LEKF [26]. LETKF separate the entire global grid into independent local patches. The LETKF scheme first separates an entire grid vector into local patch vectors with observations. The basic idea of LETKF is perform analysis at each grid point simultaneously using the state variables and all observations in the local region centered at that point. Each member of the ensemble gets its forecast:

$$x_{n-1}^{f(i)} \ : \ i = 1, 2, 3 \cdots k$$

where k is the total members at time t_n, to estimate the state vector $\bar{x}^f$ of the reference model is used the mean the ensemble forecasts:

$$\bar{x}^f = k^{-1} \sum_{i=1}^{k} x^{f(i)}, \tag{4}$$

then the model error covariance matrix is:

$$P^f = (k-1)^{-1} \sum_{i=1}^{k} (x^{f(i)} - \bar{x}^f)(x^{f(i)} - \bar{x}^f)^T. \tag{5}$$

and the ensemble analysis is:

$$x_{n-1}^{a(i)} \ : \ i = 1, 2, 3 \cdots k$$

with the its average and error covariance. LETKF in the local analysis, allows different linear combinations of the ensemble members in different regions, and comprehensive analysis explores a larger spatial scale. For local implementation separates groups of neighboring observations to a central point for a region of the model grid. The LETKF scheme first separates an entire grid vector into local patch vectors. Each grid point has on local patch, the number of vectors equals the number of global grid points, each local patch is treated independently [12].

2.3 Multilayer Perceptron

Different ANN architectures are dependent upon the learning strategy adopted. Detailed introduction on ANNs can be found in [27] and [28].

Multilayer perceptrons (MLP) with backpropagation learning algorithm are feedforward networks composed of an input layer, an output layer, and a number of hidden layers, whose aim is to extract high order statistics from the input data [29] Neural networks will solve nonlinear problems, if nonlinear activation functions are used for the hidden and/or the output layers.

A feedforward network can input vectors of real values onto output vector of real values. The connections among the several neurons have associated weights that are adjusted during the learning process, thus changing the performance of the network. Two distinct phases can be devised while using an ANN: the training phase (learning process) and the run phase (activation of the network). The training phase consists of adjusting the weights for the best performance of the network in establishing the mapping of many input/output vector pairs. Once trained, the weights are fixed and the network can be presented to new inputs for which it calculates the corresponding outputs, based on what it has learned.

The backpropagation training is a supervised learning algorithm that requires both input and output (desired) data. Such pairs permit the calculation of the error of the network as the difference between the calculated output and the desired vector. The weight adjustments are conducted by backpropagating such error to the network, governed by a change rule. The weights are changed by an amount proportional to the error at that unit, times the output of the unit feeding into the weight. 6 shows the general weight correction according to the socalled the delta rule

$$\Delta w_{ij} = \eta \delta_j y_i \tag{6}$$

where, δ_j is the local gradient, y_i is the input signal of neuron j, and η is the learning rate parameter that controls the strength of change.

3 MLP-ANN and experimental settings

The experiment was conducted with the forecast SPEEDY model mentioned and LETKF core modules are applicable to any dynamical model to obtain training data to ANN. The Fortran90 codes (SPEEDY and LETKF) originally developed by [12]. The error covariance at upper levels and surface is treated in the same way as in [12] to LETKF data assimilation.

The observational data used in data assimilation simulating satellite data. These data include vertical observations from radiance (temperature) profiles. The satellite data have been essential of weather forecasts.

The observations were generated from "true" model fields, adding random noise with standard deviation (1) to temperature values. The variables were located at some grid points model. The grid points chosen were simulating satellite observations, obtaining values for a merged point model (a grid point has observation and a grid point no observation) shown in Figure 1. Both assimilation scheme, LETKF or ANN, have the same numbers of observations at the same grid points.

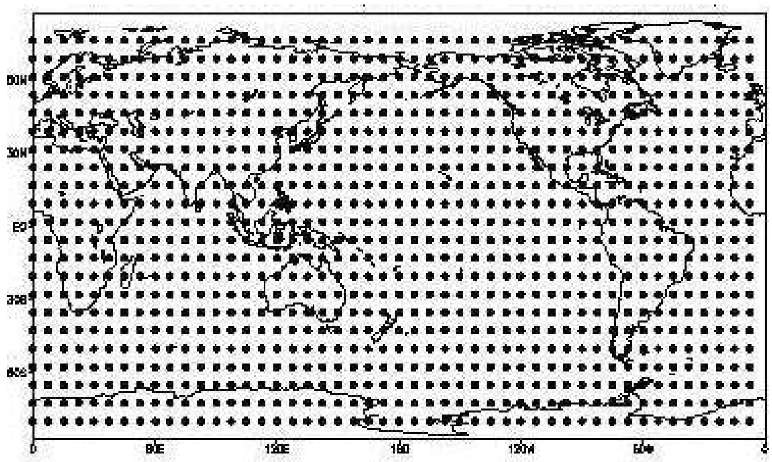


Figura 1: Dense network of observations synthetic regularly distributed grid points in the SPEEDY model. Representing 1056 stations of satellite data (23 per cent of total observations).

In this configuration SPEEDY was run for a long time integrations of the state to create "true" to start the integrations of the model. The true integration of the model was made for three years: from 1 January 1982 to 31 January 1985, generating outputs in four times a day (00, 06, 12 and 18 UTC). The data assimilation (DA) LETKF was performed with these synthetic observations (about 6,600 points) of the temperature variables to generate the vector analysis and obtain the desired output to train the neural network. The executions of the model with LETKF were made for the periods mentioned to true model. The ensemble of forecasts LETKF has 30 members and the "perturbations of ensemble" consist of random numbers with Gaussian distribution. These integrations were obtained input vectors for ANN, recent forecast from SPEEDY with LETKF analysis.

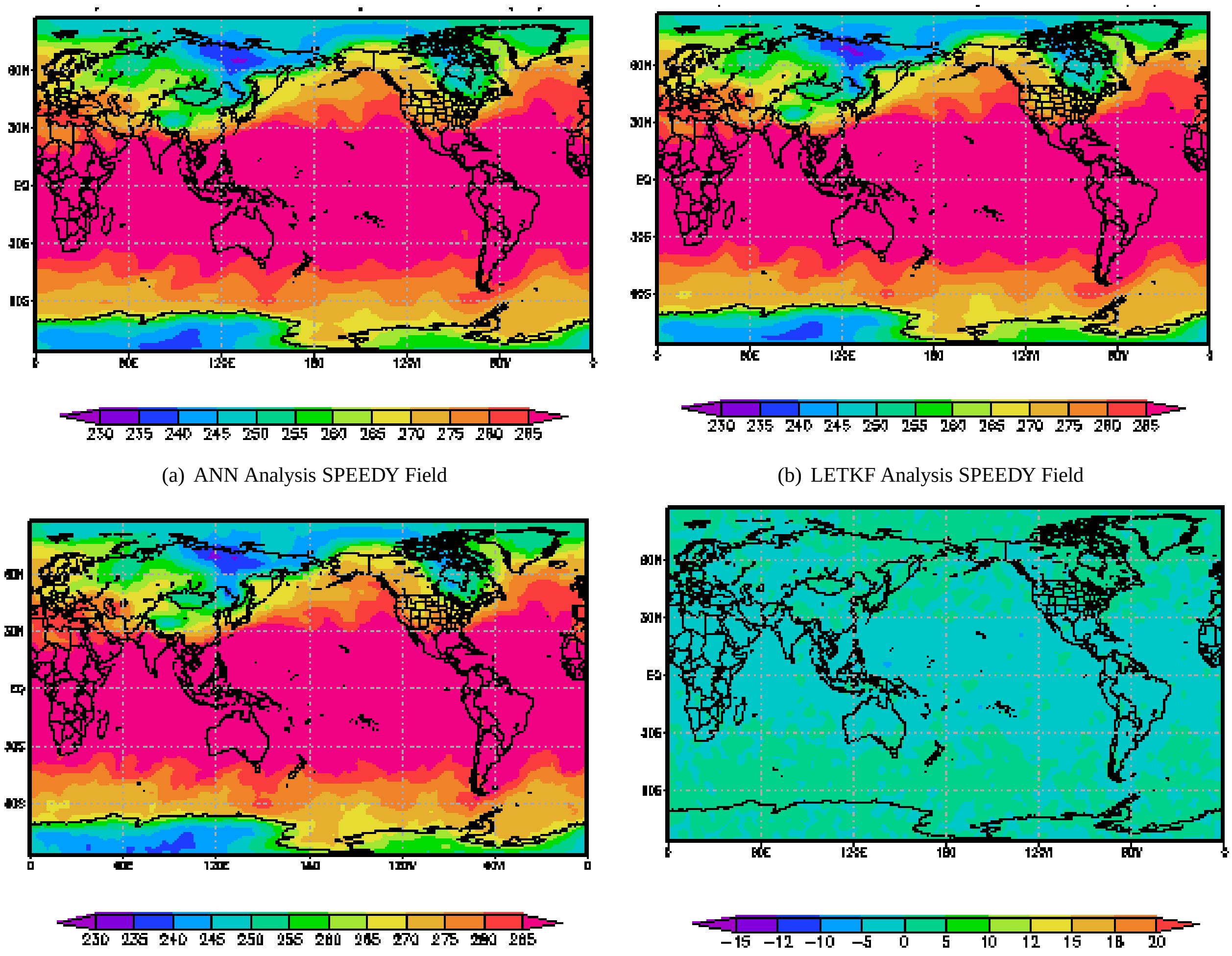
In ANN data assimilation scheme we need the local observation influence at neighboring grid points like observations point, in zonal and meridional indices, as for the vertical boundaries at the bottom and top levels. We eliminate indices at poles (i.e. no observations at boundaries grid points). This calculation was based on the distance from its neighbor:

$$\hat{y}^o = \frac{y_l^o + \sum_{l=1}^{N}(y_l^o + \delta)}{r^2} \tag{7}$$

$$r = \text{distance}(y_l^o + \delta,\ y_l^o) \tag{8}$$

where N is the total neighbors' grid points **without** observations, δ is a cubelike shape characterized by the horizontal and vertical grid lengths, this is based on the distance of each observation point: $\delta = (x_i^f - y_i^o)^2 + (x_j^f - y_j^o)^2 + (x_k^f - y_k^o)^2$ where x^f is grid point without observation that has the forecast variable value, y^o is observation in a grid point; the subscripts i, j, k, are the indices of localizations points as latitude, longitude, level respectively. The input vectors (observations, forecast value and analysis value at grid points) was collected for two vertical regions in the globe(the northern hemisphere and southern hemisphere), or 90^o to north and 90^o to south; three horizontal regions to each hemisphere, or 90^o to each region. This division is based on size of regions not in number os observations. Because of this division we developed a set of thirty MLP to temperature prognostic variable. These MLP has two inputs (model and forecast vectors), one neuron in the output (to analysis vector), eleven neurons in a hidden layer, the activation function to ensure nonlinearity of the problem used was the sigmoid tangent hyperbolic.

The network was trained by entering input values of each grid point once, i.e, the analysis is done at grid point where it has observation. The training was made with collected data of the first three months of 1982, 1983 and 1984. The MLP generalization is initiated with ANN data assimilation cycle in the first 00 UTC on 1 January 1985, generating the prediction model and the initial condition of SPEEDY (new global analysis). In the experiment the MLP generated analysis and forecasts until 01/31/1985



(a) ANN Analysis SPEEDY Field

(b) LETKF Analysis SPEEDY Field

(c) True SPEEDY Field

(d) ANN-LETKF

Figura 2: Analysis fields for Temperature to 03/01/1985 12 UTC to level 950 hPa.

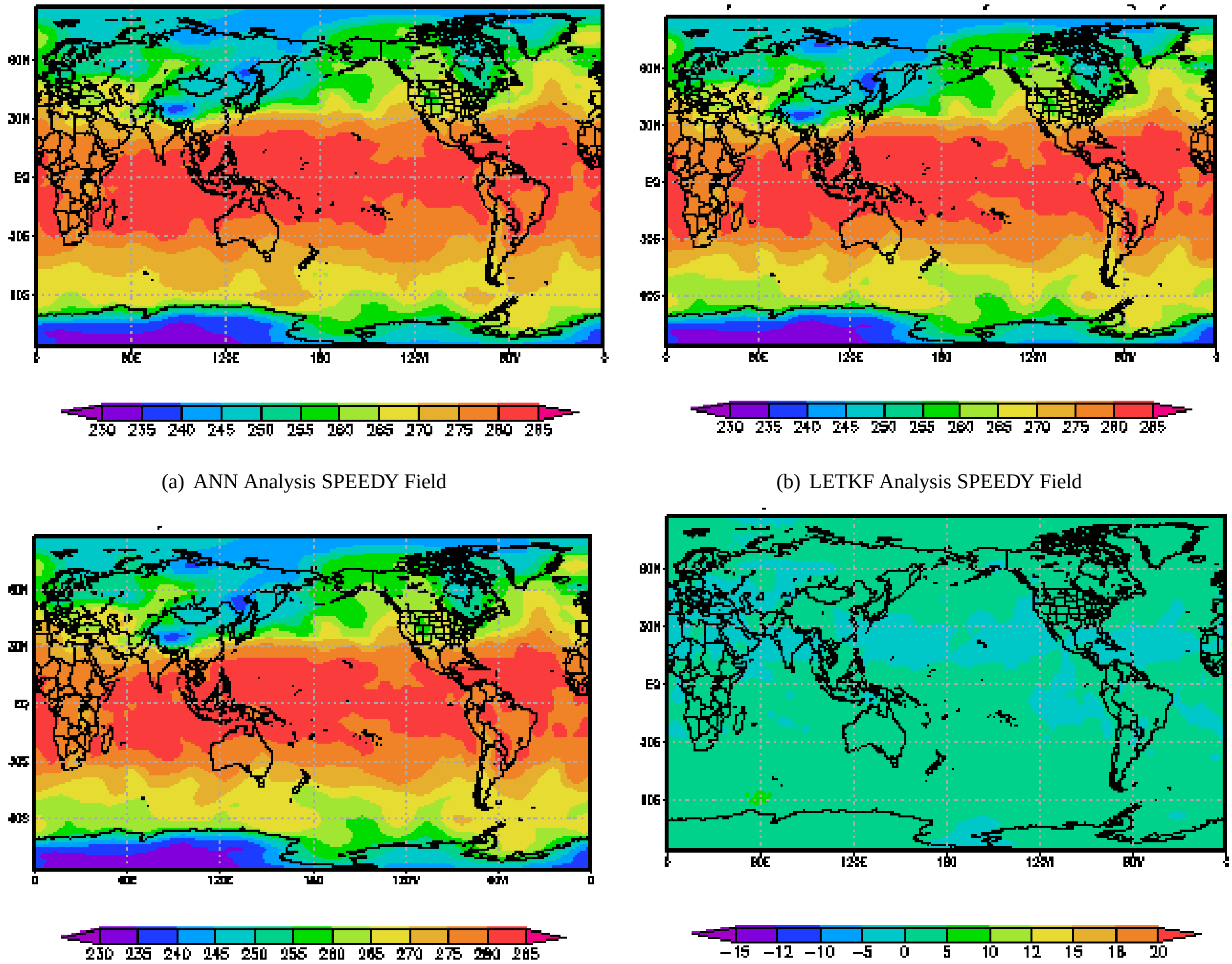
4. RESULTADOS

The input and output values were processed in grid points in time integrations for an intermittent forecasting and analysis cycle to the temperature (T) prognostic variables. The results show analysis fields generated by the activation of the MLP-ANN and analysis fields generated by LETKF data assimilarion for 03/Jan/1985 at 12 UTC at level 950 hPa (near surface) and hight level 500 hPa. Figures 2 to 3 present global fields of analysis, true model fields and the differences between analysis fields. These results show that the application of MLP-ANN as assimilation system generates data analysis similar to the assimilation system LETKF. The first conclusion of this ecxperiment is: the MLP-ANN can emulate the analysis of LETKF to temperature observations from sattelites sensors.

There are several aspects of the modeling and assimilation problem that stress computational systems and push capability requirements. The common ones in modeling are increased resolution, improved physics, inclusion of new processes, and integration and concurrent execution of Earth-system components. Often, real-time needs define capability requirements. When considering data assimilation the computational requirements become much more challenging. The use of observations from the earth-orbiting satellites in operational numerical prediction models is performed for improving weather forecasts; however, the use of this amount of data increases the computational effort. As a result there is a need an assimilation method able to get the initial field for the numerical model in time to make a prediction. At present most numerical weather prediction centers cannot assimilate all the data due to computational costs and limitations in storing the data.

The figure 4 shows a cycle of 124 data assimilation, made for 31 days with 7032 satellite observations run in hours (00, 06, 12 and 18 UTC). The time was measured in milliseconds, for LETKF data assimilation and MLP-ANN activation for data assimilation These measures show the computational performance of the ANN. It was higher than the performance of the system LETKF. These results show that the computational efficiency of neural network to the problem of atmospheric data assimilation is better with the similar quality in analysis field. The CPU-time assimilation with MLP-NN is 75 times faster than LETKF in our numerical experiment, see table 1. Total run time including forecasting and 30 members for ensemble to SPEEDY model. Analysis time is only LETKF run time or MLP-ANN activation run time (both for 124 runs).

Actually, considering the supervised ANN for data assimilation, the most relevant issue is the computational speed-up for



(a) ANN Analysis SPEEDY Field

(b) LETKF Analysis SPEEDY Field

(c) True SPEEDY Field

(d) ANN-LETKF

Figura 3: Analysis fields for Temperature to 03/01/1985 12 UTC to level 500 hPa.

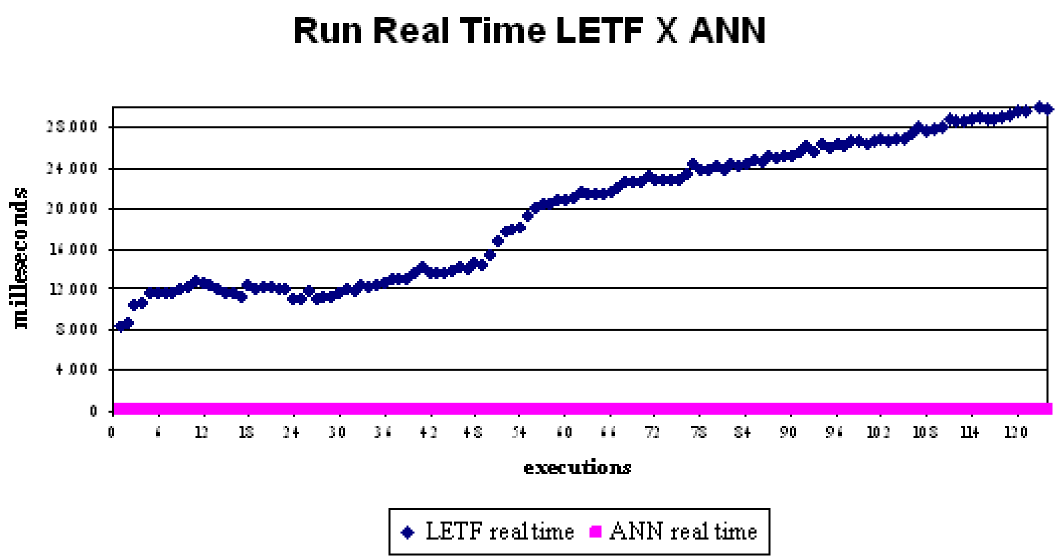


Figura 4: Real time for experiment: MLP-NN (pink color) and LETKF (blue color) methods for one month (124 cycles) assimilation

computing the analyzed initial condition.

Tabela 1: CPU-run time of 124 cycles of data assimilation (analysis and forecasting).

CPU-time of 124 cycles	MLP-NN (hh:mm:ss)	LETKF (hh:mm:ss)
Total run time	**00:04:11**	**04:47:43**
Total real Analysis time	**00:00:22**	**02:45:50**

5 Acknowledgment

First I would like to thank Prof. Dr. José Demísio Simões da Silva who gave me the honor to work at INPE with him. The authors thank Dr. Takemasa Miyoshi and Prof. Dr. Eugênia Kalnay for providing routines for the SPEEDY model and LETKF system.

REFERENCES

[1] D. A. Randall. "An Introduction to The General Circulation of the Atmosphere", 2009. Course Outline AT605 – Colorado State University.

[2] K. H. W. W. Hsieh, B. Tang. "A neural network atmospheric model for hybrid coupled modeling". *Climate Dynamics*, vol. 17 (5–6), pp. 445–455, 2001.

[3] Y. Tang and W. W. Hsieh. "Coupling neural networks to incomplete dynamical systems via variational data assimilation". *Mon. Wea. Rev.*, vol. 129, pp. 818–834, 2001.

[4] A. Liaqat, M. Fukuhara and T. Takeda. "Application of neural network collocation method to data assimilation". *Comput. Phys. Commun.*, vol. 141, pp. 350–364, 2001.

[5] Y. Tang, W. W. Hsieh, B. Tang and K. Haines. "A neural network atmospheric model for hybrid coupled modelling". *Climate Dynamics*, vol. 17, Numbers 5-6, pp. 445–455, March 2001.

[6] A. Nowosad, A. R. Neto and H. F. Campos Velho. "Data Assimilation in Chaotic Dynamics Using Neural Networks". In *Anais...*, pp. 212–221. International Conference on Nonlinear Dynamics, Chaos, Control and Their Applications in Engineering Sciencesa, 2000.

[7] N. L. Vijaykumar, H. F. Campos Velho, S. Stephany, A. J. Preto and A. G. Nowosad. "A neural network implementation for data assimilation using MPI, application of high performace computing in engineering". In *Application of high performace computing in engineering*, edited by C. A. Brebia, P. Melli and A. . Zanasi, pp. Section 5, 211–220. WIT Press, Southampton, 2002.

[8] F. Härter and H. F. Campos Velho. "Recurrent and Feedforward Neural Networks Trained with Cross Correlation Applied to the Data Assimilation in Chaotic Dynamic". *Revista Brasileira de Meteorologia*, vol. 20, pp. 411–420, 2005.

[9] F. P. Härter, E. L. Rempel, H. F. Campos Velho and A. Chian. "Application of Artificial Neural Networks in Auroral Data Assimilation". *Journal of Atmospheric and Solar – Terrestrial Physics*, vol. 70, no. 10, pp. 1243–1250, março 2008.

[10] H. C. Furtado. "Redes neurais e diferentes métodos de assimilação de dados em dinâmica não linear". Master's thesis, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2008.

[11] R. S. Cintra, H. F. Campos Velho and R. Todling. "Nova implementação em filtro do kalman estendido para assimilação de dados com redes neurais". *Learning and Nonlinear Model*, vol. 7, pp. 30–37, 2010.

[12] T. Miyoshi. "Ensemble Kalman filter experiments with a primitive–equation global model". Ph.D. thesis, University of Maryland, College Park, Maryland, USA, 2005. 197 p.

[13] F. Molteni. "Atmospheric simulations using a GCM with simplified physical parametrizations. I:Model climatology and variability in multi-decadal experiments". *Climate Dyn.*, vol. 20, pp. 175–191, 2003.

[14] R. Daley. *Atmospherics aata analysis*. Cambridge University Press, New York, USA, 1991. 471 p.

[15] B. Hunt, E. Kalnay, E. J. Kostelich, E. Ott, D. Patil, T. Sauer, I. Szunyogh, J. A. Yorke and A. V. Zemin. "Four-dimensional ensemble kalman filtering". *Tellus*, vol. 56A, pp. 273–277, 2004.

[16] H. C. Bishop, B. J. Etherton and S. J. Majumdar. "Adaptive Sampling With the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects". *Monthly Weather Review*, vol. 129, pp. 420–436, 2001.

[17] W. Bourke. "A multilevel spectral model: I. Formulation and hemispheric integrations". *Monthly Weather Review*, vol. 102, pp. 687–701, 1974.

[18] I. M. Held and M. J. Suarez. "A proposal for the intercomparison of dynamical cores of atmospheric general circulation models". *Bull. American Meteorological Society*, vol. 75, pp. 1825–1830, 1994.

[19] G. Evensen. "Sequential data assimilation with a nonlinear quasi–geostrophic model using Monte Carlo methods to forecast error statistics". *J. Geophys. Res*, vol. 99, pp. 10143–10162, 1994.

[20] G. Burgers, P. J. van Leeween and G. Evensen. "Analysis Scheme in the Ensemble Kalman Filter". *Monthly Weather Review*, vol. 126, pp. 1719–1724, 1998.

[21] G. Evensen. "The ensemble Kalman filter: theoretical formulation and practical implementation". *Ocean Dynamics*, vol. 53, pp. 343–367, 2003.

[22] R. E. Kalman and R. S. Bucy. "New results in linear filtering and prediction theory." *Trans. of the ASME–Journal of Basic Engineering*, vol. 83(Series D), pp. 95–108, 1961.

[23] R. E. Kalman. "A new approach to linear filtering and prediction problems." *Trans. of the ASME–Journal of Basic Engineering*, vol. 82(Series D), pp. 35–45, 1960.

[24] P. L. Houtekamer and H. L. Mitchell. "Data Assimilation Using an Ensemble Kalman Filter Technique". *Monthly Weather Review*, vol. 126, pp. 796–811, 1998.

[25] P. L. Houtekamer, H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek and B. Hansen. "Atmospheric data assimilation with the ensemble Kalman filter: results with real observations". *Monthly Weather Review*, vol. 133, pp. 605–620, 2005.

[26] E. Ott, B. R. Hunt, I. Szyniogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil and J. York. "A local ensemble Kalman filter for atmospheric data assimilation". *Tellus A*, vol. 56, pp. 415–428, 2004.

[27] S. Haykin. *Redes neurais princípios e prática*, volume 2. Editora Bookman, Porto Alegre, 2001.

[28] L. Tsoukalas and R. Uhrig. *Fuzzy and Neural Approaches in Engineering*. John Wiley and Sons, New York, 1997.

[29] F. Miki, E. Issamoto, J. da Luz, P. de Oliveira, H. F. de Campos Velho and J. da Silva. "A Neural Network Approach in a Backward Heat Conduction Problem". In *Proceedings of the Braz. Conf. Neural Networks, Proc. in CD-ROM - paper code 0008,* São José dos Campos-SP, 1999.