New learning strategy for supervised neural network: MPCA meta-heuristic approach

Juliana Anochi*, Sabrina Sambatti*, Eduardo Luz* and Haroldo Campos Velho*
*Laboratory for Computing and Applied Mathematics
National Institute for Space Research, São José dos Campos, São Paulo, SP, Brazil
Email: {juliana.anochi, sabrina.sambatti, eduardo.luz, haroldo} @lac.inpe.br
Telephone: 55 (12) 3208–6555, Fax: 55 (12) 3208–6375

***Abstract*—The problem of parameter optimization for a feed-forward artificial neural network (ANN) to determined its best architecture is addressed. A new metaheuristic called Multiple Particle Collision Algorithm (MPCA), introduced by Luz et al. [12], was applied to design an optimum architecture for two models of supervised neural network: the Multilayer Perceptron (MLP), and recurrent Elman network. The NN obtained using this approach is said to be self-configurable. In addition, two strategies are employed for calculating the connection weights to the MLP and Elman networks: MPCA, and backpropagation algorithm. The resulting ANNs were applied to predict the monthly mesoscale climate for the precipitation field. The comparison is performed between the ANN configuration obtained by automatic process and another configuration proposed by a human specialist.**

I. Introduction

The precipitation is considered one of most important and difficult meteorological variable to be predicted. It is linked with heavy floods and extreme events, such as windstorms, landslides, among other occurrences of natural disasters. Conversely, low levels of precipitation may cause to dry, directly affecting the agriculture sector. Considering this complexity, new techniques are used to estimate and predict the precipitation. In this paper, the problem of monthly climate forecast precipitation field is performed using supervised neural networks.

Climate prediction is the forecasting of the state of the atmosphere for a future time in a given location. This is a task that has been pursued by humans, informally and formally, for a longer a time and it is accomplished by acquiring quantitative data about the current state of the atmosphere to be analyzed using scientific understanding of the atmospheric processes in order to project how it will be in the future.

During the last decades, artificial neural networks models have been one of the most tecniques commonly used from the Artificial Intelligence, and nowadays, it is under intense research worldwide. Although there are a lot research in this area, there are still many questions about the ANN models that need to be better treated. One of the main topics on ANN models is the architecture complexity optimal or close optimal for the training process. The process of obtaining an adequate neural network to solve a specific problem is a complex task and usually requires a great effort by the expert mainly to determine the best parameters, and it is necessary a previous knowledge about the problem to be treated. The process for searching and definition of an optimal architecture for an ANN is very relevant, demanding an intensive research about computational efficiency of the model [3].

There are many algorithms in the literature for the ANN training aimed the improving of the ability of generalization and for the control of an adequate architecture specification, such as: Pruning [4], makes adjustment of neural network by modifying its structure, the training begin with an oversized architecture and the weights are eliminated until the capacity of generalization can be increased; Weight Decay suggested by [4], the algorithm is similar to the Pruning, where the cost function and weight vector are modified; Early Stopping proposed by [13], the scheme performs the early interruption of training, without changing the ANN architecture, similarly; Cross Validation proposed by [10], it is known to improve the generalization, where the data set is separated in two sets: training data and validation data. However, all these algorithms still suffer from slow convergence. In addition, these are based on gradient techniques and can easily stick at a local minimum.

The identification of the best configuration for a given neural network could be formulated as an optimization problem.

Teixeira et al. [18] present a new learning algorithm to improve the generalization of the model of multilayer perceptron. This algorithm uses the training techniques of multi-objective optimization, which proposes to control the complexity of ANN using simultaneous minimization of training error and norm of weight vector. Costa et al. [5] presents a new constructive method and pruning approaches to control the design of MLP without loss in performance. Costa et al. [6] also have developed an optimization technique for multiobjective training of ANN which uses the control algorithm for sliding mode control. This algorithm controls the convergence of the system to the point of minimum. Carvalho et al. [3] proposes an approach to configure the architecture of the neural network, using learning algorithms based on optimization techniques mono-objective. The authors used four metaheuristic search: the generalized extremal optimization, the variable neighborhood search, simulated annealing and genetic algorithm.

These different approaches to search and definition of an optimal architecture presented above, cause the classical techniques of estimation becomes unsuitable. Checking the considerations described in this paper proposes the use of the a method based on stochastic optimization techniques the Multiple Particle Collision Algorithm (MPCA) to self-configure the optimal architecture for an ANN. A penalty term is used to evaluate in the objective function to avoid very complex network architectures, where the concept of network complexity is associated to the number of free adjustable parameters (weights) in the network, such as, the number of intermediate layers, the number of neurons in each intermediate layer, the learning rate parameter, among others. The minimization of this function involves the balance between the training error and generalization error.

Optimization problems have the goal of finding the best set within a variable set to maximize or minimize a function, defined as an objective function or cost function. Optimization problems can be classified as: continuous optimization (where the variable has real or continuous values); discrete optimization (where the variable has integer or discrete values); and mixed optimization (with integer and continuous values at the same time) [4].

Additionally, this work proposes a new methodology for developing an empirical model of monthly climate prediction precipitation field, from reanalysis of historical data, the database National Oceanic & Atmospheric Administration (NOAA).

II. Configuring the supervised networks by metaheuristics

Currently, one of the main topics of research for supervised neural network is the search and definition of an architecture optimal or nearly optimal.

Neural networks is a prestigious area of Artificial Intelligence, and have shown their efficiency in various applications in different areas, bringing great benefits. Recently, they have been applied in the climate prediction [1], [7].

There is no a clear indication of how we find the best architecture among the many possible choices. Further, we not even know in advance the correct network topology to be applied. In practice, this problem is usually solved in part by using empirical methods based on repetitive trial and error method. Where the configuration is determined during preliminary tests with different topologies, modifying network parameters, until satisfactory results are obtained. This configuration consists of an iterative process where the specialist changes the values of model parameters for each trial and compared the results with the observed values until you reach a set of parameters for which, in their view, the model results are the most suitable.

The problem to identify the best configuration of the supervised NN can be treated as an optimization problem. The goal is to find the optimum value, which represents the best combination of variables for the NN architecture, and the definition of the set of weights. This self-configured NN is determined as the minimization of cost function defined by Equation 1.

The main advantage in using an automatic procedure to configure an ANN is the ability to define an architecture near-optimal ANN, without needing the help of an experts on the NN approach and/or the application. Such approach avoids this time consuming and tiring process of trial and error to find the optimal neural network architecture.

For the problem of construction of climate prediction model, as proposed in this paper, the networks multilayer perceptron and recurrent Elman are used.

The objective function used in this study is the square difference between the target values and the ANN output. The latter factor is expressed by equation [3]:

$$f_{\text{obj}} = penalty \times \left(\frac{\rho_1 \times E_{\text{train}} + \rho_2 \times E_{\text{gen}}}{\rho_1 + \rho_2} \right) \tag{1}$$

where, $\rho_1 = 1$ e $\rho_2 = 0.1$ the same values proposed by [3]. These are adjustment factors that modify the attributes relevant to the training (E_{train}) and generalization (E_{gen}) errors. There is great flexibility in the evaluation of the objective function, because the training error is directly related to the network memory capacity, and generalization error refers to the ability of ANN to identify the patterns that are similar to patterns used in training. The factor *penalty* is applied to compute the NN with the lowest complexity possible. The computational complexity of a supervised ANN architecture can be defined as the total number of weights and the epochs number present in its structure. To this purpose, a penality factor was developed to favor lightweight architectures, and it can be expressed by:

$$penalty = C_1 \left(\varepsilon^{neurons}\right)^2 \times C_2 \left(epochs\right) + 1 \tag{2}$$

where $C1 = 1$ and $C2 = 0.1$ are adjustment factors that modify the attributes relevant.

The MPCA is employed to evolve: the number of intermediate layers, the number of neurons in each intermediate (hidden) layer, the learning rate parameter η, momentum constant α, and the activation function. Allowed values for these parameters are shown in Table I:

TABLE I. Parameters to define a network architecture.

Parameter	Value
Hidden Layers	| 1 | 2 | 3 |
Neuron in each hidden layer	| 1 | ... | 32 |
Learning ratio: η	| 0.0 | ... | 1.0 |
Momentum constant: α	| 0.1 | ... | 0.9 |
Activation function	| $Tanh$ | $Logistic$ | $Gauss$ |

The goal function 1 is solved by a metaheuristic. This paper presents the use of MPCA for automatic configuration of a neural network architecture, minimizing the cost function, producing a configuration of the NN with the best possible performance.

A. Multiple particle collision algorithm

The Multiple Particle Collision Algorithm was developed by Luz [13], inspired the canonical Particle Collision Algorithm (PCA) [14], [16], [15], [17] but a new characteristic is introduced: the use of several particles, instead of only one particle to act over the search space. In both algorithms are greatly inspired by two physical behaviours, namely absorption and scattering, that occurs inside a nuclear reactor. They are similarities with basic characteristics of Simulated Annealing [25], [26].

The PCA starts with a selection of an initial solution (Old-Config), it is modified by a sthocastic perturbation (Perturbation{.}), leading to the construction of a new solution (New-Config). The new solution is compared (function Fitness{.}) and the new solution can ou cannot be accepted. If the new solution is not accepted, the particle can be send to a different location of the search space, giving the algorithm the capability of escaping a local optimum, this scheme is inspired on the scaterring (Scaterring{.}). If a new solution is better than the previous one, this new solution is absorbed (Absorption.}). The exploration around closer positions is guaranteed by using the functions Perturbation{.} and Small-Perturbation{.} [12], [13].

The implementation of the MPCA algorithm is similar to PCA, but it uses a set with n particles, where a mechanism to share the particles information is necessary. A blackboard strategy is adopted, where the Best-Fitness information is shared among all particles in the process. Luz et al. [12] have showed which the MPCA is able to computing good solutions were which PCA cannot found a correct answer.

The MPCA is intended to be implemented using Message Passing Interface (MPI) libraries in a multiprocessor architecture with distributed memory.

The pseudo-code for the MPCA is presented by Table II [13].

TABLE II. PSEUDO-CODE FOR MPCA.

```
Generate an initial solution: Old-Config
Best-Fitness = Fitness{Old-Config}
Update Blackboard
For n = 0 to # of particles
    For n = 0 to # iterations
    Update Blackboard
    Perturbation{.}
        If Fitness{New-Config} > Fitness{Old-Config}
            If Fitness{New-Config} > Best-Fitness
                Best-Fitness = Fitness{New-Config}
            End If
            Old-Config = New-Config
            Exploration{.}
        Else
            Scattering{.}
        End If
    End For
End For
```

The MPCA, as well as the PCA, is based on the methodology presented by Metropolis et al. [25], which defines an algorithm used to generate a sequence of samples from a probability distribution function.

The MPCA is part of a class of algorithms inspired by natural evolutionary algorithms (EA) [24]. While the PCA is a trajectory-based method, the MPCA is a population-based method, that starts from a set of initial solutions (initial population) and try to find a better solution by evolving elements of this population.

III. New learning strategic for networks with supervised learning

A. Learning strategic by backpropagation

The error backpropagation algorithm is used as training for networks with supervised learning multilayer perceptron and Elman.

The property of primary significance for a neural network supervised is the ability to learn from the environment, improving its performance through the learning. A neural network learns from its environment through an interactive process of adjustments applied to its synaptic weights and bias levels [8]. In the context of neural networks, the learning process can be defined as a set of well defined rules for solving a specific problem of learning.

The backpropagation algorithm is widely used as training algorithm for the models with supervised learning. The first one is the forward phase where the activations are propagated from the input to the output layer, and it's effect propagates on the entire the network, layer by layer. In the next phase, backward phase, the synaptic weights are all adjusted in accordance with an error correction rule. A set of outputs is produced as the actual response of the network. During the backward phase, the synaptic weights are all adjusted in accordance with an error correction rule. The actual value is subtracted from a desired response to produce an error signal. This way, the error signal is backward neuron through the network, against the direction of synaptic connections [8].

To minimize the error, a correction is applied $\Delta w_j(n)$ weight synaptic in accordance with rule delta [8]:

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \tag{3}$$

Where η is the learning rate parameter of the error backpropagation algorithm. The use of the negative sign 3 denotes the gradient descent in the space of weights, so as to reduce the value of $\varepsilon(n)$.

B. Learning strategic by MPCA

The training process consists to find the connection weights for minimazing the objective function, and it is expressed as square difference between the target values and the network output – which depends on the weight values. The objective function is given by:

$$J(w) = \left[y_k^{\text{ann}}(w) - y_k^{\text{obs}}\right]^2 \tag{4}$$

where y_k^{ann} and y_k^{obs} are the network out and target value for the artificial neuron k, respectively, and w is the unkown connection weights.

The initial guest to estimate the optimal set of weights w^* are randomly chosen. The MPCA is a stochastic optimization procedure. Thefore, several realizations are performed. For the present application, 10 realization are executed with MPCA. The same parameters to set up the MPCA are used to identify the best NN configuration, as to calculate the connection weights.

IV. Application: monthly climate precipitation prediction

Climate prediction is the estimation of the average behavior of the atmosphere for a future period of time (more than one month ahead). For instance, in making a seasonal climate forecast, one may evaluate if the next season (e.g. winter) will be colder (or more than warm) than the climatological average, or else, if there will be more (or less) rain fall than in the previous season. Thus, the objective of climate prediction is to estimate the statistical properties of the climate in a future period of time [22].

Climate prediction centers conduct climate prediction using models that try to describe the behavior of the physical-chemical conditions of the atmosphere. Such models require high performance computational resources to generate possible future status of the atmosphere in high resolution scales.

Thus, the empirical prediction model developed in this study provides the development of future scenarios that support the studies of impacts and vulnerability and can still enable the preparation of projections of climate extremes of atmospheric state.

A. Description of data

The data was downloaded from the reanalysis data repository from National Center for Environmental Prediction & National Center for Atmospheric Research (NCEP/NCAR). The world is divided into parallels (latitudes) and meridians (longitude). The global data reanalysis grid uses a set with horizontal resolution of $2.5^{o} \times 2.5^{o}$ (latitude $\times$ longitude) [10].

The data consists of monthly means from January 2005 to January 2013. The analysis region comprises a selected subregion with coordinates (Lat 10S, 0) to (Lon 47W, 40W) as illustrated in Figure 1. The available variables are: zonal wind components at vertical levels 300 hPa, 500 hPa, meridional wind components at vertical levels 300 hPa, 500 hPa, and precipitation.

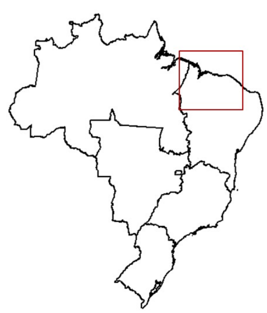

Fig. 1. The analysis region comprises a northeast subregion

During the training phase, the capacity of generalization is assessed using a set of validation data. The training data are divided into a subset of examples and a subset of validation. The subset of examples is used to train the network, the training session is interrupted periodically after a certain number of times, and the network is tested with the subset of validation after each training period. When the validation error tends to increase, characterized generalization loss and training is terminated.

The data set was the training data set was formed with data from January 2005 up to December 2010. This set was used to derive a neural network model. From January to December 2011 was used to derive the generalization data set. The test (validation) set corresponds to the period data from January 2012 up to January 2013.

B. Results for the prediction of precipitation

As the MPCA algorithm is a stochastic method for solution of optimization problems, we conducted 10 experiments with seeds generate different random numbers and experimental data generated artificially. The parameters used are: 6 particles; 6 processors; 500 iterations. The stopping criterion used was the maximum number of evaluations of the objective function.

The MPCA is used to generate a set of candidate solutions that correspond to an ANN architecture. For each solution, the ANN is activated, and the training process starts until the stopping criterion is satisfied (error minimum or total epochs). With the values obtained by ANN, the MPCA calculates the objective function (Equation 1 and 2), up dating the parameters for the ANN. This process is repeated until an optimal value (minimum) for the objective function is found.

Table III shows the settings obtained by using the optimization algorithm of the following networks: MLP-MPCA and Elman-MPCA, had their topologies configured using the MPCA algorithm. The MPCA-learning refers to the architecture obtained with training based on metaheuristics, that is, the backpropagation learning algorithm was replaced by MPCA; and the networks models MLP-expert and Elman-expert were defined empirically:

TABLE III. CONFIGURATION ARCHITECTURE OF THE SUPERVISED NEURAL NETWORKS: MLP AND ELMAN

Parameters	Self-configuring	Learning by MPCA	Configured by expert
intermediate layer	2	2	1
neuron in first layer	9	9	6
neuron in second layer	8	8	0
neuron in third layer	0	0	0
learning rate η	0.53	0.53	0.4
constant momentum α	0.20	0.20	0.6
activation function	$Tangent$	$Tangent$	$Logistics$

C. Results for the first experiment

In the experiment first ANN used backpropagation algorithm where topology parameters were defined empirically by an expert [1].

Among the various outputs provided by the models of weather and climate, precipitation is one of the variables of interest to society. In this paper we deal with an optimization problem for automatic configuration of the architecture of an ANN with minimal cost and to present best possible performance.

Figure 2 shows the results obtained with the architecture configuration determined by empirical realization, where performed by an expert. As can be seen, the results are in a good agreement.

The configuration of networks supervised by specialist was performed based by testing, where the parameters were modified to obtain a good result.

D. Results for the second experiment

In the experiment second, the MPCA was used for self-configuring the parameters of the networks MLP and Elman using the algorithm backpropagation.

In Figure presents the results obtained in the prediction process, using MLP and Elman networks with their topologies defined using the MPCA with training based in metaheuristics. The advantage of using of the procedure of an automatic method to identify a good configuration for the ANN does not require the supervision of an expert.

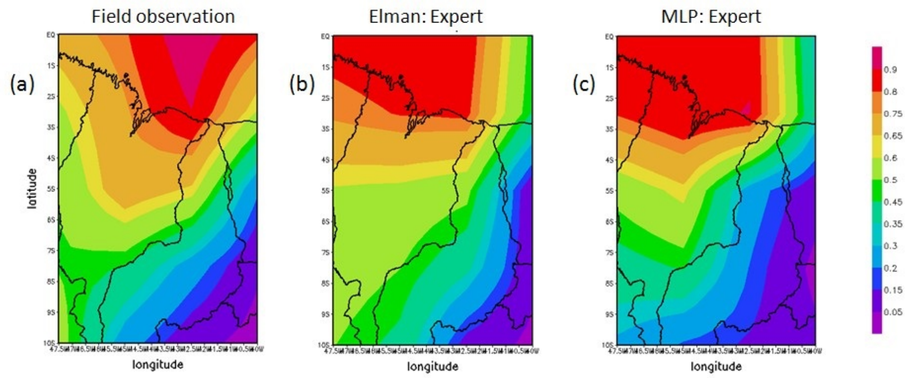

Fig. 2. (a) Field the observated precipitation in April 2012 in a subarea of Northeast; (b) Result for prediction the april using network Elman which was set up by the expert; (c) Result of the prediction using the model MLP set up by expert.

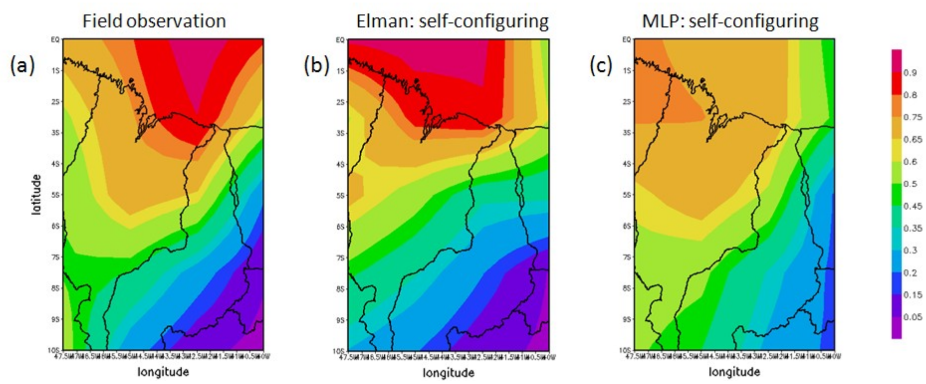

Fig. 3. (a) Field the observated precipitation in April 2012; (b) Result for prediction the april using network Elman self-configuring by MPCA; (c) Result of the prediction using the model MLP self-configuring by MPCA.

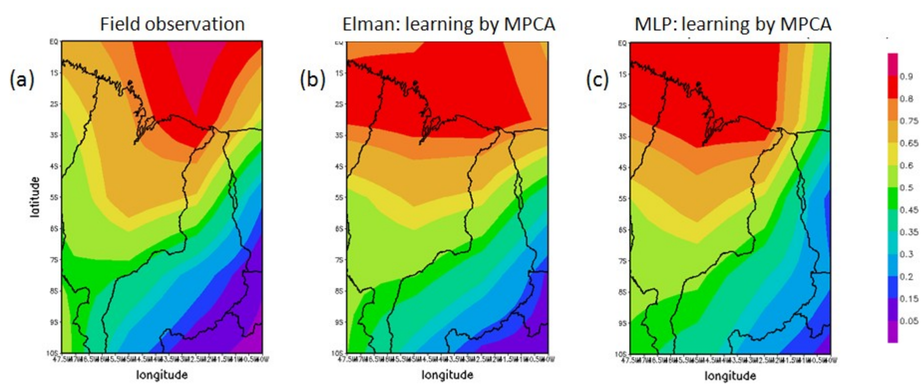

Fig. 4. (a)Field the observated precipitation in April 2012; (b) Elman: learning by MPCA; (c) MLP: learning by MPCA.

E. Results for the third experiment

In the experiment third, the learning strategic of the networks MLP and Elman were based in the metaheuristic, where the set of weights was defined using the optimization algorithm MPCA.

Figure 4 shows the results obtained with the new learning strategic of the networks, where the training the neural networks were based on metaheuristics, ie, weights were defined by using the optimization MPCA algorithm.

V. Final Remarks

The supervised networks were applied to monthly climate prediction of precipitation field. Climate prediction is an important issue with strong impact for the society, in particular the precipitation field, one of more difficult meteorolical variable to be predicted, due to its large variability in space and time. This is a task that has been pursued by humans, for a longer a time.

Anochi and Silva [1], [2] have applied ANN for sasonal climate prediction for precipitation field. The authors demonstrated the efficacy of the proposed method, with estimates of predicted similar to climatological conditions considered as available observations in the database.

However, the configuration of a supervised ANN is not a easy task and usually requires a great effort by the expert mainly to determine the best parameters, and it is necessary a previous knowledge about the problem to be treated. Many real-world problem involves the optimization of several incommensurable and often conflicting objectives. For this reason, optimization the parameters of ANN has become an important topic and, very challenging for researchers from several applied sciences

The problem to identify the best configuration of a supervised network to the application cited above is formulated as an optimization problem. The stochastic technique MPCA was employed to address the solution of the optimization problem.

The MPCA was also employed as a learning strategy, that is, the backpropagation training algorithm has been replaced by the MPCA, where the weights were defined by using the optimization algorithm MPCA.

The automatic method to identify a configuration for the ANN does not require the help of an expert. This special feature allows the application of ANN technique for a larger community. Additionally, some solution could be found that could never be tested by a human being.

Initial tests using the reanalysis data NCEP reveals great potential for the model presented here. Other databases will be used in the experimental part to ensure the competitiveness of the ANN self-configurable by MPCA.

ACKNOWLEDGMENT

The authors would like to thank CNPq and CAPES for the financial support.

REFERENCES

[1] Anochi JA (2010) Models based on neural networks for the study of seasonal climatic patterns from data processed with the Rough Set Theory. M.Sc. dissertation on Applied Computing – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.

[2] Anochi JA, Silva JD (2009) Using artificial neural networks and rough set theory in the study of seasonal weather patterns. Learning and Nonlinear Models, 7:83-91.

[3] Carvalho AR, Chaves AA, Ramos FM (2011) Metaheuristics for the feedforward artificial neural network (ANN) architecture optimization problem. Neural Comput & Applic, 20:1273-1284.

[4] Becceneri JC. 2008. Chapter Meta-Heurísticas e Otimização Combinatória: Aplicações em Problemas Ambientais. INPE, São José dos Campos.

[5] Costa MA, Braga AP, Menezes BR (2003) Improving neural networks generalization with new constructive and pruning methods, Intelligent and Fuzzy Systems, 13:75-83.

[6] Costa MA, Braga AP, Menezes BR, Texeira RA, Parma GG (2003) Training neural networks with a multi-objective sliding mode control algorithm, Neurocomputing, 51, 467-476.

[7] Gardner MW, Dorling SR (1998) Artificial neural networks (the multilayer perceptron) – A review of applications in the atmospheric sciences. Elsevier Science, 2627-2636p.

[8] Haykin S (1999) Neural Networks: A Comprehensive Foundation. Prentice-Hall Inc., New York, 842p.

[9] Hinton GE (1989) Connectionist Learning Procedures. Artificial Intelligence, v.40, 185-234p.

[10] Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, Branco G, l J, Zhu Y, Leetmaa A, Reynolds B, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society, v.77, 437-472p.

[11] Luz EFP, Becceneri JC, Velho HFC (2008) Uma nova meta-heurísticas baseada em algoritmo de colisão de multiplas partículas. Anais do XI Simpósio de Pesquisa Operacional e Logística da Marinha. Rio de Janeiro: CASNAV.

[12] Luz EFP, Becceneri JC, Velho HFC (2008) A new multi-particle collision algorithm for otimization in a high-performance environment. Journal of Computacional Interdisciplinary Sciences, v.1, 1-7p.

[13] Luz, EFP (2012) Meta-heurísticas paralelas na solução de problemas inversos. D.Sc. thesis on Applied Computing – Instituto Nacional de Pesquisas Espaciais.

[14] Sacco WF, Oliveira CREA (2005) A new stochastic optimization algorithm based on a particle collision metaheuristic. WCSMO, 6th World Congress of Structural and Multidisciplinary Optimization, Rio de Janeiro.

[15] Sacco WF, Filho HA, Pereira CMNA (2007) Cost-based optimization of a nuclear reactor core design: a preliminary model. Proceedings of the International Nuclear Atlantic Conference, ABEN, Santos.

[16] Sacco WF, Lapa CMF, Pereira CMNA, Filho HAA (2006) Two stochastic optimization algorithms applied to nuclear reactor core design. Progress in Nuclear Energy, v.48, p.525-539.

[17] Sacco WF, Lapa CMF, Pereira CMNA, Filho HAA (2008) A metropolis algorithm applied to a nuclear power plant auxiliary feedwater system surveillance tests policy optimization. Progress in Nuclear Energy, v.50, p.15-21.

[18] Texeira RA, Braga AP, and Takahashi RHC, Saldanha RR (2000) Improving generalization of MLP with multi-objective optimization. Neurocomputing, v. 35, 189-194.

[19] Loghmanian SMR, Jamaluddin HA, Yusof R, Khalid M (2012) Structure optimization of neural network for dynamic system modeling using multi-objective genetic algorithm. Neural Comput & Applic, 21: 1281-1295.

[20] Stone M (1978) Cross-validation: A review. Mathematische Operationforsch Statischen, v.9, 127-139p.

[21] Tubelis A, Nascimento FJLD. Meteorologia descritiva fundamentos e aplicações brasileiras, Nobel, 1992 São Paulo, 1992, ed. 1.

[22] Vianello RL, Alves AR. Meteorologia básica e aplicações. UFV Universidade Federal de Viçosa, 2006, 449p, Viçosa, ed. 4.

[23] Weigend AS, Huberman B (1990) Predicting the Future: a Connectionist Approach. International Journal of Neural Systems, v.1, 193-209.

[24] Eiben AE, Smith JE, Introduction to Evolutionary Computing. Springer-Verlag, Berlin, 2003.

[25] Metropolis N, Resenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. The Journal of Chemical Physics, AIP, v.21, n.6, 1087-1092, 1953.

[26] Kirkpatrick S, Gelatt CD, and Vecchi MP. Optimization by Simulated Annealing. Science, New Series, (1983) 671-680.