Brazilian Journal of Geophysics (2020) 38(2): 1–18
© 2020 Sociedade Brasileira de Geofísica
ISSN 0102-261X
www.rbgf.org.br
DOI: 10.22564/rbgf.v38i2.2042

ARTIFICIAL NEURAL NETWORK FOR DATA ASSIMILATION BY WRF MODEL IN RIO DE JANEIRO, BRAZIL

Vinícius Albuquerque de Almeida[1], Gutemberg Borges França[1],
Haroldo Fraga Campos Velho[2] and Nelson F. Favilla Ebecken[3]

ABSTRACT. This study investigates the use of neural networks for data assimilation of local data in the WRF model in Rio de Janeiro, Brazil. Surface and upper-air data (air temperature, relative humidity, wind speed and direction) from airport stations and 6-hour forecast from WRF are used as input for the model and the 3D-Var analysis for each grid point is used as target variable. Periods of 168h from 2014 and 2015 are used with 6h and 12h assimilation cycles for surface and upper-air data, respectively. The neural network model was built using the Multi-Particle Collision Algorithm (MPCA) where different topologies are tested until the optimum solution is found. Results show that the neural network is able to emulate the 3D-Var with root mean squared error (standard deviation), respectively, of 0.31 K (0.37 K), 3.10% (4.04%), 0.63 ms^{-1} (1.05 ms^{-1}), 1.10 ms^{-1} (1.56 ms^{-1}) for air temperature, relative humidity, u-component of the wind and *v*-component of the wind. Also, the results show the neural network method is able to run 71 times faster than the conventional method under similar hardware configurations.

Keywords: data assimilation; weather research and forecasting; surface data; profile data.

RESUMO. Este estudo investiga o uso de redes neurais para assimilação de dados locais no modelo WRF no Rio de Janeiro. Dados de superfície e do ar superior (temperatura do ar, umidade relativa, velocidade e direção do vento) das estações do aeroporto e previsão de 6 horas do WRF são usados como entrada para o modelo, e a análise 3D-Var para cada ponto da grade é usada como variável destino. Períodos de 168h de 2014 e 2015 são utilizados com ciclos de assimilação de 6h e 12h para dados de superfície e do ar superior, respectivamente. O modelo de rede neural foi construído usando o algoritmo de colisão de partículas múltiplas (MPCA), onde diferentes topologias são testadas até que a solução ideal seja encontrada. Os resultados mostram que a rede neural é capaz de emular o 3D-Var com raiz do erro quadrático médio (desvio padrão) de 0,31 K (0,37 K), 3,10% (4,04%), 0,63 ms^{-1} (1,05 ms^{-1}), 1,10 ms^{-1} (1,56 ms^{-1}) para temperatura do ar, umidade relativa, componente u do vento e componente v do vento. Além disso, os resultados mostram que o método de rede neural é capaz de rodar 71 vezes mais rápido que o método convencional em configurações de hardware semelhantes.

Palavras-chave: assimilação de dados; dados de superfície; dados de perfil.

Corresponding author: Vinícius Albuquerque de Almeida

[1] Universidade Federal do Rio de Janeiro, Laboratory for Applied Meteorology, Rio de Janeiro, RJ, Brazil – Emails: vinicius@lma.ufrj.br, gutemberg@lma.ufrj.br

[2] Instituto Nacional de Pesquisas Espaciais, Laboratory for Computing and Applied Mathematics, São José dos Campos, SP, Brazil – Email: haroldo.camposvelho@inpe.br

[3] Universidade Federal do Rio de Janeiro, COPPE, Rio de Janeiro, RJ, Brazil – Email: nelson@ntt.ufrj.br

INTRODUCTION

Numerical weather prediction (NWP) is considered an initial-value problem where the present state of the atmosphere is used as input to a numerical model for simulating or forecasting its evolution on space and time.

This is a remarkable key point in a scientific conquer to the geophysical fluid dynamics, with very good impact into many economic sectors: agriculture, prevention and/or mitigation of natural disasters, insure and tourism industries, just for mention few sectors.

The problem of the initial condition determination for a forecast model is essential and complex, and has become a science in itself (Daley, 1991). Several methods have been developed since the 1950s to tackle this problem. Daley (1991), Talagrand (1997), Zupanski & Kalnay (1999), and Kalnay (2003) provide to a broader review on data analysis and assimilation techniques.

Data assimilation is a relevant topic for research and operational issues for many branches of geosciences, including ocean circulation (Wang et al., 2007) – using 3D-Var method), atmosphere-ocean coupled model (Smith et al., 2015) – employing 4D-Var method), ionospheric dynamics (Lin et al., 2017) – using Kalman filter), hydrology (Vrugt & Schoups, 2018) – applying particle filter method), snow dynamics (Piazzi et al., 2018) – using particle filter method), environmental prediction (Khassenova & Kussainova, 2018) – by variational method). In meteorology, there is a wide variety of data sources to be assimilated to accurately estimate the state of the atmosphere, which includes conventional and non-conventional data. Conventional data include surface meteorological stations, balloon soundings, aircraft and ship observations. On the other hand, data retrieved from satellites (e.g. radiance), wind profilers (e.g. SODAR, LIDAR), and radar are usually known as non-conventional, due to inhomogeneity of their spatial-temporal distribution. Conventional and non-conventional data are commonly assimilated in global models. But, very often for the local conditions (regional models), the data from global models are smoothed due to interpolation methods and quality control routines. Also, not all observations are part of the global observation network and they are not processed by data assimilation routines for global models. Therefore, to accurately determine the state of the atmosphere for regional models, it is mandatory not only to employ the global model's analysis, but reinforcing the assimilation with local retrieved data. According to Cintra and Campos Velho (2012), 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 of the order $\sim O(10^{12})$. In this scenario, the applications of Artificial Neural Networks (ANN) in data assimilation is suggested for reducing the computational effort. The neural network technique is applied to implement the mapping: $x^a = F[y^o, x^f]$, where x^a is the *analysis* field – the estimated initial condition – representing the observation-based correction to the model, F is the data assimilation process, y^o is the vector of

observations of the constituent, x^f is the model forecasting field (simulation) that estimates the constituent – often called the *first guess*. A supervised neural network is trained by a set of analysis obtained from another assimilation method. Methods using ANN have been proposed showing consistent results regarding implementation in simple models, see Nowosad, 2001; Furtado, 2008; Cintra, 2010; Härter & De Campos Velho, 2012; França et al., 2018; Almeida et al., 2020. The present article is part of a sequence of studies related to nowcasting that have been executed by the Laboratory for Applied Meteorology at the Federal University of Rio de Janeiro, following Almeida, 2009; Silva et al., 2016; França et al., 2016; França et al., 2018; Paulucci et al., 2019; Almeida et al., 2020. All these studies encompass researches based on artificial intelligence and methods dealing with models focused on numerical weather forecasts. This paper relates to the latter, exploring the sensibility of the Weather Research and Forecasting (WRF) – a sophisticated mesoscale regional model - for surface and upper-air data assimilation using artificial neural networks in the metropolitan area centered at the Galeão airport of the Rio de Janeiro city, searching for efficient ways to reduce the CPU time of the assimilation process and, thus, enable faster assimilation cycles with the growing number of available datasets. The paper is organized as: section Material and Methods gives a brief description of the dataset used in this study, the WRF model and data assimilation methods using 3D-Var scheme and artificial neural network including a technique for finding optimal neural network configuration; next section presents the results and discussions; finally, the conclusions are presented with the main findings of this study.

MATERIALS AND METHODS

The study area is the metropolitan area of Rio de Janeiro and its surroundings (Fig. 1) located approximately at latitude 22°55’44.3"S and longitude 43°24’21.1"W. The most import airports in the region are located in Figure 1 identified by their International Civil Aviation Organization (ICAO) codes: Santos Dumont Airport (SBRJ), Galeão International Airport (SBGL), Santa Cruz Air Force Base (SBSC), Jacarepaguá Airport (SBJR) and Afonsos Air Force Base (SBAF). Each airport is responsible for local hourly routine and special reports surface observations of several meteorological parameters as surface wind (direction and speed), visibility, significant weather, cloud cover, air and dewpoint temperature, and station pressure. Besides, the SBGL airport has an upper-air (or sounding) station that produces regularly atmospheric soundings twice a day, the atmospheric profile of pressure, air and dewpoint temperature, relative humidity, and wind (direction and speed), from the surface up to more than 25 km.

Figure 1 – Domain and computational grid. The labels SBSC, SBAF, SBJR, SBRJ and SBGL are located at the airports in the metropolitan area of Rio de Janeiro.

The numerical experiments performed using the NCEP FNL (Final) Operational Global Analysis data. The FNL data are available on 1-degree grids prepared operationally every 6 hours. This product is from the Global Data Assimilation System (GDAS), which continuously collects observational data from the Global Telecommunications System (GTS), and other sources, for many analyses. The FNLs are made with the same model that NCEP uses in the Global Forecast System (GFS), but the FNLs are prepared for about an hour or so after the GFS is initialized. The FNLs are delayed so that more observational data can be used. The GFS is run earlier in support of time-critical forecast needs and uses the FNL from the previous 6-hour cycle as part of its initialization. The analyses are available on the surface, at 26 mandatory (and other pressure) levels from 1000 millibars to 10 millibars, in the surface boundary layer and at some sigma layers, the tropopause and a few others. More information can be found at https://rda.ucar.edu/datasets/ds083.2.

WRF: Limited-Area Atmospheric Model

The WRF model is a next-generation mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications. It features two dynamical cores, a data assimilation system, and a software architecture supporting parallel computation and system extensibility. The effort to develop WRF began in the latter 1990s and was a collaborative partnership of the National Center for Atmospheric

Research (NCAR), the National Oceanic and Atmospheric Administration (represented by the National Centers for Environmental Prediction (NCEP) and the Earth System Research Laboratory), the U.S. Air Force, the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA). Please refer to the WRF User's Guide and the Technical Note document available at http://www2.mmm.ucar.edu/wrf/users/ for completeness of the 3D-Var implementation present at WRF (Skamarock et al., 2019). The WRF model solves a set of equations modeling the state and evolution of the atmosphere, including: (i) conservation of momentum; (ii) thermodynamic energy conservation; (iii) mass conservation; (iv) geopotential relation; and (v) the equation of state. Also, several physical processes are parameterized (e.g. short and longwave radiation transfer, surface modeling, turbulence, cumulus convection, cloud microphysics and precipitation). These ones are too small, too brief, too complex, too poorly understood, or too computationally costly to be explicitly represented. In our numerical experiments, the WRF model is integrated into a 2-km grid with 35 levels in vertical, generating hourly outputs from the surface and pressure-level variables. Regarding the parametrizations the following options were chosen: Microphysics – WRF Single–moment 3 (Hong et al., 2004), Cumulus – Grell–Freitas Ensemble Scheme (Grell & Freitas, 2014), Radiation – Dudhia Shortwave Scheme (Dudhia, 1989) / RRTM Longwave Scheme (Mlawer et al., 1997), Planetary Boundary Layer – Yonsei University Scheme (YSU) (Hong et al., 2006), and Land Surface model – Unified Noah Land Surface Model (Tewari et al., 2016).

Data assimilation method: 3D-Var

The 3D-Var approach is used as implemented in the Data Assimilation component of the WRF framework. The basic ideas of variational data assimilation and specifically the WRF Data Assimilation (WRFDA) system is deeply discussed in Barker et al. (2012). Among various data assimilation methods, the variational approaches have been widely used in meteorology, specifically the method 3D-Var. In the 3D-Var approach, a cost function (Eq. 1) is defined which is proportional to the square of the distance between the analysis (x^a) and both the background (x^b) and observations (y^o) (Kalnay, 2003). The analysis field is computed by the direct minimization of such function. Important to notice that the error matrices for both the background (B) and observation (R) are considered in the minimization process. The operator H mapped the gridded analysis to the observation space for comparison against the observation matrix y^o. The analysis x^a is computed by minimizing the cost function (J) expressed as:

$$\begin{aligned} \mathrm{J} = 1/2\,\{&[y^o - H(x)]^T R^{-1}[y^o \\ &- H(x)] \\ &+ [x - x^b]^T B^{-1}[x \\ &- x^b]\} \end{aligned} \tag{1}$$

where R is the covariance matrix of the sensor errors, and B is the covariance background matrix. The latter matrix is computed as a vector product from the difference of two WRF executions for a certain initial condition (Barker, et al., 2012). The 3D-Var approach consists in processing observed information in a temporal window (typically from 1 h before the analysis time to 1 h after) over a spatial

domain. After this process, a subset of the observed data is retrieved that will be assimilated in a previous forecast grid by the minimization of a cost function.

Data assimilation method: optimal neural network

Artificial neural networks is a branch of artificial intelligence belonging to the class of machine learning algorithms – see Rosenblatt, 1958; Hopfield, 1982; Rumelhart et al., 1986; Haykin, 1999. An ANN is an arrangement of several connected processing units. These units are called *neurons*, where the weighted inputs can or not be combined with a bias to feed a nonlinear activation function. ANN can be roughly classified into two groups: supervised and unsupervised neural networks. For the first one, there is a reference dataset to be used to identify the connection weights. A very employed supervised ANN is the *multi-layer perceptron* (MLP). The MLP-NN is a supervised network, and it typically consists of a set of layers: the input layer (one or more inputs), one or more hidden layers, and the output layer (one or more outputs). The well known error back-propagation algorithm is a standard procedure to determine the connection weights – the process is named as the *training* or *learning* phase (Haykin, 1999, Section 4.3). There are many parameters or functions to be chosen for configuring the MLP-NN: number of hidden layers, number of neurons for each hidden layer, the type of activation function, and the parameters for the training phase (learning ratio and momentum). In order to find the best architecture to the MLP-NN for our application – a neural network to emulate the 3D-Var method for data assimilation, the problem is addressed as an optimization one by minimizing the functional (Anochi & De Campos Velho, 2014):

$$\mathrm{L}(\mathbf{Q}) = penality \times \left[\frac{\rho_1 E_{\mathrm{train}}(\mathbf{Q}) + \rho_2 E_{\mathrm{gen}}(\mathbf{Q})}{\rho_1 + \rho_2}\right] \quad (2)$$

$$penality = c_1 \exp\{[\#\mathrm{neurons}]^2\} + c_2\{\#\mathrm{epochs}\} + 1 \quad (3)$$

where **Q** is the unknown vector; E_{train} and E_{gen} are respectively training and generalization errors (the square difference between the NN output and the analysis produced by 3D-Var); finally *penality* is a measurement of the neural network *complextiy*. Therefore, the optimal topology for the MLP-NN is looking for the *simplest* neural network with better agreement with the reference datasets (training and generalization). The optimal solution $\mathbf{Q}^*$ is computed by minimizing the functional above (equation 2). The optimization problem is solved by the MPCA metaheurisc described in the next section.

Solving the optimization problem by the MPCA metaheuristic

The MPCA (Multi-Particle Collision Algorithm) is a metaheuristic based on the canonical Particle Collision Algorithm (PCA) developed by Sacco & De Oliveira (2005) – see also Sacco et al. (2006, 2007, 2008), inspired on a neutron traveling inside of a nuclear reactor under absorption and scattering phenomena. There are similarities with the Simulated Annealing (Kirkpatrick et al., 1983) scheme. The MPCA follows the PCA strategy, but with a new feature: the use of several particles, instead of only one particle to act over the search space. The theory behind the MPCA algorithm is detailed by Pacheco da Luz et al. (2008, 2011).

Coordination between the particles was able through a blackboard strategy, where the best fitness information is shared among all the particles in the process. The MPCA is implemented using MPI libraries in a multiprocessor architecture with distributed memory. The MPCA codification is close to the PCA. Assuming the number of calls to the *absorption operator* is equal to the number of calls of scattering operator, and both equal to N, results in a complexity $O(N \times N)$, just checking operations in the inner loops. But due to the new loop, introduced by the multiple particle technique, the number of checking operations can be increased to N^3 operations, considering the number of particles equal to the number of iterations. So, the complexity associated to MPCA will be $O(N^3)$. The parallel procedures can improve the processing by distributing the tasks among p processors. If the number of processors could be $p = N$, being N the number of particles, the computational effort is reduced to $O(N^2)$, such as the standard PCA.

The PCA starts by selecting an initial solution, and it is modified by a stochastic perturbation, leading to the construction of a new solution. The new solution is compared to the old one (the solutions are compared by calculating the fitness of each one), and the new solution can or cannot be accepted. If the new solution is not accepted, a scheme is used to find a new solution. If a new solution is better than the previous one, this new solution is absorbed (absorption is one feature involved in the real collision process). If a worst solution is found, a probability is calculated to find a particle in a different location of the search space, giving the algorithm the capability of escaping a local minimum. The latter procedure is inspired on the scattering process.

Pacheco da Luz et al. (2011) present an application of the MPCA algorithm for solving two inverse problems – formulated as optimization problems. In the conclusion, the authors state the MPCA is an alternative to determine inverse solutions. Nowadays, even personal computers are found with multicore architectures, allowing to apply the execution of an algorithm developed for high performance environments. The results also demonstrate the MPCA convergence to compute a good solution within a reasonable amount of available resources. Anochi (2015) used the MPCA for climate precipitation field prediction in the South, Southeast, and Northeast regions of Brazil. The results suggest that the optimal architecture determined by MPCA was found in a shorter time compared to time a specialist would take to find an acceptable topology. Another advantage is that the automatic strategy discards the need for a specialist in neural networks making the use of neural networks accessible to a larger audience. Additionally, the author suggests that a major advantage of using neural networks is their hardware implementation.

Description of Experiments

Experiments with 1-week data assimilation are performed using the WRF model during the years 2014 and 2015, starting at February 1st with 168h for time-integration (seven days). The data assimilation is carried out every 6 hours for surface variables (air temperature, relative humidity, and wind direction and speed) at the

airport locations, and every 12 hours for upper-air variables (air temperature, relative humidity, and wind direction and speed) at SBGL location. Figure 2 describes the flowchart for the methods performed for the numerical experiments and the neural network training and validation. The experiment steps are described as follows:

1. White-noise perturbation is applied to the background field at the airport locations for surface and upper-air data generating synthetic observations;

2. Synthetic observations are placed on the exact coordinates where real sensors are located;

3. New analysis field is generated from synthetic observations and background field using the 3D-Var data assimilation technique;

4. Steps (i)-(iii) are repeated from Feb/01 to Feb/08 00Z with surface data assimilation every 6h and upper-air data assimilation every 12h;

5. Steps (i)-(iv) are repeated for the same period of 168h for the years 2014 and 2015;

6. The impact of the synthetic observations on the surroundings is computed using 5-grid points radius with the value. In grid points under the influence of more than one station, the inverse of the distance is used as a weighting factor;

7. Synthetic observations, background field, and analysis are employed;

8. A preprocessing is executed for data cleansing and normalization;

9. A shuffle and split are performed on the dataset defining 60% for training, 20% validation, and 20% generalization; and

10. An evaluation is performed comparing the results for the data assimilation process by the 3D-Var data and self-configured neural network.

RESULTS AND DISCUSSION

Table 1 contains the results with self-configured MLP-NN for experiments performed for 4 meteorological variables using the MPCA algorithm. Table 1 is structured as follows: the column 1 lists the variable names, the column 2 presents several parameters obtained for each experiment, and the columns 3 to 7 show values retrieved from five experiments for each variable (described in the first column) and parameter (described in the second column). The MPCA software was applied to determine different parameters from a MLP-NN, such as: number of hidden layers, number of neurons in each hidden layer, type of activation function, and learning process parameters – momentum (α) and learning rate (η). In Table 1, the activation functions codes represent logistic (1), tangent (2), and Gaussian (3). The results show that the experiments number 3, 1, 1, and 1 were defined by the MPCA software as having the optimum topologies for the variables air temperature, relative humidity, u and v wind components, respectively.

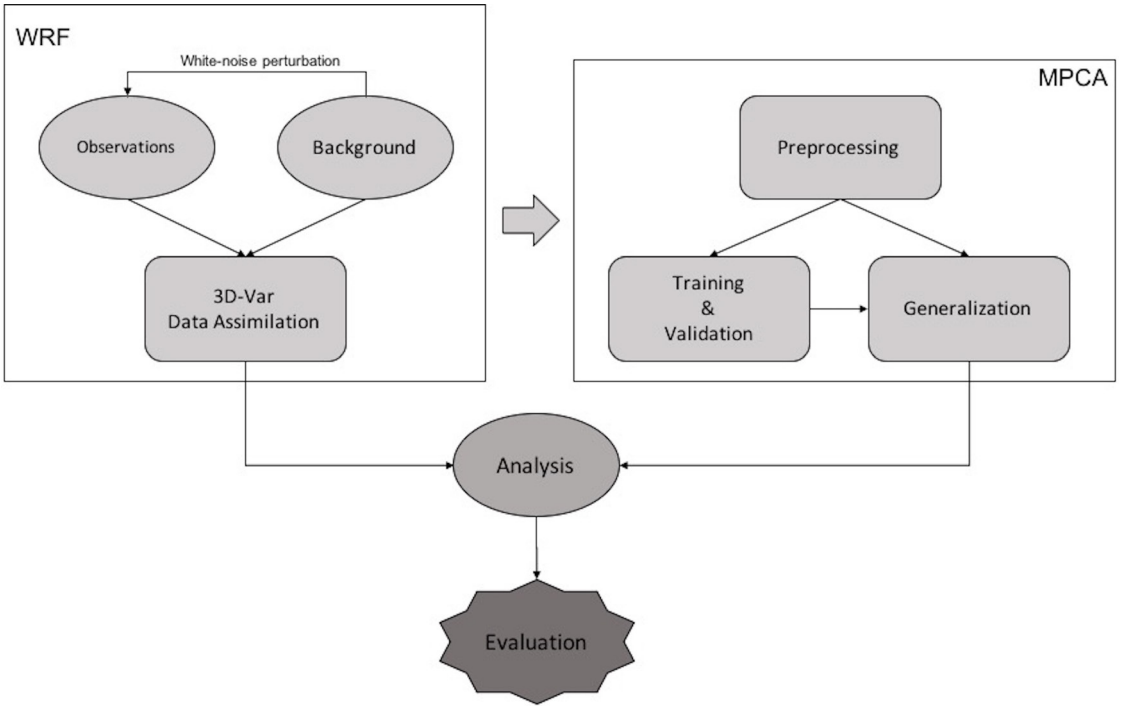

Figure 2 – Flowchart describing the method.

Table 2 presents the statistical values for mean error (ME), standard deviation (STD), root mean square error (RMSE), and Pearson correlation coefficient (CORR), all values computed for each meteorological variable from the testing dataset. The forecasts were determined using the best neural network topologies obtained from the MPCA software – see Table 1. The statistics presented in Table 2 show correlations over 90% to the target variables for the optimal trained neural data assimilation operators and errors smaller than the white-noise perturbation of the synthetic observations been assimilated. The wind variables show higher errors compared to the statistics retrieved for the other variables. It is important to note a statistical performance difference is reported to the vector variable (wind). Figures 3 to 6 present the *quantile–quantile* plot (graphic for comparing two probability distributions) for air temperature, relative humidity, *u* and *v* wind components, respectively, from the testing dataset. This kind of plot is very useful to find bias in the model forecast for specific regions of the variable distribution. Looking at Figures 3 and 4, there is an underestimation tendency for air temperature greater than 35°C), and for relative humidity in the interval [80-100%] slight tendency of relative humidity overestimation for lower percentiles (under 30%). As shown in Table 2, greater differences are found for wind forecasts. Figure 5 shows the *u*-component, where there is an underestimation tendency for values greater than 5 ms^{-1}. For the *v*-component of the wind (Figure 6), there is an underestimation tendency for values greater than 5 ms^{-1}, and a positive bias for all negative values.

The differences observed in the distributions tails are expected since for extreme values there is a greater uncertainty in the observed data.

Table 1 – MPCA results for the training and validation dataset for the meteorological variables. The activation function codes represent logistic (1), tangent (2), and Gaussian (3). The alpha parameter represents the momentum and the eta parameter represents the learning rate.

VARIABLE	PARAMETER	EXPERIMENTS				
		1	2	3	4	5
AIR TEMPERATURE	Best objective function value	0.0924	0.1233	0.0899	0.0924	0.0770
	Number of hidden layers	1	1	1	1	2
	Neurons in hidden layer 1	5	16	15	15	11
	Neurons in hidden layer 2	0	0	0	0	24
	Activation function	1	3	2	2	1
	alpha	0.2545	0.6352	0.3492	0.669	0.5306
	eta	0.0428	0.5617	0.5573	0.2724	0.8033
RELATIVE HUMIDITY	Best objective function value	0.0827	0.0908	0.0891	0.0912	0.0909
	Number of hidden layers	1	2	1	2	1
	Neurons in hidden layer 1	9	6	20	11	9
	Neurons in hidden layer 2	0	14	0	23	0
	Activation function	2	1	1	1	2
	alpha	0.6701	0.3604	0.2851	0.1981	0.0139
	eta	0.8110	0.2713	0.5577	0.0318	0.3069
WIND (*u*-component)	Best objective function value	0.0737	0.0739	0.0751	0.0754	0.0577
	Number of hidden layers	1	1	1	1	1
	Neurons in hidden layer 1	5	5	12	12	16
	Neurons in hidden layer 2	0	0	0	0	0
	Activation function	1	1	2	2	2
	alpha	0.1021	0.2733	0.2722	0.1559	0.5584
	eta	0.4563	0.3299	0.0282	0.0520	0.5277
WIND (*v*-component)	Best objective function value	0.0492	0.0841	0.0424	0.0698	0.0772
	Number of hidden layers	2	2	2	1	1
	Neurons in hidden layer 1	21	14	19	6	10
	Neurons in hidden layer 2	11	7	12	0	0
	Activation function	1	2	1	2	1
	alpha	0.0946	0.7286	0.2388	0.1467	0.3568
	eta	0.8424	0.1335	0.7601	0.7047	0.5013

Table 2 – MPCA statistics for the generalization dataset.

VARIABLE	ME	STD	RMSE	CORR
AIR TEMPERATURE	-0.12 K	0.37 K	0.31 K	0.99
RELATIVE HUMIDITY	1.02 %	4.04 %	3.10 %	0.99
WIND (*u*-component)	-0.19 ms^{-1}	1.05 ms^{-1}	0.63 ms^{-1}	0.98
WIND (*v*-component)	-0.83 ms^{-1}	1.56 ms^{-1}	1.10 ms^{-1}	0.95

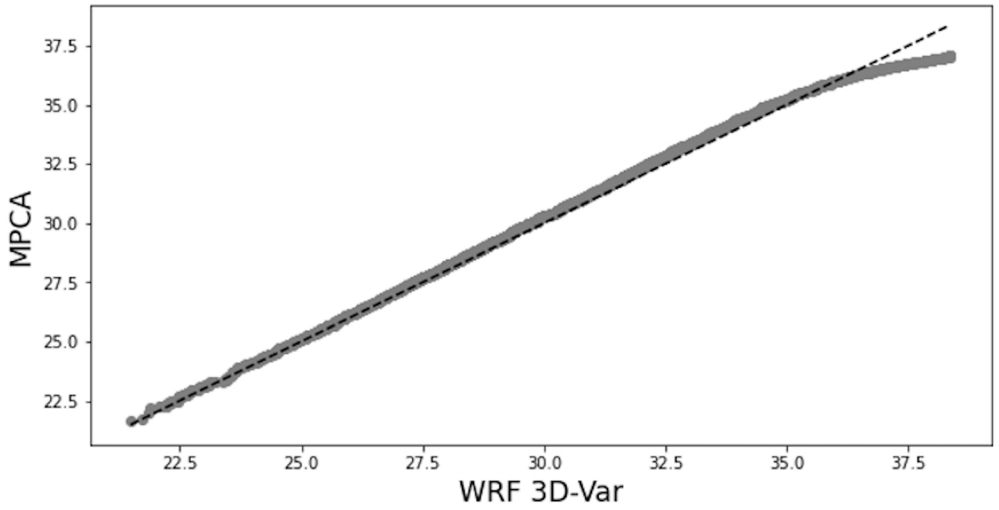

Figure 3 — Quantile-quantile plot comparing the probability distribution of the air temperature analysis generated by the 3D-Var assimilation approach and the MPCA algorithm. The dashed line represents the perfect correspondence (1:1) between the trained neural network and the 3D-Var approach.

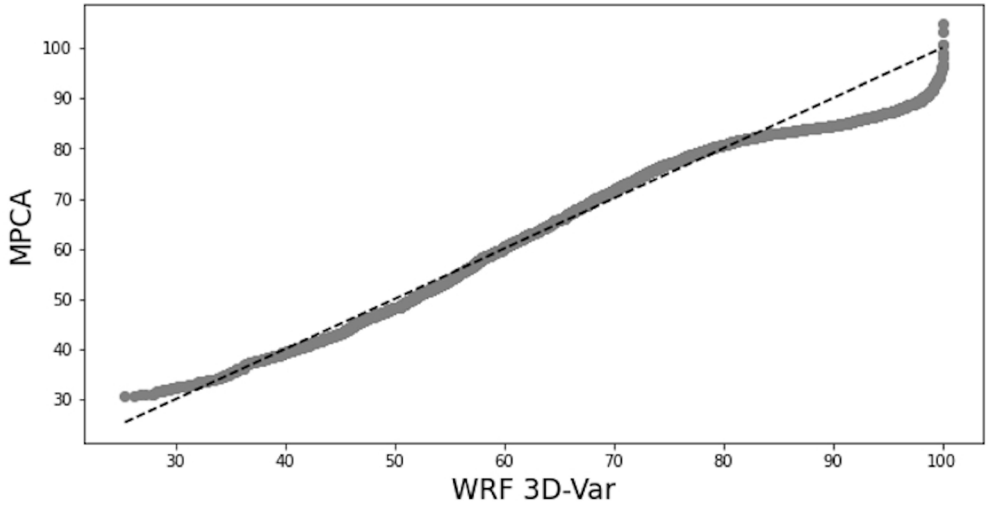

Figure 4 – Quantile-quantile plot comparing the probability distribution of the zonal component of the analysis generated by the 3D-Var assimilation approach and the MPCA algorithm. The dashed line represents the perfect correspondence (1:1) between the trained neural network and the 3D-Var approach.

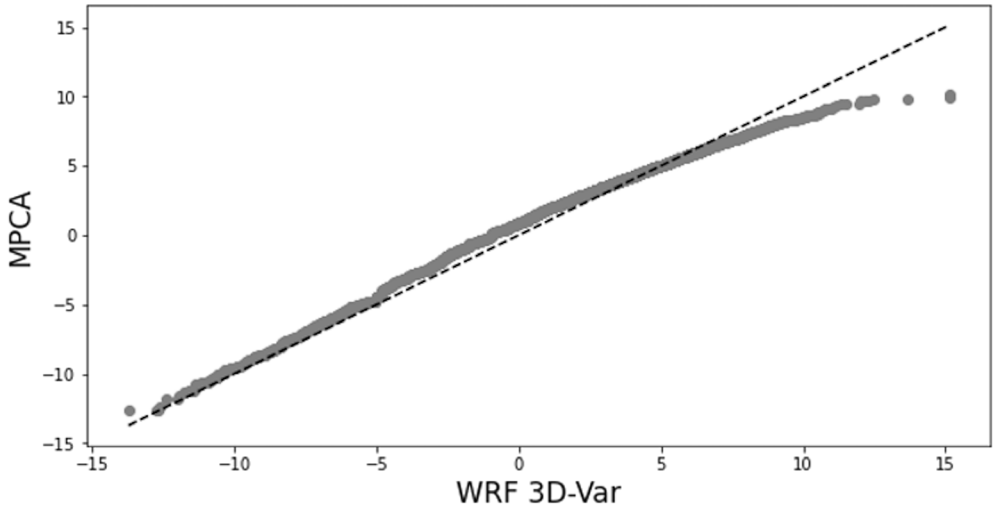

Figure 5 — Quantile-quantile plot comparing the probability distribution of the wind zonal component (*u* component) analysis generated by the 3D-Var assimilation approach and the MPCA algorithm. The dashed line represents the perfect correspondence (1:1) between the trained neural network and the 3D-Var approach.

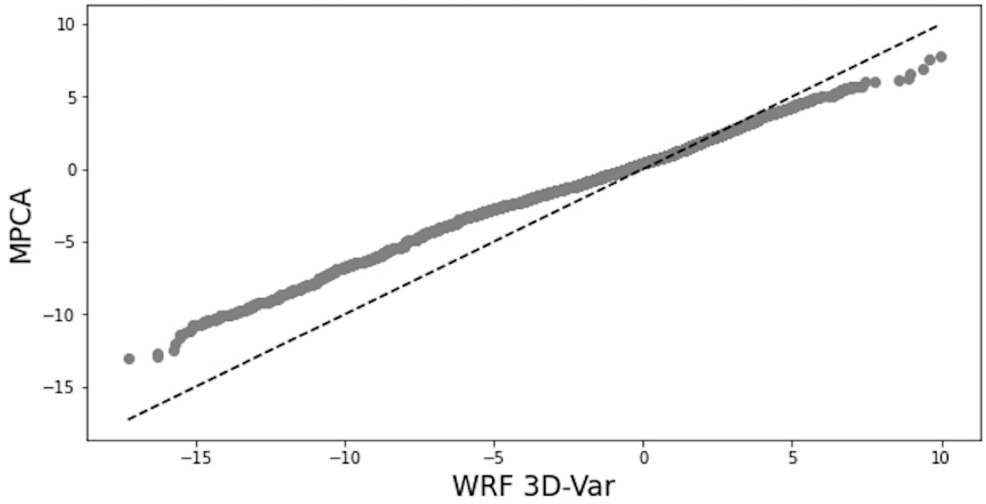

Figure 6 — Quantile-quantile plot comparing the probability distribution of the wind meridional component (*v*-component) analysis generated by the 3D-Var assimilation approach and the MPCA algorithm. The dashed line represents the perfect correspondence (1:1) between the trained neural network and the 3D-Var approach.

Figure 7 shows a case for Feb/01/2014 06Z for the control field (Fig. 7a), the 3D-Var analysis (Fig. 7b), the optimized MLP-NN analysis (Fig. 7c), and the difference between MLP-NN and the 3D-Var analysis (Fig. 7d), considering air temperature at 1000 hPa. Here, the control field is the 6-hour model integration by noiseless initial condition, which is considered as the reference field. Comparing Figures 7b and 7c to Figure 7a it is clear that there is an increase of values in the surroundings of the station locations (red dots). As expected, although the assimilation process removes a great part of the white-noise perturbation on the data, part of it still changes the variable field. Figure 7d represents the root square difference between the assimilation

performed by the 3DVar technique (WRFDA) and the results from the MPCA trained model. The difference between the two processes is under 3 K for all the regions which is around 1% of the magnitude of the assimilated air temperature data. The average execution time of each 3D-Var assimilation cycle was 00:01:11 (1 minute and 11 seconds), while the average execution time of the neural network model was close to 00:00:01 (about 1 second). Therefore, the MLP-NN method was (at least) 71 times faster than 3DVar, under similar hardware conditions, producing very similar quality analysis. Previous results using MLP-NN emulating the analysis from the local ensemble transform Kalman filter (LETKF) has obtained a computational speed-up of 79 and about 54 times faster than the LETKF for the 3D spectral global models Simplified Parameterizations, primitivE-Equation DYnamics (SPEEDY) (Cintra & Campos Velho, 2018) and Center for Ocean-Atmospheric Prediction Studies, Florida State University (COAPS-FSU) (with full physics parameterizations) (Cintra et al., 2018), respectively. We point out the relevance to have an effective and faster technique for data assimilation, allowing to include more observations on a finer model resolution.

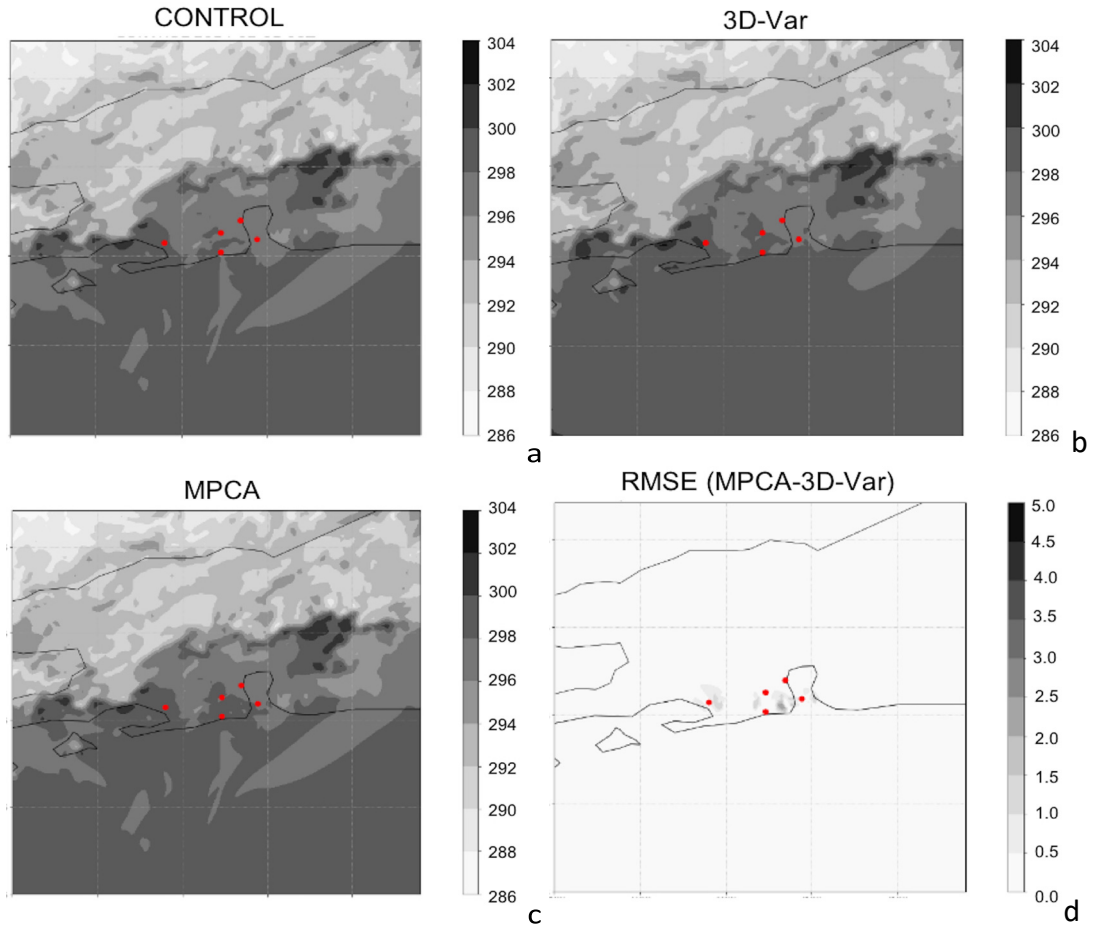

Figure 7 – Air temperature field at Feb/01/2014 06Z for: (a) control; (b) 3D-Var analysis; (c) MLP-NN analysis; and (d) difference between the MLP-NN and 3D-Var analysis. The red dots represent the location of the surface stations where the assimilated data were measured.

CONCLUSION

This work investigated the ability of a multilayer perceptron using a self-configuration strategy by MPCA metaheuristic to emulate the 3D-Var method implemented in the WRF data assimilation framework, for surface and vertical profile assimilated data. The experiments were performed on the terminal area of Rio de Janeiro for the years 2014 and 2015, with synthetic observations generated at five different airport locations.

Our results can be summarized as follows:

1. The results showed correlations over 90% between the two data assimilation techniques (3D-Var and MLP-NN) and errors smaller than the white-noise perturbation of the synthetic observations been assimilated.

2. The greater differences between optimized MLP-NN and 3D-Var were verified for the vector field (wind), in comparison to scalar variables.

3. The neural data assimilation method was 71 times faster than 3D-Var approach.

4. Although only 1-week experiments in 2014 and 2015 were used, our results also show a huge reduction of the CPU-time for the assimilation cycle, as shown in previous results (Cintra & Campos Velho, 2012; Härter & De Campos Velho, 2012). Longer periods of the year will be analyzed in a near future studies as well.

More computational effort than 3D-Var is verified by 4D-Var method. Using the test case for the native variational schemes in a quad-core computer, the 3D-Var demands 45 seconds while 4DVar demanded 9550 seconds (approximately 210 times slower than 3D-Var). Future works will investigate the performance of neural networks face on 4D-Var data assimilation, including hybrid techniques (Wang et al., 2008a; Wang et al., 2008b), with an evolving background error matrix. Other strategies for neural networks will be also studies, such as a multi-objective scheme for neural network training (Anochi et al., 2020) and deep learning approach (LeCun et al., 2015). As a final note, considering the operational centers with a relative short time window to elaborate forecast bulletins, the reduction of CPU time of order 71 times faster than a standard method — the worked example here was the 3DVar scheme – to the assimilation cycle is important for several aspects: possibility of assimilation of a greater amount of data and/or the use of finer model computer resolution.

ACKNOWLEDGMENTS

This study is funded by the Department of Airspace Control (DECEA, Brazilian Air Force, via the Brazilian Organization for Scientific and Technological Development of Airspace Control (CTCEA) (GRANT: 0022018/COPPETEC_CTCEA). Authors would also like to thank the National Council for Scientific and Technological Development (CNPq, Brazil) for the research grants to authors: HFCV (CNPq: 312924/2017-8), and GBF (CNPq: 304441/2018-0).

REFERENCES

Almeida, M. V., 2009. Forecasting short-term of visibility and ceiling for Airport Guarulhos - SP by application of techniques of artificial neural networks. D.Sc. thesis, COPPE/UFRJ, 182 pp., Rio

de Janeiro, RJ, Brazil.

Almeida, V. A., G. B. França, and H. F. de Campos Velho, 2020. Short-range forecasting system for meteorological convective events in Rio de Janeiro using remote sensing of atmospheric discharges. International Journal of Remote Sensing, 41 (11), 4372–4388, doi:10.1080/01431161.2020.1717669.

Anochi, J. A., 2015. Previsão climática de precipitação por redes neurais autoconfiguradas. D.Sc. thesis, Instituto Nacional de Pesquisas Espaciais, 135 pp., São José dos Campos, São Paulo, Brazil.

Anochi,J.A., and H. F. de Campos Velho, 2014. Optimization of feedforward neural network by multiple particle collision algorithm. In: IEEE Symposium on Foundations of Computational Intelligence (FOCI), 128–134, doi: 10.1109/FOCI.2014.7007817.

Anochi, J. A., R. Hernández Torres, and H. F. Campos Velho, 2020. Two geoscience applications by optimal neural network architecture. Pure and Applied Geophysics, 177, 2663–2683, doi:10.1007/s00024-019-02386-y.

Barker D., Huang X-Y, Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., Chen, Y., Demirtaş, M., Guo, Y-R, Henderson, T., Huang, W., Lin, H-C, Michalakes, J., Rizvi, S and Zhang, X., 2012. The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. Bulletin of The American Meteorological Society, 93, 831–843, doi:10.1175/BAMS-D11-00167.1.

Cintra, R. S., and H. F. Campos Velho, 2012. Global data assimilation using artificial neural networks in SPEEDY model. In: Proceedings of the 1st International Symposium on Uncertainty Quantification and Stochastic Modeling, Maresias, São Sebastião-SP, Brazil, p. 648-654.

Cintra, R. S., and H. F. Campos Velho, 2018. Data assimilation by artificial neural networks for an atmospheric general circulation model. Advanced Applications for Artificial Neural Networks, I. EL-SHAHAT, Ed., Intech, chap. 14, 265–286, doi: 10.5772/intechopen.70791.

Cintra, R. S., S. Cocke, and H. F. Campos Velho, 2018. Data assimilation by neural networks with ensemble prediction. In: Proceedings of the International Symposium on Uncertainty Modeling and Analysis, Florianópolis (SC), Brazil, 35-43, Available on: <http://icvramisuma2018.org/cd/web/PDF/ICVRAMISUMA2018-0095.PDF>. Access on: Sep 30th, 2020.

Cintra, R. S. C., 2010. Assimilação de dados com redes neurais artificiais em modelo de circulação geral da atmosfera. D.Sc. thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brazil.

Daley, R., 1991. Atmospheric Data Analysis. Cambridge: Cambridge University Press, 457 pp.

Dudhia, J., 1989. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. Journal of The Atmospheric Sciences, 46, 3077–3107, doi: 10.1175/15200469(1989)046<3077:NSOCOD>2.0.CO;2.

França G.B., Almeida M.V. and Rosette A.C., 2016. An automated nowcasting model of significant instability events in the flight terminal area of Rio de Janeiro, Brazil. Atmospheric Measurement Techniques, 9 (5), 2335–2344, doi: 10.5194/amt-9-2335-2016.

França, G.B., Almeida M.V., Bonnet S.M. and Albuquerque Neto F.L., 2018. Nowcasting model of low wind profile based on neural network using SODAR data at Guarulhos Airport, Brazil. International Journal of Remote Sensing, 39 (8), 2506–2517,doi: 10.1080/01431161.2018.1425562.

Furtado, H. C., 2008. Neural networks and different methods of data assimilation in dynamic nonlinear. M.Sc. thesis, INPE, 125 pp., São José dos Campos,

SP, Brazil.

Grell, G. A., and S. R. Freitas, 2014. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmospheric Chemistry and Physics Discussions, 14, 5233–5250, doi:10.5194/acp-14-5233-2014.

Härter, F. P., and H. F. de Campos Velho, 2012. Data assimilation procedure by recurrent neural network. Engineering Applications of Computational Fluid Mechanics, 6 (2), 224–233, doi:10.1080/19942060.2012.11015417, https://doi.org/10.1080/19942060. 2012.11015417.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. 2nd ed., Prentice-Hall, New Jersy, USA, 842 pp.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004. A revised approach to ice microphysical processes for the bulk parametrization of clouds and precipitation. Monthly Weather Review, 132, 103–120, doi: 10.1175/15200493(2004)132<0103:ARATIM>2.0.CO;2.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Monthly Weather Review, 134, 2318–2341, doi: 10.1175/MWR3199.1.

Hopfield, J. J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences of the United States of America, 79 (8), 2554–2558, doi: 10.1073/pnas.79.8.2554.

Kalnay, E., 2003. Atmospheric modeling, data assimilation and predictability. New York: Cambridge University Press, 341 pp.

Khassenova, Z. T., and A. T. Kussainova, 2018. Applying data assimilation on the urban environment. In: Shokin, Y. and Shaimardanov, Z. (Eds.). Computational and Information Technologies in Science, Engineering and Education, CITech 2018, CCIS 998, Springer, p. 125-134, doi:10.1007/978-3-030-12203-4₁2.

Kirkpatrick, S., C. Gelatt, and M. Vecchi, 1983. Optimization by simulated annealing. Science (New York, N.Y.), 220, 671–80, doi: 10.1126/science.220.4598.671.

LeCun, Y., Y. Bengio, and G. Hinton, 2015. Deep learning. Nature, 521 (7553), 436–444, DOI: 10.1038/nature14539.

Lin, C., T. Matsuo, J. Y. Liu, C. H. Lin, J. D. Huba, H. F. Tsai, and C. Y. Chen, 2017. Data assimilation of ground-based GPS and radio occultation total electron content for global ionospheric specification. Journal of Geophysical Research: Space Physics, 122 (10), 10,876–10,886, DOI: 10.1002/2017JA024185.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. Journal of Geophysical Research: Atmospheres, 102 (D14), 16 663–16 682, doi: 10.1029/97JD00237.

Nowosad, A. G., 2001. New approaches to meteorologic data assimilation. D.Sc. thesis, Instituto Nacional de Pesquisas Espaciais (INPE), 120 pp., São José dos Campos, SP, Brazil.

Pacheco da Luz, E., J. Becceneri, and H.F. Campos Velho, 2008. A new multi-particle collision algorithm for optimization in a high performance environment. Journal of Computational Interdisciplinary Sciences, 1(1), 03–10. Available on: <http://www.epacis.net/jcis/PDFJCIS/JCIS11art.01.pdf>. Access on: Sep 30th, 2020.

Pacheco da Luz, E., J. Becceneri, and H. F. Campos Velho, 2011. Multiple Particle Collision Algorithm Applied to Radiative Transference and Pollutant Localization Inverse Problems. In: IEEE International Symposium on Parallel and Distributed Processing Workshops and PhD Forum. Shanghai, China: IEEE, 342-345, doi: 10.1109/IPDPS.2011.171.

Paulucci T., França G., Libonati R. and Ramos, A., 2019. Long-Term Spatial–Temporal Characterization of Cloud-to-Ground Lightning in the Metropolitan Region of Rio de Janeiro. Pure and Applied Geophysics, 176, 5161–5175, doi: 10.1007/s00024-019-02216-1.

Piazzi, G., G. Thirel, L. Campo, and S. Gabellani, 2018. A particle filter scheme for multivariate data assimilation into a point-scale snowpack model in an Alpine environment. The Cryosphere, 12, 2287–2306, doi: 10.5194/tc-12-2287-2018.

Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65 (6), 386–408, doi: 10.1037/h0042519.

Rumelhart, D. E., G. E. Hinton,and R. J. Williams, 1986. Learning representations by backpropagating errors. Nature, 323 (6088), 533–536, doi: 10.1038/323533a0.

Sacco, W. F., and C. R. E. de Oliveira, 2005. A new stochastic optimization algorithm based on a particle collision metaheuristic. In: 6th World Congresses of Structural and Multidisciplinary Optimization - WCSMO6. Rio de Janeiro, Brazil. International Society for Structural and Multidisciplinary Optimization (ISSMO).

Sacco W., Oliveira C. and Pereira C.M.N.A., 2006: Two stochastic optimization algorithms applied to nuclear reactor core design. Progress in Nuclear Energy, 48, 525–539, doi:10.1016/j.pnucene.2005.10.004.

Sacco W. F., Alves Filho H. and Pereira C.M.N.A., 2007: Cost-based optimization of a nuclear reactor core design: A preliminary model. In: 2007 International Nuclear Atlantic Conference - INAC 2007. Santos, SP, Brazil. 7 pp.

Sacco W., Lapa C., Pereira C.M.N.A. and Alves Filho H., 2008: A metropolis algorithm applied to a nuclear power plant auxiliary feedwater system surveillance tests policy optimization. Progress in Nuclear Energy, 50, 15–21, doi: 10.1016/j.pnucene.2007.09.004.

Silva, W., F. Albuquerque, G. França, and M. Matschinske, 2016: Conceptual model for runway change procedure in Guarulhos International Airport based on SODAR data. The Aeronautical Journal, 120, 725–734, doi:10.1017/aer.2016.33.

Skamarock W.C., Klemp J.B., Dudhia J., Gill D.O., Liu Z., Berner J., Wang W., Powers J.G., Duda M.G., Barker D.M. and Huang X.-Y., 2019: A description of the advanced research WRF version 4. NCAR Technical Notes. Colorado. USA. 145 pp. NCAR/TN-556+STR. doi:10.5065/1dfh-6p97.

Smith P.J., Fowler A.M., and Lawless A.S., 2015: Exploring strategies for coupled 4D-Var data assimilation using an idealised atmosphere–ocean model. Tellus A: Dynamic Meteorology and Oceanography, 67(1), 27025, doi: 10.3402/tellusa.v67.27025.

Talagrand, O., 1997: Assimilation of Observations, an Introduction. Journal of the Meteorological Society of Japan, 75, 191–209.

Tewari M, Chen F, Wang W, Dudhia J, LeMone MA, Mitchell K, Ek M, Gayno G, Wegiel, J and Cuenca R, 2016: Implementation and verification of the united NOAH land surface model in the WRF model. In: 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction. 11-15.

Vrugt J.A., C. D. C., Braak, and G. Schoups, 2018: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. Advances in Water Resources, 51, 457–478, doi:10.1016/j.advwatres.2012.04.002.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A Hybrid ETKF–3DVAR Data Assimilation Scheme for the WRF Model. Part I: Observing System Simulation Experiment. Monthly Weather Review, 136 (12), 5116–5131, doi: 10.1175/2008MWR2444.1.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A Hybrid ETKF–3DVAR Data Assimilation

Scheme for the WRF Model. Part II: Real Observation Experiments. Monthly Weather Review, 136 (12), 5132–5147, doi: 10.1175/2008MWR2445.1.

Wang, Z., A. Storto, N. Pinardi, G. Liu, and H. Wang, 2007: Data assimilation of Argo profiles in a northwestern Pacific model. Natural Hazards and Earth System Sciences, 17, 17–30, doi: 10.5194/nhess-17-17-2017.

Zupanski, M., and E. Kalnay, 1999: Principles of data assimilation. In: Browning KA and Gurney RJ (Eds.). Global energy and water cycles. Cambridge: Cambridge University Press. p. 48-54.

Recebido em 2 de junho de 2020 / Aceito em 21 de julho de 2020
Received on June 2, 2020 / Accepted on July 21, 2020