



Using the Firefly optimization method to weight an ensemble of rainfall forecasts from the Brazilian developments on the Regional Atmospheric Modeling System (BRAMS)

A. F. dos Santos¹, S. R. Freitas¹, J. G. Z. de Mattos¹, H. F. de Campos Velho², M. A. Gan¹, E. F. P. da Luz², and G. A. Grell³

¹National Institute for Space Research, Center for Weather Forecasting and Climate Studies, Cachoeira Paulista, Brazil

²Laboratory for Computing and Applied Mathematics, National Institute for Space Research, São José dos Campos, SP, Brazil

³National Oceanic and Atmospheric Administration, Boulder, CO, USA

Correspondence to: A. F. dos Santos (ariane.frassoni@cptec.inpe.br)

Received: 14 February 2013 – Revised: 25 July 2013 – Accepted: 8 August 2013 – Published: 17 September 2013

Abstract. In this paper we consider an optimization problem applying the metaheuristic Firefly algorithm (FY) to weight an ensemble of rainfall forecasts from daily precipitation simulations with the Brazilian developments on the Regional Atmospheric Modeling System (BRAMS) over South America during January 2006. The method is addressed as a parameter estimation problem to weight the ensemble of precipitation forecasts carried out using different options of the convective parameterization scheme. Ensemble simulations were performed using different choices of closures, representing different formulations of dynamic control (the modulation of convection by the environment) in a deep convection scheme. The optimization problem is solved as an inverse problem of parameter estimation. The application and validation of the methodology is carried out using daily precipitation fields, defined over South America and obtained by merging remote sensing estimations with rain gauge observations. The quadratic difference between the model and observed data was used as the objective function to determine the best combination of the ensemble members to reproduce the observations. To reduce the model rainfall biases, the set of weights determined by the algorithm is used to weight members of an ensemble of model simulations in order to compute a new precipitation field that represents the observed precipitation as closely as possible. The validation of the methodology is carried out using classical statistical scores. The algorithm has produced the best combination of the weights, resulting in a new precipitation field closest to the observations.

1 Introduction

Precipitation is one of the most important meteorological variables of the climate system and directly affects human activities. It is often the variable of greatest interest to the public in a daily weather forecast. People want to know whether or not precipitation will occur, and if a precipitation event is predicted, an accurate quantitative precipitation forecast is also of great value.

However, while the precipitation forecast provides very important information for the population, it is the most difficult variable to predict, due to the fact that accuracy in the timing and spatial distribution of the precipitation is crucial. Moreover, quantitative precipitation forecasting, especially in tropical and subtropical regions in summertime, is particularly challenging because of the convective nature of much of the precipitation (Stensrud et al., 2000). Precipitation is a final result of a set of physical processes, associated with convection and cloud formation, that are not well understood or well represented in the numerical models. Some of these physical processes and parameters cannot be explicitly predicted in full detail on the model grid points (or wave numbers for spectral models), but their effects on the resolved variables in the model are crucial for obtaining a correct forecast.

The model physics, especially the convective processes, contribute to uncertainties in weather and climate forecasts due to the imperfect representation of the atmosphere in the model (Tribbia and Baumhefner, 1988). The representations of the physical and dynamical processes in the atmospheric

model are some of the sources of errors in numerical forecasts, and they are termed external errors (due to deficiencies; Reynolds et al., 1994). These errors grow with the time integration of the atmospheric flow, thus making predictability dependent on the atmospheric state (Lorenz, 1963).

In a regional model, the temporal and spatial scales are reduced compared with global circulation models, and the location and amounts of precipitation should be better represented in these models, though this is not always the case. The models have difficulty in developing and organizing convection at the correct location and time (Kain and Fritsch, 1992). The characteristics of precipitation forecasts are often directly affected by the assumptions used to develop the model parameterization schemes for convection and other processes (Stensrud et al., 2000). In this sense, the problem associated with convective parameterizations is to find a relation between the intensity of the subgrid-scale convective activity and the large-scale variables. Although extensive progress has been made since the first studies about convective parameterization, it is still an open task. Hence, cumulus parameterization is one of the most difficult problems in both weather and climate prediction.

A cumulus parameterization is an attempt to account for the net effect of an ensemble of clouds on the scale of the atmospheric model (Arakawa and Schubert, 1974). The goal of cumulus parameterization is to determine changes in the simulated large-scale environment due to the collective influence of multiple cumulus clouds (Jones and Randall, 2011). A statistical approach is used to assume the solution, which always introduces errors, providing an additional source of uncertainty to the stochastic nature of the atmosphere.

Each parameterization method derives information about the processes from the meteorological variables, considering a set of assumptions. Many schemes have been developed to represent the convection process. For some interesting review articles on convective parameterizations the reader is referred to Yanai et al. (1973), Frank (1983), Grell et al. (1991), Emanuel and Raymond (1993) and Arakawa (2004). New ideas that have recently been implemented include build-in stochasticism (Grell and Dévényi, 2002; Lin and Neelin, 2003), and the super parameterization approach (Grabowski and Smolarkiewicz, 1999; Randall et al., 2003). For our purposes we will not discuss the super parameterization approach, which implements a different design.

The more commonly used parameterization approaches differ significantly in their closure assumptions and description of the iteration between the environment and convection. Closure refers to the link between the assumptions in the parameterization and the forecast variables, and it closes the loop between the parameterization and forecast equations. The main concept of a parameterization scheme is the choice of a closure, which consists of a set of statistical equations governing the system, whose size is too large (Arakawa, 2004). The differences among the parameterizations are a consequence of the uncertainties in the understanding of the

physical and dynamical processes of convection, particularly with respect to how to express the iteration between the large-scale flow and convective clouds in parameterized terms (Bao et al., 2011). The debate regarding which assumptions are most appropriate under certain conditions and locations has led to many discussions in past studies (Grell and Dévényi, 2002).

A way to make use of the uncertainties in numerical prediction is to generate an ensemble of forecasts to span the space of possible numerical solutions (Stensrud et al., 2000). The goal of ensemble forecasting is to predict the forecast probability of future weather events by the integration of an ensemble of numerical predictions (Lorenz, 1965; Mullen and Baumhefner, 1994). The ensemble members can be created using different initial conditions, different physical parameters or by applying different model systems (Casanova and Ahrens, 2009). The use of different physical parameterizations, which creates different model configurations, has been used to generate a multi-model ensemble of experiments (Houtekamer et al., 1996; Stensrud et al., 2000). Stensrud et al. (2000) compared an ensemble using configurations with different physical processes and identical initial conditions with an ensemble of different model initial conditions and identical model configurations. Specifically, five different convective parameterization schemes were used within the ensemble. They investigated 48 h forecasts during summertime in North America. They found more skill when using a model physics ensemble than when using an initial condition ensemble when the large-scale forcing for upward motion was weak. When the large-scale forcing for upward motion was strong, improved forecast skill was found with the initial condition ensemble. According to the authors, the results suggested that varying the model physics is a potentially powerful method to use in the creation of ensembles.

Bao et al. (2011) mentioned that the use of parameterizations separately in a single model framework provides a way to take advantage of the uncertainties in numerical prediction in terms of generation of an ensemble of different realizations in order to provide an ensemble prediction. In this sense, Grell and Dévényi (2002), hereafter GD, developed a convective parameterization scheme proper to create a physical ensemble inside the same model system and using the same initial conditions if the user desired. GD developed a deterministic scheme expanding the convective parameterization of Grell (1993) to include several assumptions of classical closures and parameters commonly used in convective parameterizations, which are often used in different realizations of the same model to create an ensemble of physics. The GD parameterization scheme has been developed to provide more freedom to users to choose one or more assumptions and closures within the extensive existing options. The ensemble members are chosen to allow a large spread in terms of accumulated convective rainfall (Bao et al., 2011). The ensemble of closures consists of disturbances around the classical closures of Grell (1993); a simplified version of Arakawa

and Schubert (1974) and Kain and Fritsch (1993), low-level omega (Brown, 1979; Frank and Cohen, 1987) and integrated vertical advection of moisture (Krishnamurti et al., 1983), which is a Kuo (1974) type closure, hereafter GR, AS, KF, LO and MC, respectively. GD used these members to determine the ensemble mean. Using the ensemble mean realization, a numerical representation of precipitation and atmospheric heating and moistening rates is obtained.

However, the ensemble mean is most likely not the best choice, since some closure options can provide the best precipitation forecasts individually (Santos et al., 2013). The GD ensemble could be improved by using the best combination of ensemble members. In a first approximation, the results of the model with different closure options could be combined with weighting the ensemble of precipitation forecasts of the model system. Many methodologies have been used to combine different members of an ensemble to generate a bias-corrected ensemble. An ensemble method which requires several model outputs and is applied in several studies is called a Multimodel-Superensemble (Krishnamurti et al., 1999; Krishnamurti et al., 2000, 2002), which is weighted with an adequate set of weights calculated during a period of training (Cane and Milelli, 2010). Cane and Milelli (2010) used several model outputs to generate weights to obtain a combined estimation of precipitation by least-square minimization of the difference between the model and the observed field using the Gauss–Jordan method applied to Piemonte, Italy. The authors used the data collected in the period March 2006–August 2008 from 342 weather stations, along with different model outputs from European Centre for Medium-Range Weather Forecasts (ECMWF) and different versions of the model Consortium for Small-scale Modeling (COSMO). The weights were calculated only once, during a training period, and then applied during the forecast time integration. The results showed significant improvement for low thresholds of precipitation, with a reduction of the bias and of the false alarm rates.

In order to identify the role of each component of the convective parameterization ensemble in representing precipitation over South America using the Brazilian developments on the Regional Atmospheric Modeling System Freitas et al. (2009) (BRAMS), an inverse problem methodology is applied to estimate the parameters associated with each member of the ensemble of GD and then construct a new precipitation field with bias removal. The inverse problem is formulated as an optimization problem for retrieving the weights of a set of precipitation simulations obtained from the GD convective parameterization of the model BRAMS, and it is solved using a metaheuristic optimization method called the Firefly algorithm (FY) (Yang, 2008). The parameter estimation consists of minimizing an objective function using model parameter adjustment. The objective function is a measure of the distance between the observed data and BRAMS forecasts, the root mean square deviation of the model forecasts and observations.

The remainder of this paper is organized as follows: Sect. 2 gives a description of the FY, Sect. 3 describes briefly the model BRAMS, Sects. 4, 5, and 6 give a description of the experiments, observed data used and mathematical formulation of the proposed problem, respectively. Section 7 describes the statistical verification of the results, while Sect. 8 discusses the main results. Finally, Sect. 9 provides some conclusions concerning the solutions obtained by the FY application and directions for future work.

2 The firefly algorithm (FY)

The FY was proposed by Yang (2008), and is based on the bioluminescence process which characterizes fireflies. Preliminary studies with this methodology have produced satisfactory results for this important application (Luz et al., 2009; Santos et al., 2013). According to the literature, although the FY has many similarities with other algorithms which are based on swarm intelligence (such as Particle Swarm Optimization – PSO, Artificial Bee Colony optimization – ABC, and others), the FY has many advantages over other techniques (Apostolopoulos and Vlachos, 2011). Lukasik and Zak (2009) and Yang (2009, 2010) describe the simplicity of the concept and implementation of the FY. The algorithm is very efficient and can outperform other conventional algorithms, such as genetic algorithms and Simulated Annealing (Kirkpatrick et al., 1983), as demonstrated in Luz (2012). Its main advantage is the fact that it uses mainly real random numbers, and is based on the global communication among the swarming particles, and as a result, it seems more effective in multiobjective optimization (Apostolopoulos and Vlachos, 2011).

According to Yang (2008), scientists do not yet have complete understanding of the function of the flashing lights of fireflies. However, there are at least two important functions associated with the flashes: (a) attracting mating partners, and (b) attracting potential prey.

In order to implement the algorithm based on the flashes of fireflies, Yang (2008) used the following three idealized rules: (i) the fireflies are unisex, so one firefly will be attracted to other fireflies regardless of their sex; (ii) attractiveness is proportional to brightness, so, for any two flashes, the less brighter one will move towards the brighter one. Furthermore, since the attractiveness is proportional to the brightness, both decrease as the distance between the fireflies 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.

There are two important pieces of information for determining the FY: the light intensity variation and the attractiveness formulation. For simplicity, the attractiveness is determined by the brightness, which in turn is associated with the objective function.

For a simple case, the brightness (I) in a particular location is a function of its position x as follows:

$$I(x) \propto f(x), \quad (1)$$

and the attractiveness (β) is relative to the firefly position from which the more attractive firefly is observed. Thus, β will depend on the distance r_{ij} between firefly i attracted by the brightness of firefly j . Moreover, light intensity decreases with increasing distance from its source and depends on the propagation medium, so the attractiveness varies with the degree of absorption.

The light intensity is assumed to be $[I(r)]$ and it varies according to the inverse square distance r ,

$$I(r) = \frac{I_s}{r^2}, \quad (2)$$

where I_s is the light intensity at the source. For a given medium with a fixed light absorption coefficient γ , the light intensity varies with r as follows:

$$I = I_0 e^{-\gamma r}, \quad (3)$$

where I_0 is the original light intensity. In order to avoid the singularity at $r = 0$ in the Eq. (2), the combined effect of both the inverse square law and absorption can be approximated as the following Gaussian form:

$$I(r) = I_0 e^{-\gamma r^2}, \quad (4)$$

which can be approximated as

$$I(r) = \frac{I_0}{1 + \gamma r^2}. \quad (5)$$

The attractiveness β is proportional to the light intensity seen by adjacent fireflies as follows:

$$\beta = \beta_0 e^{-\gamma r^2}, \quad (6)$$

where β_0 is the attractiveness at $r = 0$. This function can be approximated by:

$$\beta(r) = \frac{\beta_0}{1 + \gamma r^2}. \quad (7)$$

The distance between any two fireflies i and j at x_i and x_j , respectively, is the Cartesian distance

$$r_{ij} = |x_i - x_j| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2}, \quad (8)$$

where $x_{i,k}$ is the k -th component of the spatial coordinate x_i of i -th firefly.

The movement of a firefly i attracted to another more attractive (brighter) firefly j in a given time step (t) is determined by:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2} (x_j^t - x_i^t) + \alpha \left(\text{rand} - \frac{1}{2} \right), \quad (9)$$

where α is the randomization parameter and “rand” is a random number uniformly distributed in $[0, 1]$.

The second term on the right hand side of Eq. (9) is due to the attraction, while the third term is the randomization term, associated with the random movement of firefly i toward firefly j . Without this term, the fireflies could possibly be attracted to a firefly that is not necessarily the brightest. The solution would be restricted to a local minima, the best solution in the local search space. With the randomization term, the search over small deviations makes it possible to escape from local minima, creating a higher chance of finding the global minimum of the function.

The γ parameter (in m^{-1}) is an absorption coefficient which controls the decrease of the light intensity. It characterizes the attractiveness variation and is crucially important in determining the speed of convergence and how the FY behaves. In theory, $\gamma \in [0, \infty)$, but in practice and in most applications, $\gamma = O(1)$ and is determined by the characteristic length Γ of the system to be optimized. For most applications, it typically varies from 0.1 to 10.

3 The BRAMS model

The BRAMS model (Freitas et al., 2009) is a joint project of several Brazilian institutions, including the Center for Weather Forecasting and Climate Studies (CPTEC) of the National Institute for Space Research (INPE), and was initially funded by FINEP (The Brazilian Funding Agency). BRAMS is based on the Regional Atmospheric Modeling System (RAMS) (Walko et al., 2000), with several new functionalities and parameterizations. BRAMS is a numerical model developed to simulate atmospheric circulations on many scales. It solves the time-split compressible non-hydrostatic equations described by Tripoli and Cotton (1982). The set of physical parameterizations in BRAMS is appropriate for simulating processes such as surface-air exchange, turbulence, convection, radiation and cloud microphysics (Freitas et al., 2009). The BRAMS model includes an ensemble version of a deep and shallow cumulus scheme based on the mass flux approach (the GD scheme).

The convective parameterization trigger function uses the turbulent kinetic energy (TKE) of the RAMS Planetary Boundary Layer (PBL) parameterization to modulate the maximum distance that air parcels can rise from their source level and, based on that, to determine if a grid column will be able to sustain convection (Freitas et al., 2009). The trigger function is modified using the parameter `cap_max`, which represents the maximum distance the air parcel can rise and trigger the convective portion of a column if it can reach the condensation level and, subsequently, free convection. The `cap_max` is modified with three values to be used in the ensemble. The ensemble is built with 3 perturbations of the trigger function (as described before), 3 perturbations of the precipitation efficiency and a total of 16 perturbations of the

mass fluxes at cloud base generated from 5 mass flux closure options, providing a total number of members equal to $3 \times 3 \times 16 = 144$. Finally, by taking the arithmetic mean over the trigger functions, precipitation efficiencies and closure perturbations for each of the 5 mass flux closure options, the dimension of the ensemble is reduced to 5 members, denoted by GR, MC, LO, AS and KF. Then, an arithmetic mean of these 5 members provides a unique solution denoted by ensemble mean (EN). A detailed description of the 144 members of the convective parameterization is given in GD.

4 Experimental design

The BRAMS model version 4.3 was used to simulate precipitation over South America. The model was run for a forecast length of 24 h, once a day, from 1–31 January 2006, initialized at 12:00 UTC, with a restart every 24 hours. The following configuration was used: the GD convective parameterization scheme, a model grid with 25 km horizontal resolution covering South America and 100 m vertical resolution in the first level. The vertical resolution varied telescopically with higher resolution at the surface with a ratio of 1.1 up to a maximum vertical resolution of 950 m, with the top of the model at approximately 19 km (a total of 40 vertical levels). For initial and boundary conditions, we used the CPTEC/INPE Atmospheric General Circulation Model (AGCM) analysis with T126L28 resolution, where T126 is the rhomboidal truncation at wave number 126, and L28 indicates the number of model vertical levels.

One approach to generating an ensemble is to use different model physical process parameterization schemes to construct various versions of a model and produce an ensemble of simulations that start from the same initial condition (Stensrud et al., 2000). If all configurations of the model are equally skillful, the ensemble mean should be a good approximation. As this assumption cannot be verified, since each member has a different skill, the ensemble members should be weighted unequally (Thompson, 1977). To create this physics ensemble, the focus of this study is the choice of different parameterization assumptions for deep convection, yet there is uncertainty in other aspects of the model, including the initial and boundary conditions, and parameterizations of turbulence, the planetary boundary layer (as shown in Stensrud et al., 2000), microphysics and radiation, among others. The possible role of these other parameterizations in the ensemble system has not yet been investigated but will be the focus of future studies.

In this paper, six different model simulations were performed. From this, a set of distinct precipitation forecasts (P) was created using five different closure options (i.e. GR, MC, LO, AS and KF), hereafter P_{GR} , P_{MC} , P_{LO} , P_{AS} and P_{KF} , respectively, and a run was performed using the EN closure. The EN precipitation field was used as the control. All simulations started from the same model initial condition. The

weighted ensemble (P_M) based on the set of P was created using the FY. P_M is defined as the linear combination of P_{GR} , P_{MC} , P_{LO} , P_{AS} and P_{KF} . The methodology proposed to obtain a well-weighted ensemble is detailed in Sect. 6.

5 Observations

The following definition of a 24 h period was chosen for this study: from 12:00 UTC on the previous day to 12:00 UTC on the current day (precipitation accumulated in 24 h) for January 2006. During summertime (austral summer December–February), more than 50 % of the total annual precipitation over most of tropical and subtropical South America typically occurs, mainly in central and southern South America (Rao and Hada, 1990; Figueroa and Nobre, 1990; Gan et al., 2004), due to the climatological characteristics of the South American monsoon system (Zhou and Lau, 1998; Gan and Moscati, 2003; Gan et al., 2004). This period was chosen because the precipitation in January spans the range from weak to strong forced convective forecast problems routinely handled by operational forecasters (Stensrud et al., 2000).

The precipitation data is from the CPTEC/INPE, and is a combined dataset using a technique called MERGE (Rozante et al., 2010). The authors combine data from the Tropical Rainfall Measuring Mission (TRMM) satellite precipitation estimates (Huffman et al., 2007) with rain gauge observations over South America. Although TRMM is highly valuable for numerical model evaluation, the MERGE technique is used to correct TRMM precipitation estimation problems, because systematic errors have been identified, in particular in the eastern portion of Northeast Brazil, where precipitation is underestimated due to the formation of warm clouds (Huffman et al., 2007). Warm clouds are clouds that purely consist of liquid water. The temperature inside warm clouds is warmer than 0 °C. The only requirement for a cloud to be classified as a warm cloud is the absence of ice crystals (IAC, 2013).

6 Solving the inverse problem: weight estimation

To find the best combination of P in order to obtain P_M , we applied the FY to solve an inverse problem of parameter estimation, computed as an optimization problem. The goal is to identify the optimum weight values associated with each member of the ensemble. The weight set is the unknown vector of parameters denoted by W^T with 5 dimensions ($d = 5$). We assume the objective function $J(P)$ as the inverse of the light intensity of a firefly, since we need minimize the square of the difference between P_M and the observed precipitation field or MERGE data (P_O). The estimator $J(P)$ is a random variable that minimizes the Euclidian norm square of $P_M - P_O$, i.e. $J(P)$ minimizes

Table 1. Parameter values used in the FY.

Parameter	Value
α	0.1
β_0	0.5
γ	10.0
n	50
G	100

$$\begin{aligned}
 J(P) &= |P_M(W^T) - P_O|^2 \\
 &= [P_M^W - P_O]^T [P_M^W - P_O] \\
 &= \sum_{i=1}^{w_i} [P_M(W^T) - P_O]^2, P_M = \sum_i w_i P_i \quad (10)
 \end{aligned}$$

where $W^T \equiv [w_i, i = 1, \dots, d]^T$, P_i ($i = 1, 2, \dots, N_w$) and $N_w = d$. Each member of P_M is associated with a weight in the parameter vector, represented by $W^T = [w_{GR}, w_{MC}, w_{LO}, w_{AS}, w_{KF}]^T$. Therefore,

$$P_M = w_{GR} P_{GR} + w_{MC} P_{MC} + w_{LO} P_{LO} + w_{AS} P_{AS} + w_{KF} P_{KF}. \quad (11)$$

The subscripts GR, MC, LO, AS and KF denote the different choices of closures (see Sect. 3).

Each firefly represents a candidate solution (the W^T vector), and the brightest firefly identifies the best weight set for weighting P_M . The best solution (the brightest and most attractive firefly) is the minimum value of $J(P)$.

The parameters used in the FY are: the number of generations (G), i.e. the number of iterations; the number of fireflies (n) for each G , β_0 , α and γ . The total number of function evaluations was 5000 ($n \times G$). The parameters used are summarized in Table 1. The sensitivity of the algorithm with respect to the chosen parameters was previously tested, allowing the choice of the best parameters to be used together to solve the proposed inverse problem. More details can be found in Santos et al. (2012). However, the FY is very sensitive to parameter changes. There is no guarantee that the parameters used in this work will be the best choice for another condition. The parameters used are valid only for the BRAMS forecasts with specifications described in Sects. 3 and 4.

Figure 1 shows a flowchart of the proposed inverse problem solution using FY. The first step is to set the FY parameters. Here, we associate each firefly with the vector W^T , with each component between the interval $[0, 1]$. The BRAMS precipitation simulations (P_{GR} , P_{MC} , P_{LO} , P_{AS} and P_{KF}) are supplied as input to FY and the output is the best value for W^T .

The initial candidate solution for the iterative optimization process is randomly generated. The fireflies are sorted based on the $J(P)$ evaluation, where the brightest firefly is classified on the first position, and so on. This process is the evaluation of the first firefly population (first generation), and it

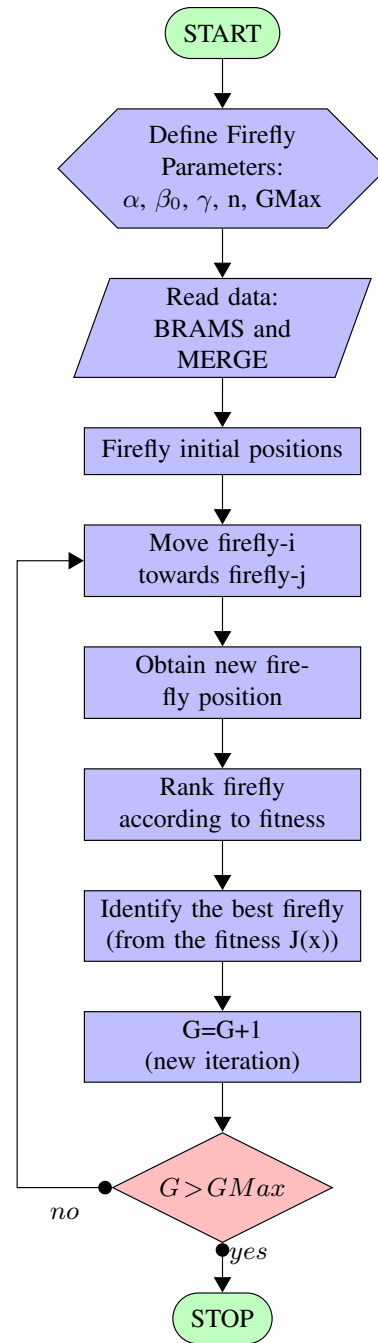


Fig. 1. The firefly algorithm implemented.

is repeated for next generations. If the iteration number G is less than the maximum number of generations ($GMax$), the process continues. This is the termination criterion that needs to be satisfied to end the iterative process. The generation of a new solution, i.e. the movement of a firefly, is given by Eq. (9). For each generation, the swarm of $n = 50$ fireflies is ranked based on their light intensity, and the firefly with the maximum light intensity (i.e. the solution with the minimum

$J(P)$) is chosen as the potential optimal solution. In the last iteration, the firefly with the brighter light intensity among the swarm of fireflies is considered the optimal solution of the inverse problem.

Finally, with the computed optimized solution, the vector W^T for each grid point of the domain is saved, and the final precipitation field is retrieved. The retrieved precipitation is P_M given by Eq. (11), i.e. the linear combination between each precipitation simulation and the corresponding weight.

7 Verification forecasts

In this paper we used some classical statistical scores to compare results from the individual ensemble members, the GD EN and the ensemble weighted using the FY.

To evaluate model skill in simulating the actual rainfall, the bias score and the Equitable Threat Score (ETS) were computed at each grid point of the domain. The ETS measures the skill in predicting the precipitation amounts over a given threshold with respect to a random forecast (Rogers et al., 1995). The bias score is the ratio of the forecast precipitation to the observed precipitation for amounts exceeding a given precipitation category (Wilks, 1995). Bias greater than one indicates that the event was forecast more often than observed, or overpredicted.

The bias score and the ETS were calculated for different precipitation thresholds (0.254, 2.54, 6.53, 12.7, 19.05, 25.4, 38.1, 50.8 mm). These thresholds were chosen because they are commonly used and are standard thresholds used operationally at CPTEC/INPE. A brief description of these scores is given in Appendix A. Both indices were computed over the entire domain of South America and the adjacent oceans (both Pacific and Atlantic Oceans) and the domain was divided into five small areas. In these areas, there are different precipitation regimes, and the partition areas follow criteria used by Chou et al. (2002) with modifications to separate southern South America from central South America. These areas show completely different precipitation regimes during the entire year. Particularly during the summer months, it is known that when the South Atlantic Convergence Zone (SACZ) is causing precipitation in Southeastern Brazil, there is no precipitation in the southern part of Brazil (Herdies et al., 2002). It is very important to consider the different precipitation regimes in central and southern South America. The corresponding geographical locations are indicated below and the areas are indicated in Fig. 2:

- South America: 82.6° W/32.6° W, 49.9° S/13.6° N;
- Southern South America (area 2): 82.6° W/32.6° W, 49.9° S/24° S;
- Central South America (area 3): 82.6° W/32.6° W, 24° S/13.5° S;

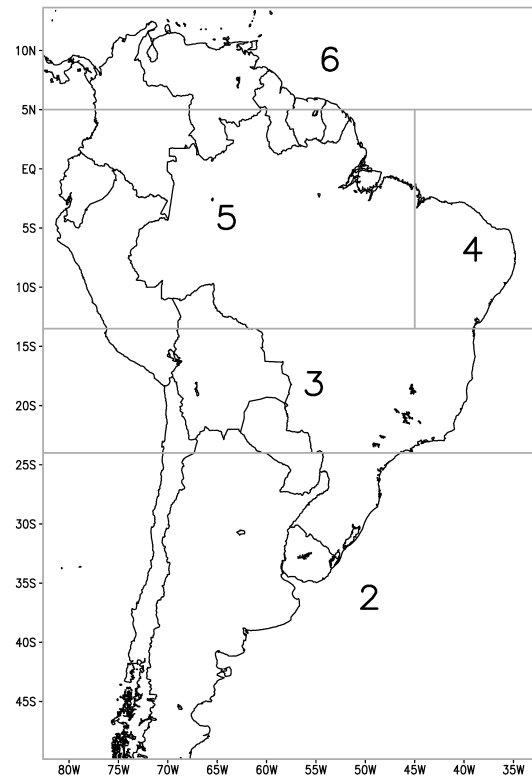


Fig. 2. Areas used to calculate the statistical scores. The numbers represent corresponding areas indicated in the text.

- Northeastern Brazil (area 4): 45° W/32.6° W, 13.5° S/5° N;
- Northwestern South America (area 5): 82.6° W/45° W, 13.5° S/5° N;
- Northern South America and tropical Atlantic Ocean (area 6): 82.6° W/32.6° W, 5° N/13.6° N.

8 Results

Figure 3 shows the January 2006 average 24-h accumulated precipitation as simulated by the BRAMS model using 5 different closures as well as the EN. A small spread is observed among all ensemble members (Fig. 3a–e), except for the AS member (Fig. 3b), which deviates more substantially. In the EN (Fig. 3f), the spatial pattern of the precipitation forecast is very similar to that of the individual members, as expected. Each ensemble member and the EN simulated a huge area of precipitation over the preferential position of the Intertropical Convergence Zone (ITCZ) during January, except for AS, which simulated intense rainfall over northern Brazil. The ITCZ in January is typically observed between the equator and 5° N (Waliser and Gautier, 1993; Nobre and Srunkla, 1996).

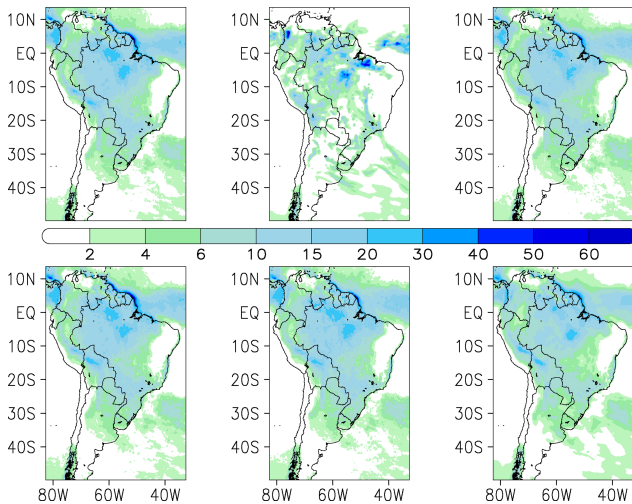


Fig. 3. 24-h accumulated precipitation for January 2006 (in mm) using the closures: (a) GR, (b) AS, (c) MC, (d) LO, (e) KF and (f) EN.

A comparison among the spatial distributions of the observed rainfall, that simulated by the GD ensemble mean (EN) and the ensemble with mean weighted with FY is shown in Fig. 4. The observed rainfall (Fig. 4a) shows a near uniform distribution of precipitation covering most of South America and adjacent oceans. More precipitation is observed between the equator and 10° N, associated with the ITCZ.

During January 2006, the ITCZ had been located around its climatological position near the African coast and slightly to the north near the Brazilian coast. According to the estimation in pentads of the mean position of the ITCZ in January 2006 from the location of the minimum values of Outgoing Longwave Radiation (OLR) along the Equatorial Atlantic Ocean (Climanálise, 2006), in most pentads it was north of the climatological position near the Brazilian coast. However, during the first half of the month, the iteration between the ITCZ and the upper tropospheric cyclonic vortices favored the occurrence of rain on the north coast of Brazil.

Due to the intense activity of the ITCZ, there is more precipitation over the northern part of South America, with an accumulated amount of 6 mm to 10 mm (Fig. 4a). More intense accumulated precipitation is observed over the Atlantic Ocean (around 20 to 30 mm). Over the central part of South America, the elongated precipitation band from northwestern to southeastern South America is associated with the SACZ (Casarin and Kousky, 1986; Kodama, 1992), which is typically observed during the austral summer. Convective activities associated with typical thermodynamic summer cloud systems and mountain effects are found in western and northeastern Bolivia and southeastern Peru, producing rainfall amounts of around 20 to 30 mm. Intense precipitation is also found over northern and northeastern Argentina, as well as over Uruguay.

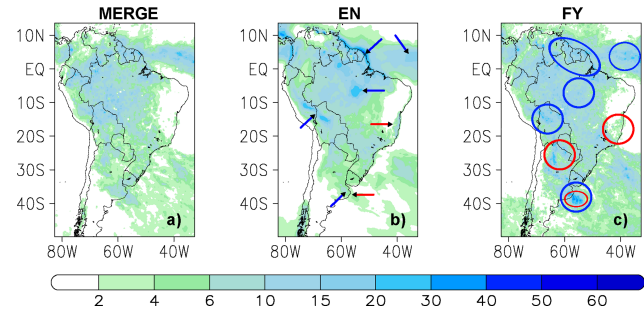


Fig. 4. 24-h accumulated precipitation for January 2006 (in mm): (a) observation (MERGE), (b) using the ensemble simple mean (EN) and (c) field weighted by FY. The blue (red) arrows and circles point out areas where the FY performance is better (worse) in comparison with EN.

The precipitation simulation by GD using the EN is shown in Fig. 4b. Comparing the EN field with the observed precipitation (Fig. 4a), one can see an overestimation of precipitation on the northern coast of South America and in the central-eastern Amazon basin, southeastern Peru and central Bolivia. Overestimation is also observed over the eastern part of Southeast Brazil, where the largest urban centers of Brazil and concentrated populations are located. Also, overestimation of light precipitation is evident in a large area of the ITCZ. Underestimation is observed over the western Amazon basin, southeastern Argentina and Uruguay.

The weighted precipitation generated by FY is shown in Fig. 4c. The results are clearly better when compared to the EN. Blue (red) arrows in Fig. 4b point out areas where the FY performance is better (worse) in comparison with the EN. Over the northern coast of South America and adjacent Atlantic Ocean, over the central Amazon basin and southeastern Peru and northern Bolivia, the FY reduced the overestimation observed in the GD EN field, generating a simulated precipitation field very close to the observation. The method showed good performance in increasing precipitation over the southeastern and northern portions of Argentina, where the GD EN underestimated precipitation. However, in these areas, FY overestimated precipitation, as shown by the red circles in Fig. 4c. Overestimation of light precipitation was observed (not shown) over central-southeastern Brazil, where precipitation amounts of about 2 to 6 mm were observed. FY did not reduce the total overestimation over the eastern part of Southeast Brazil, as highlighted by the red circle in Fig. 4c, but compared with the GD EN, highlighted by the red arrow in Fig. 4b, a reduction of the overestimation was observed.

For the bias score computed over all domains (Fig. 5), the magnitudes of the scores are comparable with the bias scores of other regional models shown in other studies (Chou and da Silva, 1999). Figure 5a shows the bias score mean averaged over the entire domain. LO and KF members produce bias scores higher (lower) than 1 for the thresholds lower (higher) than 25 mm, with a maximum (minimum)

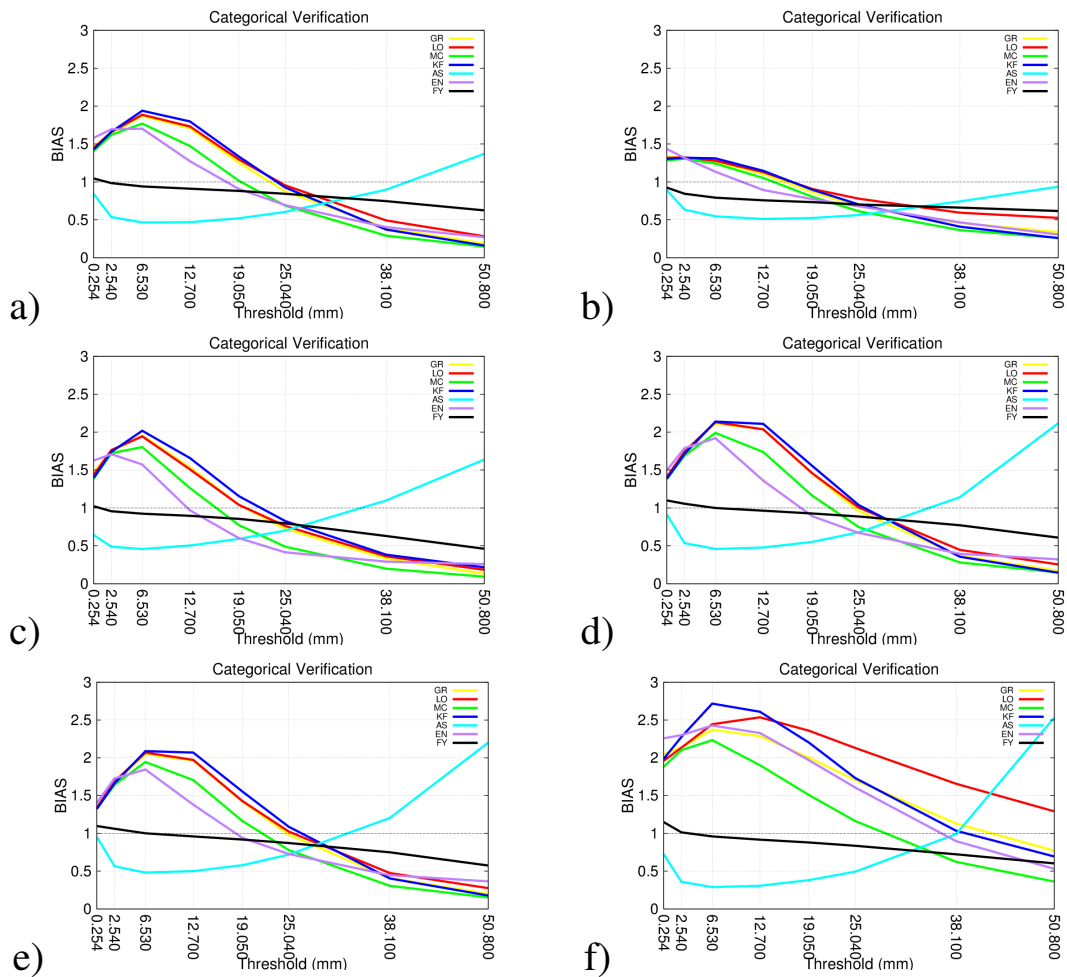


Fig. 5. Bias score vs precipitation thresholds (0.254, 2.54, 6.53, 12.7, 19.05, 25.4, 38.1, 50.8 mm) for (a) South America, (b) area 2, (c) area 3, (d) area 4, (e) area 5, (f) area 6.

value of about 2 (0.25) for threshold 6.53 mm (50.8 mm). The MC member produces a similar performance, but bias higher (lower) than 1 is observed at thresholds lower (higher) than 19.05 mm, with minimum value of about 0.15. The EN produces a maximum bias of 1.75 for thresholds between 2.54 and 6.53 and bias lower than 1 for thresholds higher than 17 mm. The EN solution reduces bias observed in individual members for thresholds between 2.54 mm up to about 22 mm. The member AS exhibits an opposite behavior, with bias lower (higher) than 1 for thresholds lower (higher) than about 40 mm. All members overestimate lower precipitation amounts and underestimate higher precipitation amounts, while AS produces heavy precipitation in localized areas and underestimates light precipitation amounts in most parts of the South America domain.

When the bias score is computed to verify the impact of the weighted precipitation using the FY over the entire domain, a substantial increase in the skill of the model compared with the skill based on individual members of the

ensemble is observed (Fig. 5a). Of particular interest is a comparison with the EN simulation. FY produces bias for thresholds between 0.254 and 38.1 mm between 1 and 0.75. For the highest threshold, the FY bias decreases to about 0.6. The FY bias exhibits a smooth decrease from the best value of bias at lower thresholds to small bias at higher thresholds. Although the FY underestimates precipitation at the highest thresholds, the other members underestimate it as well. However, with FY the underestimation is smaller than that of the other members. The one exception to this is the AS member, which overestimates precipitation at the 50.8 mm threshold and underestimates it at all lower thresholds.

Analyzing as well the bias score applied to the five smaller areas over South America, the effect of the EN is to decrease the bias as observed over the entire domain, except for area 6. In addition, the increase of the skill of FY in all domains compared to the skill of the EN closure is observed (Fig. 5b–f). The bias behavior in areas 3, 4 and 5 is very similar (Fig. 5c–e), however, different behaviors are found

in areas 2 and 6 (Fig. 5b and f). In the former, all members show similar performance for smaller thresholds, with bias higher (lower) than 1 for thresholds lower (higher) than about 25 mm. The bias in domain 2 was higher (lower) than 1 for thresholds lower (higher) than 15 mm. However, the domain 6 bias exhibits values higher than 2 for thresholds between 0.254 and 19 mm for most of the members, including the EN. Bias between 1 and 2 is observed for thresholds between 25 and 38 mm for most members, and bias lower than 1 for higher thresholds, except for member LO, which exhibits a maximum bias of 1.3.

The effect of FY weighting on lower thresholds is to decrease higher bias, and this is observed in all domains. For higher thresholds, FY improves the values of bias, increasing lower bias, or in other words, decreasing the underestimation. In comparison with the other ensemble members, the increase in skill is most evident in area 6 (Fig. 5f), where in general we observe the ITCZ during January and the precipitation pattern is dominated by convective circulations. The EN and all members (except AS) overestimate precipitation lower than 25 mm. Larger bias is observed in area 6 indicating a deficiency in the forecast precipitation from deep convection mainly over the ITCZ. On the other hand, in this region there are no conventional rain gauge observations, and the precipitation data is provided by satellite estimation. As mentioned in Sect. 5, there are deficiencies in satellite products to estimate precipitation from warm clouds. Thus, the observed precipitation could be underestimated and the bias could be inflated. Area 4 is, like area 6, a preferred location for warm clouds, but has the highest surface observation density. As the MERGE technique combines rain gauge with satellite estimation, the representation of precipitation in tropical regions is best in area 4.

The precipitation weighted by FY showed comparable skill compared with all individual members and the EN closure in area 2 (Fig. 5b – South Brazil), where the precipitation pattern is influenced mainly by synoptic scale systems. The ensemble weighted by FY increases the skill of the model for the full range of the thresholds.

During the warm season, precipitation is dominated by small-scale convective processes in tropical and subtropical regions, that occur with greater frequency at night (Maddox et al., 1979). The models tend to underestimate the higher thresholds and tend to overestimate events of lighter precipitation. Thus, many of the model schemes have a strong high bias. According to Anthes et al. (1989), as the area predicted diminishes in size when increasing precipitation thresholds, it becomes more difficult to forecast. The high values of bias produced by member AS are associated with overestimation of high precipitation amounts, while low values of bias at lower thresholds can be explained by the underestimation of light precipitation, both situations observed in Fig. 3b.

The large variations in most of the model schemes reflect the complexity of the precipitation process and the difficulty

of predicting accurately quantitative precipitation patterns (Anthes et al., 1989), mainly when these patterns are intense.

According to Mason (1989) and Hamill (1999), the ETS can be affected by model bias. Comparing the ETS values from competing forecasts may be misleading if their biases are dissimilar (Mason, 1989). As shown in Fig. 5, similar bias scores are observed at all thresholds, except for member AS. Mason (1989) and Mesinger (1996) have noticed that typically a model with high bias should normally exhibit higher ETS than if the model had less bias. As ETS is affected by the bias of the model, a response of the ETS behavior to the model bias would be expected, in the sense that at lower thresholds, ETS would be higher, and at higher thresholds, ETS would be lower. The results using the ETS agree with previous studies as discussed before, showing higher ETS when higher bias is observed. Even if the ETS has been inflated by higher bias, the ETS showed better skill for the precipitation weighted by FY over all domains (Fig. 6a–f). The highest gain was in the range of small to medium thresholds, but the increase of the skill for higher thresholds was clear.

Unfortunately, in the scope of this paper, it was not possible to verify if the increase of the skill is significant. The FY incorporates more characteristics of the observed rainfall into the new rainfall estimation; therefore the performance of the precipitation simulation was substantially improved.

9 Conclusions

The BRAMS regional atmospheric model incorporates the GD convection scheme, which has several assumptions and closures to simulate sub-grid scale convective precipitation. As a result, the precipitation forecasts can be combined in multiple ways, generating a numerical representation of precipitation and atmospheric heating and moistening rates. This scheme also allows the simple mean ensemble average to be used as the effective simulated rainfall amount and rate. The GD approach also allows optimization methods to be applied in order to determine the optimum combination of the ensemble members which, frequently, is not the simple mean.

In this work, the Firefly optimization method was applied to capture the best combination of the ensemble members in order to minimize the error of the rainfall predicted by the GD scheme. The use of the Firefly algorithm led to a weighted ensemble that captures more of the real precipitation pattern than the GD ensemble simple mean and has more skill than individual ensemble members in most of the precipitation threshold ranges.

Results from one-month simulations using five different closures to generate a new 5-member ensemble of precipitation simulations suggest that the parameter estimation using the Firefly algorithm as an optimization method is a reasonable and potentially powerful methodology to use in producing an ensemble with improved ETS and bias score,

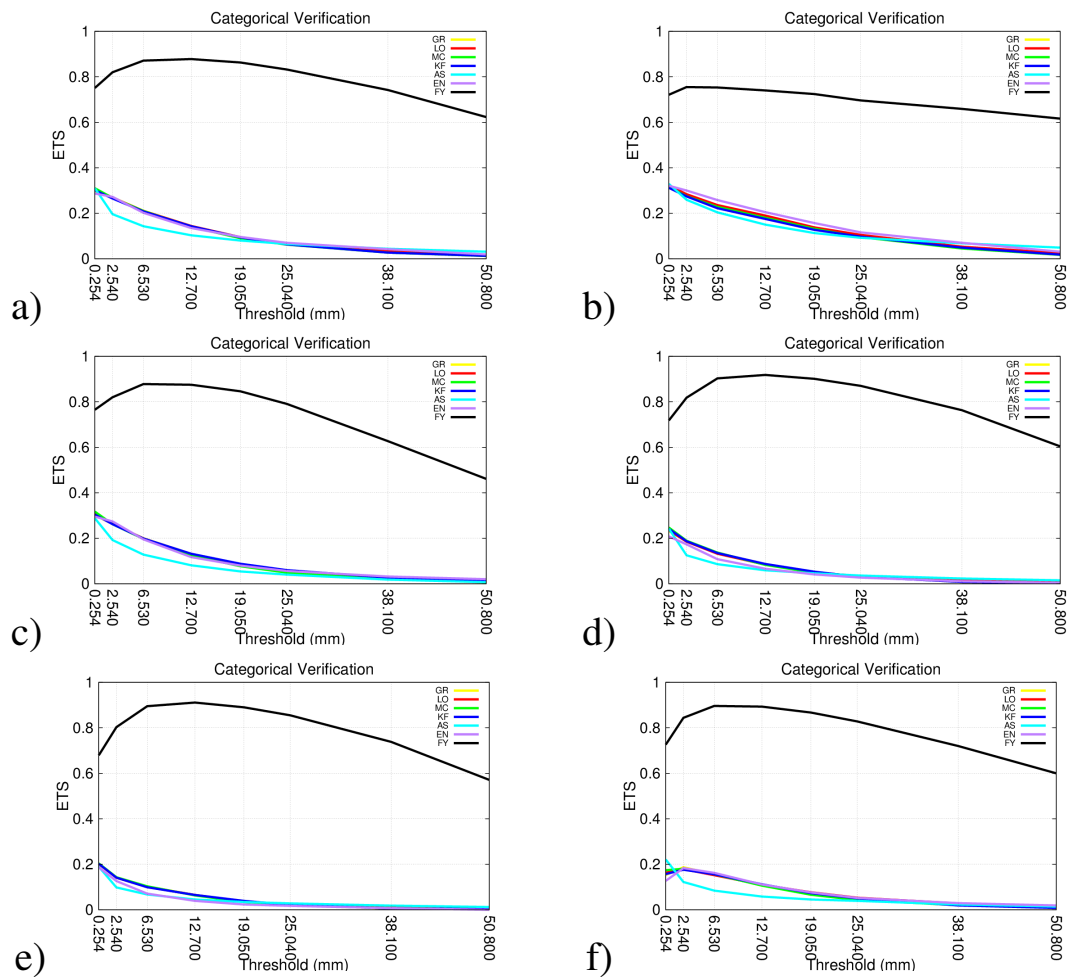


Fig. 6. ETS score vs. precipitation thresholds (legends are the same as Fig. 5).

especially since there are so many uncertainties in the parameterizations used in any numerical model. This methodology is innovative in using the observations as a priori information to generate a set of weights to be applied to the members of an ensemble of physics parameterizations, and the results indicate that the Firefly algorithm is a robust method that can be used for this kind of problem.

In this paper, no independent data set was used to verify the implementation of the Firefly algorithm. The method proposed has demonstrated beneficial hindcast improvement, which is a very important result. On the other hand, we are now working on the application of the use of a priori data to generate a set of weights that can be applied to the members of the ensemble physics of the model BRAMS to improve the forecasting of future events. Thus far, the results are encouraging and we are planning to apply the method to weighting the closure members of the GD convective parameterization of the model BRAMS in order to generate a more realistic precipitation response. With such improvement, we expect not only to improve the precipitation forecasts, but also to

improve the forecasts of other variables, such as temperature, surface pressure and chemical constituents, because convective parameterization plays an important role in the vertical distribution of heat and moisture as well as in the mechanism of vertical transport of pollutants.

Appendix A

Verification measures

In this Appendix, we briefly describe the statistical scores used to verify the results obtained in this paper, according to Wilks (1995).

A1 Bias score

The bias score is the ratio of the number of “yes” forecasts (F) to the number of “yes” observations (O) given a specific criterion, such that

$$\text{Bias} = \frac{F}{O}. \quad (\text{A1})$$

A forecast with no bias has a value of 1. In this paper, we computed the F and O values for each grid point using the EN, each individual ensemble member, the weighted precipitation using FY, and the observations. The bias was calculated for the precipitation thresholds described in Sect. 7.

A2 Equitable threat score

The ETS measures the skill in predicting the area of precipitation amounts over a given threshold with respect to a random forecast (Rogers et al., 1995) and is calculated according to the precipitation thresholds described in Sect. 7. The ETS is defined as:

$$\text{ETS} = \frac{H - CH}{F + O - H - CH}, \quad (\text{A2})$$

where H is the number of correctly simulated points and CH is the expected number of correct points from a random forecast, $CH = (F/V) \cdot O$, where V is a sample size of verification points. A perfect forecast occurs when $\text{ETS} = 1$, and any forecast with $\text{ETS} \leq 0$ has no skill. Forecasts with $\text{ETS} > 0$ have skill relative to a random forecast.

Acknowledgements. This study was supported by the *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP), process 2008/01983-0. Authors Ariane F. dos Santos, Haroldo F. de Campos Velho, Saulo R. Freitas (306340/2011-9), Manoel A. Gan thank the Brazilian Federal (Conselho Nacional de Desenvolvimento Científico e Tecnológico – National Counsel of Technological and Scientific Development – CNPq and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES) and State (FAPESP) agencies for supporting this study. Thanks are also due to Evelyn Grell for manuscript review as well as the anonymous reviewers for their helpful scientific contributions.

Edited by: K. Tokeshi

Reviewed by: A. B. Siebert and another anonymous referee

References

- Anthes, R. A., Kuo, Y.-H., Hsie, E.-Y., Low-Nam, S., and Bettge, T. W.: Estimation of skill and uncertainty in regional numerical models, *Q. J. Roy. Meteorol. Soc.*, 115, 763–806, doi:10.1002/qj.49711548803, 1989.
- Apostolopoulos, T. and Vlachos, A.: Application of the Firefly Algorithm for Solving the Economic Emissions Load Dispatch Problem, *Int. J. Comb.*, 2011, 1–23, doi:10.1155/2011/523806, 2011.
- Arakawa, A.: The Cumulus Parameterization Problem: Past, Present, and Future, *J. Climate*, 17, 2493–2525, 2004.
- Arakawa, A. and Schubert, W. H.: Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part I, *J. Atmos. Sci.*, 31, 674–701, 1974.
- Bao, J.-W., Grell, E. D., Grell, G. A., and Michelson, S.: Application of the Grell-Devenyi Convective Parameterization in Ensemble QPFs, available at: <http://www.mmm.ucar.edu/wrf/users/workshops/WS2007/abstracts/p1-25Bao.pdf>, last access: 9 January 2011.
- Brown, J. M.: Mesoscale Unsaturated Downdrafts Driven by Rainfall Evaporation: A Numerical Study, *J. Atmos. Sci.*, 36, 313–338, 1979.
- Cane, D. and Milelli, M.: Multimodel SuperEnsemble technique for quantitative precipitation forecasts in Piemonte region, *Nat. Hazards Earth Syst. Sci.*, 10, 265–273, doi:10.5194/nhess-10-265-2010, 2010.
- Casanova, S. and Ahrens, B.: On the Weighting of Multimodel Ensembles in Seasonal and Short-Range Weather Forecasting, *Mon. Weather Rev.*, 137, 3811–3822, 2009.
- Casarin, D. P. and Kousky, V. E.: Anomalias de precipitação no sul do Brasil e variações na circulação atmosférica, *Precipitation anomalies over Southern Brazil and variations in the atmospheric circulation*, *Rev. Bras. Meteor.*, 1, 83–90, 1986.
- Chou, S. C. and da Silva, M. G. A. J.: Objective Evaluation of the Eta Model Precipitation Forecasts over South America, *Revista Climanalise*, 01, 1–17, 1999.
- Chou, S. C., Tanajura, C. A. S., Xue, Y., and Nobre, C. A.: Validation of the coupled Eta/SSiB model over South America, *J. Geophys. Res.*, 107, 56(1)–56(21), 2002.
- Climanalise: *Climanalise: Boletim de Monitoramento e Análise Climática*, São José dos Campos, 2006.
- Emanuel, K. and Raymond, D. J. (Eds.): *The Representation of Cumulus Convection in Numerical Models of the Atmosphere*, vol. 24 of *Meteor. Mono.*, Amer. Meteor. Soc., Boston, 1993.
- Figueroa, S. N. and Nobre, C. A.: Precipitation distribution over central and western tropical South America, *Climanalise*, 5, 36–45, 1990.
- Frank, W. M.: The Cumulus Parameterization Problem, *Mon. Weather Rev.*, 111, 1859–1871, 1983.
- Frank, W. M. and Cohen, C.: Simulation of Tropical Convective Systems. Part I: A Cumulus Parameterization, *J. Atmos. Sci.*, 44, 3787–3799, 1987.
- 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., and 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, doi:10.5194/acp-9-2843-2009, 2009.
- Gan, M. A. and Moscati, M. C. L.: Estação chuvosa de 2001/02 na Região Centro-Oeste do Brasil, *Rev. Bras. Meteor.*, 18, 181–194, 2003.
- Gan, M. A., Kousky, V. E., and Ropelewski, C. F.: The South America Monsoon Circulation and Its Relationship to Rainfall over West-Central Brazil, *J. Climate*, 17, 47–66, 2004.
- Grabowski, W. W. and Smolarkiewicz, P. K.: CRCP: a Cloud Resolving Convection Parameterization for modeling the tropical convecting atmosphere, *Physica D*, 133, 171–178, 1999.
- Grell, A. G.: Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations, *Mon. Weather Rev.*, 121, 764–787, 1993.

- Grell, G. A. and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophys. Res. Lett.*, 29, 38(1)–38(4), doi:10.1029/2002GL015311, 2002.
- Grell, G. A., Kuo, Y.-H., and Pasch, R.: Semi-prognostic tests of cumulus parameterization schemes in the middle latitudes, *Mon. Weather Rev.*, 119, 5–31, 1991.
- Hamill, T. M.: Hypothesis Tests for Evaluating Numerical Precipitation Forecasts, *Weather Forecast.*, 14, 155–167, 1999.
- Herdies, D. L., Silva, A., Silva Dias, M. A. F., and Ferreira, R. N.: Moisture budget of the bimodal pattern of the summer circulation over South America, *J. Geophys. Res.*, 107, 42(1)–42(-10), 2002.
- Houtekamer, P. L., Lefavre, L., Derome, J., Ritchie, H., and Mitchell, H. L.: A System Simulation Approach to Ensemble Prediction, *Mon. Weather Rev.*, 124, 1225–1242, 1996.
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scale, *J. Hydrometeorol.*, 8, 38–55, 2007.
- IAC: Microphysical processes of warm clouds, http://www.iac.ethz.ch/edu/courses/bachelor/vertiefung/atmospheric_physics/Script-2012/Script-warm-microphysics.pdf, last access: 6 June 2013.
- Jones, T. R. and Randall, D. A.: Quantifying the limits of convective parameterizations, *J. Geophys. Res.*, 116, 1–19, 2011.
- Kain, J. S. and Fritsch, J. M.: The role of the Convective “Trigger Function” in Numerical Forecasts of Mesoscale Convective Systems, *Meteorol. Atmos. Phys.*, 49, 93–106, doi:10.1007/BF01025402, 1992.
- Kain, J. S. and 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.
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P.: Optimization by Simulated Annealing, *Science*, 220, 671–680, doi:10.1126/science.220.4598.671, 1983.
- Kodama, Y. M.: Large-scale common features of subtropical precipitation zones (the Baiu frontal zone, the SPCZ, and the SACZ) Part-I: Characteristics of subtropical frontal zones, *J. Meteorol. Soc. Jpn.*, 70, 813–836, 1992.
- Krishnamurti, T. N., Low-Nam, S., and Pasch, R.: Cumulus Parameterization and Rainfall Rates II, *Mon. Weather Rev.*, 111, 815–828, 1983.
- Krishnamurti, T. N., Kishtawal, C. M., LaRow, T. E., Bachiochi, D. R., Zhang, Z., Williford, C. E., Gadgil, S., and Suren-dran, S.: Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble, *Science*, 285, 1548–1550, doi:10.1126/science.285.5433.1548, 1999.
- Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., LaRow, T., Bachiochi, D., Williford, E., Gadgil, S., and Surendran, S.: Multimodel Ensemble Forecasts for Weather and Seasonal Climate, *J. Climate*, 13, 4196–4216, doi:10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2, 2000.
- 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.
- Kuo, H. L.: Further studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow, *J. Atmos. Sci.*, 31, 1232–1240, 1974.
- Lin, J. W.-B. and Neelin, J. D.: Toward stochastic deep convective parameterization in general circulation models, *Geophys. Res. Lett.*, 30, 11(1)–11(4), doi:10.1029/2002GL016203, 2003.
- Lorenz, E. N.: Deterministic Nonperiodic Flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
- Lorenz, E. N.: A study of the predictability of a 28-variable atmospheric model, *Tellus*, 17, 321–333, 1965.
- Lukasik, S. and Zak, S.: Firefly Algorithm for Continuous Constraint Optimization Tasks, in: *Proceedings of the International Conference on Computer and Computational Intelligence*, vol. 5796, Wroclaw, Poland, 97–106, 2009.
- Luz, E. F. P.: Meta-heurísticas paralelas na solução de problemas inversos, *Doutorado em computação aplicada*, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, 2012.
- Luz, E. F. P., Becceneri, J. C., and de Campos Velho, H. F.: Conceitualização do algoritmo vagalume e sua aplicação na estimativa de condição inicial da equação de calor, in: *Anais do XI Workshop de Computação Aplicada do Laboratório Associado de Computação e Matemática Aplicada, XI WORCAP*, São José dos Campos, SP, 2009.
- Maddox, R. A., Chappell, C. F., and Hoxit, L. R.: Synoptic and Meso- α Scale Aspects of Flash Flood Events, *B. Am. Meteorol. Soc.*, 60, 115–123, 1979.
- Mason, I.: Dependence of the Critical Success Index on sample climate and threshold probability, *Austr. Meteorol. Mag.*, 37, 75–81, 1989.
- Mesinger, F.: Improvements in Quantitative Precipitation Forecasting with the Eta Regional Model at the National Centers for Environmental Prediction: The 48-km Upgrade, *B. Am. Meteorol. Soc.*, 77, 2637–2649, 1996.
- Mullen, S. L. and Baumhefner, D. P.: Monte Carlo Simulations of Explosive Cyclogenesis, *Mon. Weather Rev.*, 122, 1548–1567, 1994.
- Nobre, P. and Strukla, J.: Variations of Sea Surface Temperature, Wind Stress and Rainfall over the Tropical Atlantic and South America, *J. Climate*, 9, 2464–2479, 1996.
- Randall, D., Khairoutdinov, M., Arakawa, A., and Grabowski, W.: Breaking the Cloud Parameterization Deadlock, *B. Am. Meteorol. Soc.*, 84, 1547–1564, 2003.
- Rao, V. B. and Hada, K.: Characteristics of Rainfall over Brazil: Annual Variations and Connections with the Southern Oscillation, *Theor. Appl. Climatol.*, 42, 81–91, 1990.
- Reynolds, C. A., Webster, P. J., and Kalnay, E.: Random Error Growth in NMC’s Global Forecasts, *Mon. Weather Rev.*, 122, 1281–1305, 1994.
- Rogers, E., Deaven, D., and Dimego, G. J.: The Regional Analysis System for the Operational “Early” Eta Model: Original 80-km Configuration and Recent Changes, *Weather Forecast.*, 10, 810–825, 1995.
- Rozante, J. R., Moreira, D. S., de Goncalves, L. G. G., and Vila, D. A.: Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America, *Weather Forecast.*, 25, 885–894, doi:10.1175/2010WAF2222325.1, 2010.

- Santos, A. F., Velho, H. F. d. C., de Mattos, J. G. Z., Freitas, S. R., Gan, M. A., Passos, H. L., and Luz, E. F. P.: A parametric study for firefly algorithm by solving an inverse problem for precipitation field estimation, in: Proceedings 1st International Symposium on Uncertainty Quantification and Stochastic Modeling (Uncertainties 2012), <http://urlib.net/dpi.inpe.br/pluta0/2012/06.21.20.37>, last access: 17 January 2013, 1–9, 2012.
- Santos, A. F., de Campos Velho, H. F., Luz, E. F. P., Freitas, S. R., Grell, G., and Gan, M. A.: Firefly optimization to determine the precipitation field on South America, *Inverse Probl. Sci. Eng.*, 21, 451–466, doi:10.1080/17415977.2012.712531, 2013.
- Stensrud, D. J., Bao, J.-W., and 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.
- Thompson, P. D.: How to Improve Accuracy by Combining Independent Forecasts, *Mon. Weather Rev.*, 105, 228–229, 1977.
- Tribbia, J. J. and Baumhefner, D. P.: The Reliability of Improvements in Deterministic Short-Range Forecasts in the Presence of Initial State and Modeling Deficiencies, *Mon. Weather Rev.*, 116, 2276–2288, 1988.
- Tripoli, G. J. and Cotton, W. R.: The Colorado State University three-dimensional cloud/mesoscale model, Part I: General theoretical framework and sensitivity experiments, *J. de Rech. Atmos.*, 16, 185–220, 1982.
- Waliser, D. E. and Gautier, C.: A Satellite-derived Climatology of the ITCZ, *J. Climate*, 6, 2162–2174, 1993.
- Walko, R. L., Band, L. E., Baron, J., Kittel, T. G. F., Lammers, R., Lee, T. J., Ojima, D., Pielke, R. A., Taylor, C., Tague, C., Tremback, C. J., and Vidale, P. L.: Coupled Atmosphere-Biophysics-Hydrology Models for Environmental Modeling, *J. Appl. Meteorol.*, 39, 931–944, 2000.
- Wilks, D. S.: Statistical methods in the atmospheric sciences – an introduction, vol. 2 of International Geophysics Series, 59, 1st Edn., Academic Press, San Diego, 1995.
- Yanai, M., Esbensen, S., and Chu, J.-H.: Determination of Bulk Properties of Tropical Cloud Clusters from Large-Scale Heat and Moisture Budgets, *J. Atmos. Sci.*, 30, 611–627, 1973.
- Yang, X.-S.: Nature-Inspired Metaheuristic Algorithms, Luviner Press, Cambridge, 2008.
- Yang, X. S.: Firefly algorithms for multimodal optimization, in: Proceedings of the Stochastic Algorithms: Foundations and Applications, vol. 5792, Sapporo, Japan, p. 178, 2009.
- Yang, X. S.: Firefly Algorithm, Stochastic Test Functions and Design Optimisation, *Int. J. Bio-Inspired Comput.*, 2, 78–84, 2010.
- Zhou, J. and Lau, K. M.: Does a Monsoon Climate Exist over South America?, *J. Climate*, 11, 1020–1040, 1998.