

Optimization by Firefly with Predation for Ensemble Precipitation Estimation Using BRAMS

Eduardo F. P. da Luz¹, Ariane F. dos Santos¹, Saulo R. Freitas¹,
Haroldo F. de Campos Velho^{1,*}, Georg A. Grell²

¹Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos (SP), Brazil

²National Oceanic and Atmospheric Administration (NOAA), Boulder (CO), USA

Abstract The precipitation is predicted by using a weighted combination of different convective parameterization schemes. The best prediction estimation is obtained by calculating the appropriated weights for the ensemble of parameterizations. The weight identification process is a type of inverse problem: parameter estimation. The inverse solution is computed by minimizing the functional of square difference between observations (measured precipitation) and precipitation calculated by a mathematical model (BRAMS model). The optimization problem is solved by a new meta-heuristic: Firefly algorithm with predation (FAP), where the best set for ensemble weights is found. The method is applied to the BRAMS (Brazilian developments on the Regional Atmospheric Model System) over South America during January 2006.

Keywords Parameterizations for precipitation, Weighted ensemble, Inverse problem, Optimization, Firefly with predation meta-heuristic

1. Introduction

In 2005 and 2010, large domains of the Amazon region experienced.

The numerical weather and climate forecasting is a significant scientific advance from the century XX. However, many aspects are waiting to have a better representation in the meteorological computer codes. One of these issues is the quantitative precipitation forecasting – in fact, it is a permanent challenging. The rainfall prediction is a final result of a set of physical processes in the atmosphere dynamics. However, such processes are not well represented in the numerical models, in particular during summer period in the tropical/equatorial zones. Therefore, the precipitation process is parameterized, and the approaches differ on several aspects. In general, operational weather centres deal with two types of atmospheric dynamics models: global and regional (limited area) models. Regional models have a finer spatial resolution than global atmospheric models, and the location and amounts of precipitation should be better represented in limited area models. As already mentioned, the convection process is not correctly described in the models, and errors are verified for the latter process corresponding to a correct amount, location and time [1].

The convection parameterization is a development to express a relation between the intensity of the subgrid-scale convective activity and the large-scale variables. In particular, the precipitation forecasting depends on the assumptions employed for parameterization schemes for convection, cloud turbulence, and other features [2]. Even with the progress on this modelling, the convection process remains unsolved. The cumulus parameterization, for instance, is one more difficult question to be addressed into numerical prediction. Hence, cumulus parameterization is one of the most difficult problems in meteorological models. Basically, the parameterization is a relation among assumptions to describe a process (how the phenomena start, and the iteration between the environment and the phenomena) with the associated forecast variables. Parameterizations are constitutive frameworks to be applied into a mathematical model used to simulate a physical process, here, atmospheric dynamics.

The forecasting fails mainly due to the uncertainties in the model, and the uncertainties are everywhere: topography representation, type of soil and soil moisture mapping, type of soil covering, iteration surface-atmosphere, dynamical forcing (ocean circulation, for example), numerical values of model parameters, processes under-represented, and numerical methods. One way to quantify the uncertainty is to employ the *ensemble prediction*, for computing a confidence interval. The ensemble members can be obtained using different strategies: with a set of different initial conditions, different numerical values for model parameters, or by

* Corresponding author:

haroldo@lac.inpe.br (Haroldo F. de Campos Velho)

Published online at <http://journal.sapub.org/ajee>

Copyright © 2015 Scientific & Academic Publishing. All Rights Reserved

employing different models [3].

The application of different physical parameterizations is one technique to develop a multi-model ensemble [1, 4]. Two ensemble strategies were compared by Stensrud *et al.* [1]: (i) a multi-model ensemble (five different convective parameterization schemes) with identical initial conditions, (ii) an ensemble with different initial conditions and the same model configurations. Better results were obtained with strategy-(i) or strategy-(ii) depending on the applied conditions [1]. From the results, the model change approach is an option as a method to identify the members of ensemble.

The method named Multimodel-Superensemble [5-7] is an ensemble-based approach, where the response from several models is combined, and a weighting average is calculated. Cane and Milelli [8] applied the method to predict the precipitation on the Piemonte (Italy) region. They used the outputs from the ECMWF (European Centre for Medium-Range Weather Forecasts) model and different versions from the COSMO (Consortium for Small-scale Modeling) model. The least-square minimization was employed to estimate the weights for the ensemble average. The estimated weights were applied during the forecast time integration. The precipitation forecasting was improved.

Grell and Dévényi [9], for simplicity GD, developed a scheme based on an ensemble of models for convection/precipitation. Several parameterizations are the ensemble members to the GD scheme, where an ensemble mean (EN) is computed. One improvement for the GD technique is to use a weighted average of precipitation forecasts from the several parameterizations [10, 11]. Each weight in this average represents a new unknown, characterizing an inverse problem for parameter identification.

The GD approach is applied to predict the precipitation field over South America using the BRAMS model [12]. The ensemble weights are identified by minimizing the square difference between combined forecasting from several convection parameterizations, applied in the BRAMS model, and the observations. The optimization problem is solved with the Firefly algorithm (FA) [13]. A new version of the FA formulation is employed here: the FA with predation (FA-P).

The next section is a brief description of the BRAMS model. Section 3 presents the FA meta-heuristic, and Section 4 draws the numerical experiment – showing the results. In the last section, final remarks are addressed.

2. The BRAMS Model

The BRAMS code [12] is supported and developed by the CPTEC-INPE (CPTEC: Center for Weather Prediction and Climate Study, INPE: National Institute for Space Research), and it is based on the Regional Atmospheric Modeling System (RAMS) [23], with new functionalities, parameterizations, and a very efficient parallel version – it is

able to run on massively parallel machines, with gain of performance, even considering 10000 processing cores. This mesoscale meteorological code can simulate several atmospheric processes at many scales. BRAMS is the operational on limited area code used by the CPTEC-INPE for producing the numerical prediction on South America. Our simulations were carried out with BRAMS version 4.3 (see: <http://brams.cptec.inpe.br/>), but the CPTEC-INPE employs the version 5 for the operational forecasting. In our experiments, the horizontal resolution was 25 km, with 38 levels for the vertical coordinate, where the first level placed at 100 m over the surface.

The BRAMS model includes, among other parameterizations, an ensemble version of a deep and shallow cumulus scheme based on the mass flux approach (GD). The convective parameterization trigger function uses the turbulence kinetic energy (TKE) from the Planetary Boundary Layer (PBL) parameterization. The model can run with five popular closure parameterizations: moisture convergence (MC: Kuo, 1974) [25], low-level Omega (LO: Frank-Cohen, 1987) [26], Kain-Fritsch (KF) [27], Arakawa-Schubert (AS) [28], and Grell (GR) [29]. Of course, the user can select only one closure option, choosing among the five available options. However, these five parameterizations can be combined to estimate the precipitation. The convective parameterization employed was the GD approach, where the weighted ensemble was constituted by a GR, LO, MC, KF, and AS schemes.

The BRAMS model was run with 5 parameterization schemes separately, obtaining 5 different simulations. Each simulation is a member of the forward problem (BRAMS model). The final predicted precipitation is the weighted ensemble following the GD approach. The best estimation for the weights used to estimate the precipitation filed is obtained by an inverse analysis, described in the next Section.

3. Precipitation Field: Inverse Analysis

The precipitation field estimation is computed using the Grell and Dévényi multi-model ensemble approach [9]. In this technique, the net effect from the feed-back cycle (environment-cloud) is obtained from statistical method to compute the optimal values to the set of parameters to express the precipitation in the model. The GD parameterization here is applied using five different closure precipitation schemes (the ensemble members) together.

In the procedure of the DG scheme, some parameters are selected (radiation, surface flux, and turbulent diffusion fluxes in the Planetary Boundary Layer – PBL) to calculate the total forcing term for all grid points, and applied to each precipitation parameterization. Total forcing represents the budget from the sub-grid scale forcing (radiation, surface flux, and turbulent flow in the PBL) added to the large scale terms (advection on the grid scale). Each parameterization has its threshold for starting the precipitation process. During

the latter process, the mass flux and the precipitation are calculated according to parameterization from the ensemble members. These five precipitation predictions are combined to have a numerical forecasting for precipitation and the atmospheric ratios for the heat and moisture fluxes. For the current BRAMS version, there are some options to be defined by the user:

1. Mass flux for a given ensemble member- i : $(m_b)_i$,
2. Ensemble average (ENS):

$$\langle m_b \rangle = (1/N) \sum_{i=1}^N (m_b)_i \quad (3)$$

The identification of weight values for a linear combination of the five different parameterizations is formulated as an optimization problem. The smoothest precipitation field is found by adding to the objective function a regularization operator with the help of a Lagrange multiplier:

$$J(W) = \sum_{i=1}^N [P_0 - w_i (P_W^{\text{Mod}})_i]^2 + \eta \Lambda [P_W^{\text{Mod}}] \quad (4)$$

where the vector $W = [w_1 \ w_2 \ \dots \ w_N]^T$ represents the unknown weights with $N = 5$, $(P_W^{\text{Mod}})_i$ is the simulated precipitation according the i -th parameterization, $\Lambda[\cdot]$ is the regularization operator, and η is the regularization parameter (Lagrange multiplier). The mass flux is also computed from a weighted linear combination with the mass flux associated at each parameterization. Hence, the equation for the mass flux is expressed by

$$m_s = w_{AS} m_{AS} + w_{GR} m_{GR} + w_{KF} m_{KF} + w_{LO} m_{LO} + w_{MC} m_{MC} \quad (4)$$

here m_s is the mass flux on the basis of the cloud, m_k and w_k are the mass flux and the weight associated to each parameterization, respectively, with $k = AS, GR, KF, LO, MC$.

4. Firefly Algorithm with Predation

The Firefly Algorithm (FA) belongs to the class of meta-heuristic methods to compute a solution for an optimization problem. A heuristic¹ is a procedure for computing a solution under a finite number of steps, where a proof of convergence for an optimal (or global) solution is not usually described. The prefix *meta* means *after*, indicating an upper level of discovering. The expression meta-heuristic was introduced by Glover to denote a general searching strategy [14], and it is not specific for a certain problem. Such strategies use a mechanism to avoid falling into a local minima or maximum.

Xin-She Yang has developed the FA [13] based on the bioluminescence process, a marvel characteristic of fireflies. Preliminary studies with this methodology have produced satisfactory results on important applications [10, 11, 13, 15]. FA has many similarities with other meta-heuristics, such as Particle Swarm Optimization (PSO) [16] – see

Artificial Bee Colony optimization (ABC), information available in the webpage: <http://mf.erciyes.edu.tr/abc/>. Some FA advantages are the simplicity of its implementation [17–19], and the efficiency faced on other algorithms, for instance Genetic Algorithms [20] and Simulated Annealing [21], as expressed by Luz [22].

There is no full understanding of the function of the flashing lights of fireflies [13]. However, the experts have pointed out at least two important issues associated with the flashes: (a) attracting mating partners, and (b) attracting potential prey. For the FA implementation, Yang [13] has considered: (i) fireflies attracts each other; (ii) the attractiveness is proportional to their brightness, and both decrease as their distance increases – if there is no brighter firefly, it will move randomly; (iii) the brightness of a firefly is affected or determined by the landscape of the objective function. The attractiveness is determined by the brightness, for practical issue it is associated with the objective function.

The light intensity variation and the attractiveness formulation are two important aspects to the FA. The attractiveness is determined by the brightness, where the latter is associated with the objective function. The light intensity is exponentially decreasing with the distance. The firefly position represents a candidate solution. The new position ($m+1$) of a firefly (x_i) is computed as following (there is a firefly population under consideration):

$$x_i^{m+1} = x_i^m + \beta_0 e^{-\gamma r_{ij}^2} (x_i^m - x_j^m) + \alpha (\text{rand} - 1/2) \quad (1)$$

where m denotes the interaction step, the parameter α is the degree of influence for the stochastic forcing, and rand is a random number with uniform distribution on $[0, 1]$. The second term in RHS in Eq. (1) is the attractiveness, representing the decay of light intensity seen by other fireflies, with β_0 the attractiveness at $r = 0$, and γ represents the absorption light coefficient for the medium. Parameters α , β_0 , and γ are free ones to the algorithm. The distance between two fireflies is calculated in a standard way:

$$r_{ij} \equiv |x_i - x_j| = \sqrt{\sum_k (x_{k,i} - x_{k,j})^2} \quad (2)$$

$$i, j = 1, 2, \dots, N \quad \text{and} \quad k = 1, 2, \dots, N_p$$

with $x_{i,k}$ being the k -th component of the vector x_i – the i -th firefly, a vector containing the parameters to be identified. The set $\{x_i\}_{i=1}^N$ is the firefly population.

4.1. Predation Operator

A new predation operator was developed to the FA. The idea is to emulate the natural world, where predation action is present in the nature. Predators, in general, seek easier prey, that is, the youngest, slow, and elements with diseases. In analogy with the natural world, the solutions in the population (firefly) further away from the optimality condition are eliminated, during the predation cycle. The resulting population is rebuilt with new elements, randomly generated, preserving the best elements from the earlier population. The new version of the firefly algorithm with

¹ Heuristic: a Greek word meaning finding, discovery.

predation is shown in Figure 1.

```

Define the cost function:  $J(x^i)$ , with  $x^i = [x_1, x_2, \dots, x_d]^T$ 
Define parameters:  $\alpha, \beta_0, \gamma, GMax, m_{pred}, t_{pred}$ 
Generate an initial firefly population:  $x^i$  ( $i=1, 2, \dots, N_f$ )
For  $t = 1$  up to  $GMax$ 
  If  $t_{pred} = .true.$ 
    For  $i = (m_{pred} + 1)$  up to  $N_f$ 
      Generation of new firefly
    End-For
  End-If
  Compute light intensity for  $x^i$ : proportional  $J(x^i)$ 
  For  $i = 1$  up to  $N_f$ 
    Compute the attractiveness:  $\beta = \beta_0 e^{-\gamma r_{ij}^2}$ 
    Move firefly- $i$  for the position of the brightest one
  End-For
End-For
End

```

Figure 1. Pseudo-code: FA with predation, where $m = \#$ fireflies, $t_{pred} =$ predation cycle, $N_f - m_{pred} =$ number of fireflies to be eliminated

5. Results for Precipitation Retrieval

BRAMS executions are carried out with the initial and boundary conditions from the analysis of the general circulation model (MCGA) from the CPTEC-INPE. The simulation is performed with resolution T126L28, where T126 represents the spherical harmonic expansion with the rhomboidal truncated with 126 terms, and L28 represents the 28 vertical levels in the MCGA/CPTEC model.

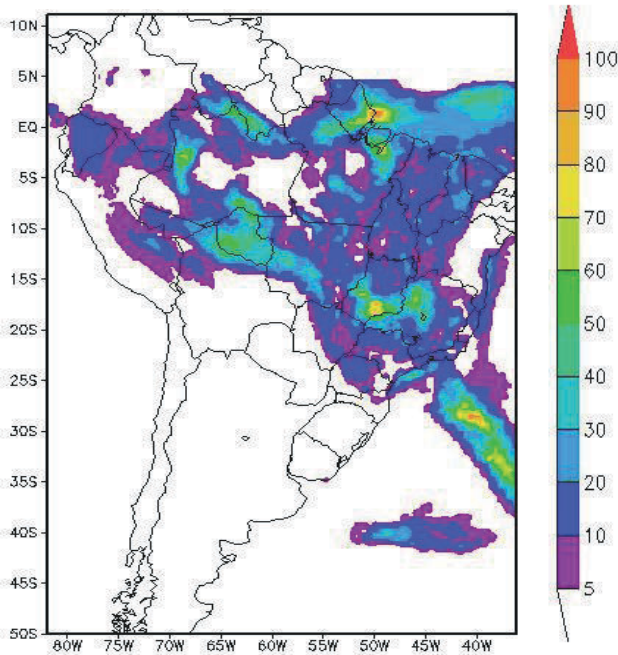


Figure 2. Accumulated precipitation (mm) for 24 h by BRAMS using GD approach (arithmetic mean) at the day 21/February/2004

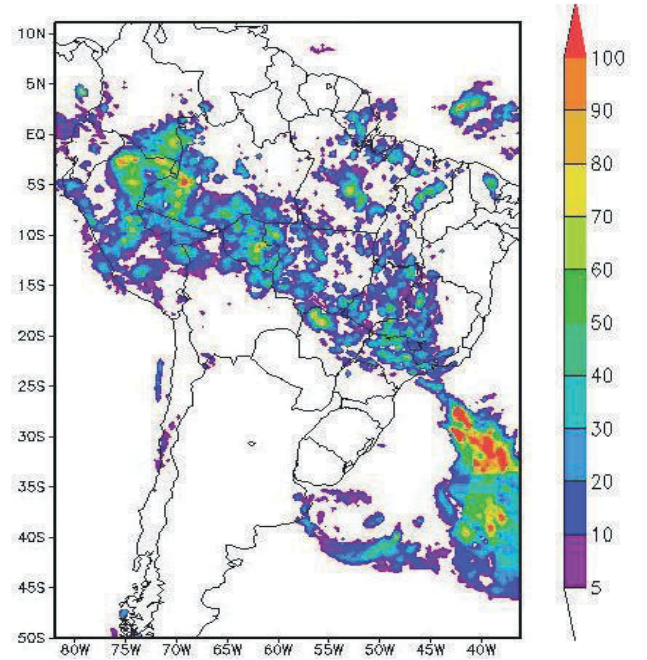


Figure 3. Accumulated precipitation (mm) for 24 h by TRMM satellite data at the day 21/February/2004

The experiment is designed to simulate one day of intense precipitation on the South America region, associated to the South Atlantic Convergence Zone (SACZ, or ZCAS for the acronym in Portuguese: Zona de Convergência do Atlântico Sul), during 21 up to 24/February/2004. Figure 2 shows the BRAMS simulation with GD scheme, but only a simple arithmetic mean was considered (all parameterization has the same relevance).

The evaluation of the methodology using the FA with predation is done with synthetic observation, following the same numerical inverse analysis described by Souza *et al.* [10]. Synthetic precipitations were obtained from combining of 5 BRAMS simulations for each parameterization. Table 1 shows the indexed weight for each cumulus parameterization. From such analysis, the results indicated that parameterizations AS, GR, and KF had a better agreement with the precipitation field measured by the TRMM data at 21-February-2004. Therefore, these parameterizations will be represented by heavier weights having a greater contribution for the ensemble.

Table 1. Weights for the synthetic precipitation filed

Parameterization	weight
W_{AS}	0.25
W_{GR}	0.35
W_{KF}	0.20
W_{MC}	0.15
W_{LO}	0.05
Summation	1.00

The resulting precipitation field is given by weighted combination:

$$P_s = w_{AS}P_{AS} + w_{GR}P_{GR} + w_{KF}P_{KF} + w_{LO}P_{LO} + w_{MC}m_{MC} \quad (5)$$

where the precipitation fields P_k are computed from the parameterizations codified in the BRAMS. Observational data contains noise. For a better representation of the measured (precipitation) data, the inversion deals with noisy precipitation field:

$$P_o = P_s(1 + \sigma\xi) \quad (6)$$

where $\sigma=0.02$ is the level of the noise (2%), and ξ is a random value associated to a Gaussian distribution.

Therefore, the synthetic precipitation field is given by Eq. (6), and the weight identification is performed with the inverse analysis detailed in Sections 3 and 4. With the use of FA with predation, the regularization was not necessary ($\eta = 0$). The inverse solutions were obtained with 40 fireflies in the population, and other parameters for the FA were: $\alpha = 0.2$, $\beta_0 = 1$, and $\gamma = 1$.

Table 2 shows the average results (μ) obtained using 25 realizations for regular FA and for FA with predation (FAp). It is clear that only for the weight to the GR parameterization a better estimated weight was found. Figure 4a depicts the convergence for the FA with predation, and it is much faster than standard FA (Fig. 4b). Figure 5 displays the precipitation field as computed by the BRAMS model, with exact weights adding 2% of level of noise (Fig. 5a), and the precipitation calculated with the weights estimated by FA with predation.

Table 2. Estimated weights using the synthetic precipitation filed

Parameterization	Exact	μ -FA	μ -FAp
W_{AS}	0.25	0.2399	0.2402
W_{GR}	0.35	0.3533	0.3423
W_{KF}	0.20	0.2051	0.2007
W_{MC}	0.15	0.1939	0.1561
W_{LO}	0.05	0.0619	0.0506

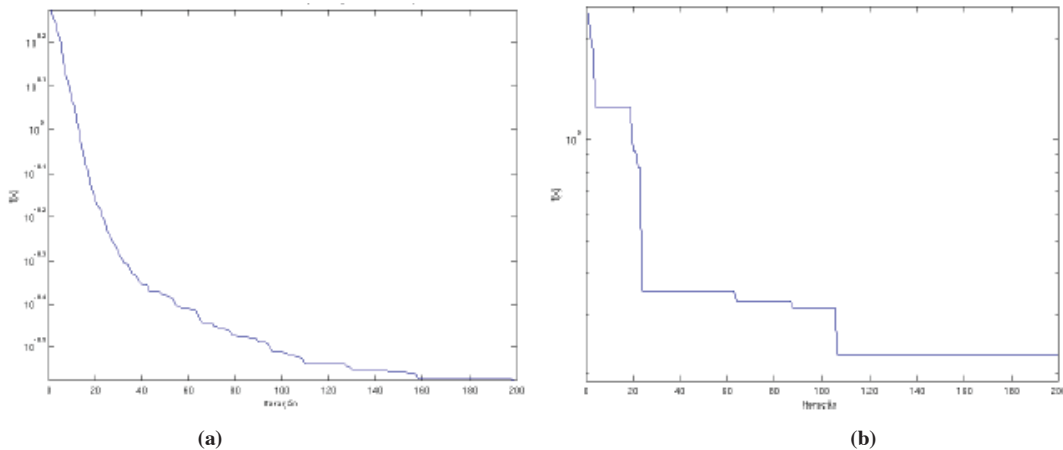


Figure 4. Convergence for the FA: (a) with predation operator, (b) regular FA

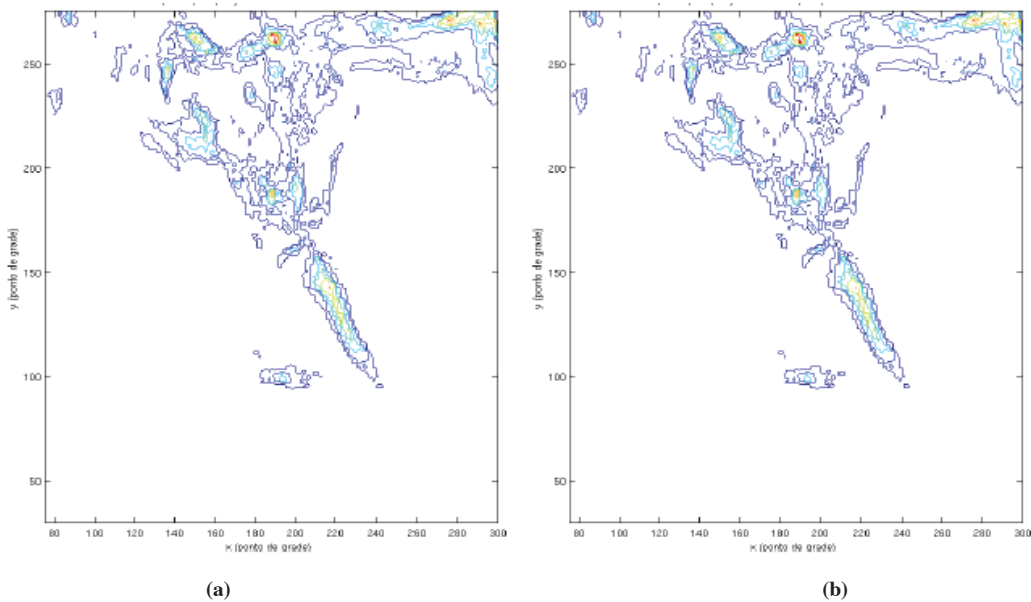


Figure 5. Precipitation field: (a) from the BRAMS with 2% of noise, (b) reconstruction with weights estimated by FA with predation

6. Final Remarks

This paper addresses the estimation of the precipitation field applying the GD's approach with several parameterizations. The goal is to obtain a better representation for the predicted precipitation employing a weighted linear combination from the parameterizations, where the best combination is found through an inverse analysis. The inverse problem is formulated as an optimization problem, where the objective function is given by the square difference between measurements and the precipitation field computed from a mathematical model. The optimal inverse solution was determined by using a new version of Firefly Algorithm (FA): the FA with predation, developed to be a general procedure to extract the bad elements in the population, introducing new fireflies (candidate solutions). This enhanced diversity is designed to search better inverse solutions.

Synthetic observations with noise were employed to evaluate the methodology. Indeed, the data from the satellite TRMM data presents errors on space and time over the South America [32]. The FA with predation produced better results than the standard FA. In addition to find better numerical values for the estimated weights, a faster convergence was noted with FA with predation.

Other schemes to represent the precipitation, such as super parameterization approach [30, 31], are out of our purposes, and such formulations were not discussed here.

The focus is to have a better forecasting to the most difficult meteorological variable to be predicted: rainfall, due to its variability on space and time. The tropical zone has intense convection activity, in particular on the North part of the Brazil, on the Amazon region. This is our main motivation to look for developing new methodologies.

Finally, similar to other meta-heuristics, FA – including FA with predation – has many free parameters. Santos *et al.* [33] have done a study to determine the best choice for the meta-heuristics parameters. This is our recommendation mainly for implementation of the technique for the operational prediction centers.

ACKNOWLEDGEMENTS

This work was supported Brazilian agencies for research support: CAPES, CNPq, FAPESP.

REFERENCES

- [1] Kain, J. S., Fritsch, J. M., Convective parameterization for mesoscale models: The Kain–Fritsch scheme, edited by: Emanuel, K. A. and Raymond, D. J., *Meteor. Monogr., American Meteorological Society*, Boston, 24, 165–170 (1993).
- [2] Stensrud, D. J., Bao, J.-W., Warner, T. T., Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems, *Mon. Weather Rev.*, 128, 2077–2107 (2000).
- [3] Casanova, S., Ahrens, B., On the Weighting of Multimodel Ensembles in Seasonal and Short-Range Weather Forecasting, *Mon. Weather Rev.*, 137, 3811–3822 (2009).
- [4] Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H., and Mitchell, H. L.: A System Simulation Approach to Ensemble Prediction, *Mon. Weather Rev.*, 124, 1225–1242 (1996).
- [5] Krishnamurti, T. N., Kishtawal, C. M., LaRow, T. E., Bachiochi, D. R., Zhang, Z., Williford, C. E., Gadgil, S., and Surendran, S.: Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble, *Science*, 285, 1548–1550 (1999).
- [6] Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., LaRow, T., Bachiochi, D., Williford, E., Gadgil, S., Surendran, S., Multimodel Ensemble Forecasts for Weather and Seasonal Climate, *J. Climate*, 13, 4196–4216 (2000).
- [7] Krishnamurti, T. N., Stefanova, L., Chakraborty, A., Vijaya Kumar, T. S. V., Cocke, S., Bachiochi, D. R., and Mackey, B.: Seasonal Forecasts of precipitation anomalies for North American and Asian Monsoons, *J. Meteorol. Soc. Jpn.*, 80, 1415–1426 (2002).
- [8] Cane, D., Milelli, M., Multimodel Super Ensemble technique for quantitative precipitation forecasts in Piemonte region, *Nat. Hazards Earth Syst. Sci.*, 10, 265–273, (2010).
- [9] Grell, G. A., Dévényi, D., A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, 29(14), article 1693, (2002).
- [10] Santos, A. F., Campos Velho, H. F., Luz, E. F. P., Grell, G. A., Gan, M. A., Firefly optimization to determine the precipitation field on South America. *Inverse Problems in Science and Engineering*, 21(3), 451–466 (2013a).
- [11] Santos, A. F., Freitas, S. R., Mattos, J. G. Z., Campos Velho, H. F., M. A. Gan, M. A., Luz, E. F. P., Grell, G. A., Using the Firefly optimization method to weight an ensemble of rainfall forecasts from the Brazilian developments on the Regional Atmospheric Modeling System (BRAMS). *Adv. Geosci.*, 35, 123-136, (2013b).
- [12] Freitas, S. R., Longo, K. M., Silva Dias, M. A. F., Chatfield, R., Silva Dias, P., Artaxo, P., Andreae, M. O., Grell, G., Rodrigues, L. F., Fazenda, A., Panetta, J., The Coupled Aerosol and Tracer Transport model to the Brazilian developments on the Regional Atmospheric Modeling System (CATT-BRAMS) Part 1: Model description and evaluation, *Atmos. Chem. Phys.*, 9, 2843-2861 (2009).
- [13] Yang, X.-S. *Nature-Inspired Metaheuristic Algorithms*. Luviner Press, Cambridge, 2008.
- [14] Glover, F., Future paths for integer programming and links to artificial intelligence. *Comput. & Ops. Res.*, 13(5), 533-549 (1986).
- [15] Luz, E. F. P., Becceneri, J. C., Campos Velho, H. F., Conceitualização do algoritmo vagalume e sua aplicação na estimativa de condição inicial da equação de calor, Proceedings of XI Workshop de Computação Aplicada do Laboratório Associado de Computação e Matemática Aplicada (XI WORCAP), São José dos Campos, SP, Brazil (2009).