Analysis of the Breeding Technique applied to the CPTEC-AGCM Model

Luis F. Salgueiro Romero[1], Sandra A. Sandri and Haroldo F. de Campos Velho
National Institute for Space Research, São José dos Campos, SP, Brazil

Abstract

The increase of understanding of the climate and weather processes also supported with the evolution of the computation power has led in a systematic improvement on the Numerical Weather Prediction, with highly positive impact on the society activities. The NWP is becoming a country-policy concern level. The models developed to simulate weather behaviour became more accurate and complex, but the errors in the initial conditions will be propagated during the forecasting process. One approach to previous evaluation about the forecasting reliability can be addressed by the breeding technique, which consist in the generation and rescaling of Bred Vectors (BV).

Bred Vectors are the difference between reference and perturbed simulations of the same model, after a time inteval of integration, measured with a choosen norm. The BV are peridocally re-scaled to be the same size of the initial perturbation for restarting the process. The evolution of the BV magnitude could be used to evaluate predictability and sensibility of variables.

The breeding technique will be applied to the Atmospheric General Circulation Model (AGCM) from the CPTEC-INPE to evaluate the goodness of the prediction.

Keywords: Bred vector, Predictability, atmospheric general circulation models.

1 Numerical Weather Prediction

Numerical Weather Prediction (NWP) is a growing science with a remarkable progress in the last 4 decades, where with the advent of supercomputers, had led the improvement of the models generated for research and operations to reach a level of huge importance for our lives.

Since in the beginning of last century, when Vilhelm Bjerknes [1] publish a paper about the possibility of doing weather forecast, introducing concepts of hydrodynamics and thermodynamics into meteorology to solve a set of

[1]E-mail Corresponding Author: luis.romero@inpe.br

non-linear partial differential equations where the initial value problem can be managed by observations of the true state of the atmosphere. [2]

Not until a few years later, Lewis Fray Richardson took this theory and performed a 6 Hs. forecast of the evolution of the pressure, by hand and after weeks of hard work, he came with totally wrong results. Later, he published a book (in 1922) about his results and the difficulty to do practical NWP. [2, 3].

Not until around the 50's that, with John Von Newmann in charge in the Institute for Advance Study in Princeton, it was demonstrated that the NWP was possible by producing the first 24-hours forecast using the ENIAC computer. NWP gain popularity in the scientific community and by the end of the decade, the forecast produced began to show steadily increasing and useful skill. [2].

Nowadays, improvement of skills and accuracy of the different models arrive as an effect of the increasing computation power which permits the possibility of having high-resolutions models, accurate representations of physical processes, high availability of data to represent the true state of the atmosphere, etc.

In Brazil since 1994, The CPTEC (Centro de Previsão de Tempo e Estudos Climáticos) has increased its role in NWP developing a new Global Atmospheric model know as BAM (Brazilian Global Atmospheric Model) in operation since January-2016. With this tool weather events such as rains, flooding and others will be more accurate and help to mitigate the negative impacts in the society and economy.

2 Uncertanties in the model

At the beginning, NWP has been used to describe the simulation of the processes in the atmosphere even if the models are used for research or operational purpose [4]. Today, we have models that also include oceanic and land processes coupled together and this interaction has an improvement in the understanding and skills of the forecast.

The Modelling of the governing processes as a set of non-linear differential equation (knows as the primitive equations) consist only an approach about the true state of the atmosphere where the complexity of the problem is far from being a closed problem.

The chaotic behaviour, proved by Lorenz [5], of the atmospheric dynamics shows a sensitivity to initial conditions, and, as the true state of the atmosphere is never known for sure (due errors in the measure instrumentation, calibrations, others) this help in the increment of the complexity of the model and limits the predictability [2] of the weather in a period of approximately two weeks at most.

Dealing with a such a complex model where the initial condition is subject to errors, the equations of the model aren't known completely and it has a chaotic behavior, has all the kind of unavoidable uncertainty sources that we can mention. As described in [3], "even if the model is perfect and even if the initial condition is almost known perfectly, like any dynamic system with instabilities still has a finite limit of predictability".

Another focus of uncertainties are numerical round-off, the different numerical schemes and even the compiler used that represent the numbers in machine code have an impact in this type of systems.

3 Atmospheric Global Circulation Model - AGCM

Kalnay in [3] categorized the models in two categories, regional and global models. Regional Models are used for short-range forecast (typically 1-3 days) with high resolution discretization, even two or more times higher than Global models, which are used generally for guidance in medium-range forecast (more than 2 days) and for climate simulations.

The CPTEC has both models, the regional BRAMS (Braziliam Regional Atmospheric Modeling System) and the AGCM-CPTEC Global model. This work use the AGCM model, which has a T126 horizontal configuration of spectral discretization with triangular truncation (approx 105 km of space step) and a 28 vertical layer with a time step of 1200 seconds.

The Global AGCM-CPTEC model operates with a higher spectral resolutions but it was chosen T126 because it has to be in agreement with the Ensemble model used to contrast the results of the breeding technique (see section 5) The principal processes including in the AGCM-CPTEC model are hydrodynamic processes (movement equations, horizontal diffusions, etc.), superficial physics of the ocean and land, vertical diffusions, deep convection and the initial (spectral analysis, initialization) and boundary conditions. [6].

The prognostic variables of the AGCM-CPTEC model are the Surface pressure logarithm, the vorticity, the divergence of the horizontal wind, the

[2]The term predictability, as defined in [4], consist of the time required for solutions from two models that are initialized with slightly different initial conditions to diverge to the point where the objective (e.g. RMS) difference is the same as that between two randomly chosen observed states of the atmosphere.

virtual temperature and the specific humidity also with the topography. The models has three interconnected parts:

- Pre-processing, dealing with the initial set-up of the variables for the model
- The model, the core of the system which has the numerical equations
- Pos-processing, necessary to perform the transform of the model output in spectral coordinates to geographic coordinates for visualizing the results.

4 Ensemble Prediction Model

The Ensemble prediction model is a technique used for overcoming the sources of error or uncertainties mentioned in 2, by sampling the error space associated and performing several simulations of the model with different initial conditions values.

Each simulation with a particular initial condition is known as a member of the ensemble and in the overall, the ensemble method is most useful that an individual, deterministic forecast [4].

Using statistics, i.e. mean, spread and variance, the accuracy is improved and the difference in the members is an indication of the quantitative uncertainty of the forecast.

For example the Mean of a X-variable (Temperature, Moisture, Wind, etc.):

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n},$$

and spread (Standard Desviation) of X:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})}{n-1}}$$

In the CPTEC, The forecast for ensemble method is operative since 2001, with EOF (Empiric Orthogonal Function) as a method to develop the several initial conditions. The Ensemble Model operates in a T126 resolution with 28 vertical layer and 15 (fifteen) members, where the perturbations are added (7 members) and subtracted (7 members) consistently to the control initial condition.

The operation of the Ensemble forecast starts for every day at 00 UTC and at 12 UTC and perform a forecasting for 15 days in advance. Figure

1 shows the variability of the Ensemble Surface Temperature Global Mean and its spreads (the standard deviation) starting in 02 of December of 2014 at 00 UTC and integrating to 17 of December of 2014 at 00 UTC. It worth to mention about the variability of the mean and the growing spread as a synonymous of a decay in the reliability of the forecast.

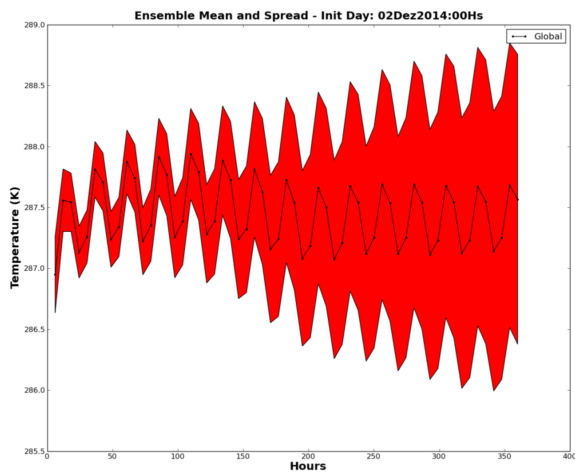
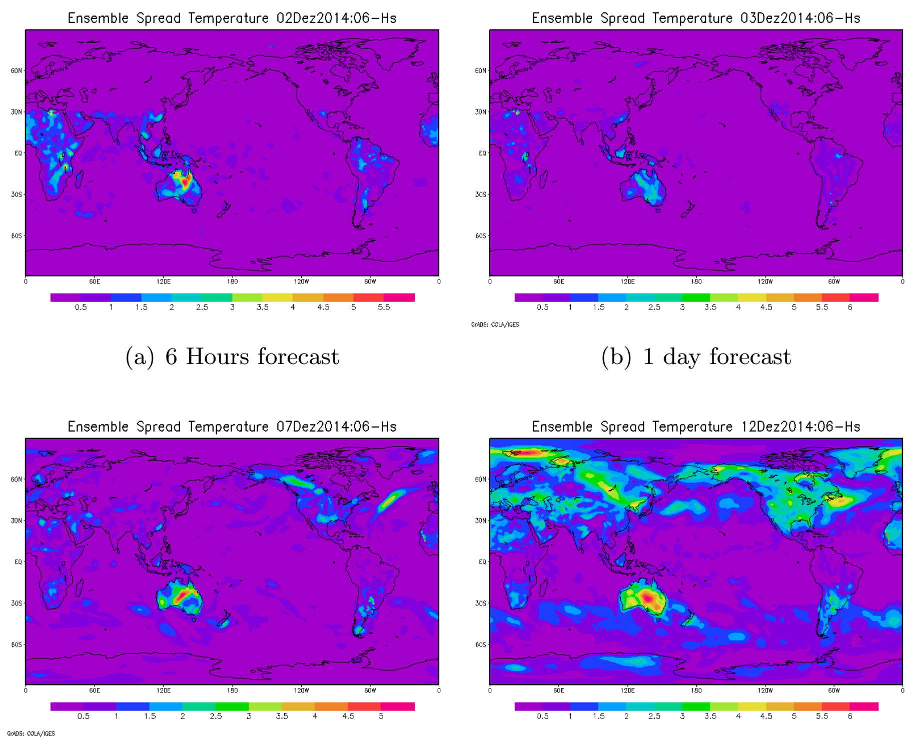


Figure 1: Ensemble Forecast of the AGCM-CPTEC model. The Ensemble Global Surface Temperature mean in black and surrounded by the spread of the 15 members in red. Simulation performed for 15 days (360 Hours)

Another view of the evolution of the Global Temperature Spread in the Ensemble Prediction is shown in Figure 2. As can be noticed, the spread of the Ensemble grows in the north hemisphere and in the region of Australia with a high amplitude, also but with a lower amplitude in the south-America tropical region. This growing amplitude of the spread can be associated with the loss of reliability of the prediction in those areas.



(a) 6 Hours forecast

(b) 1 day forecast

(c) 5 days forecast

(d) 10 days forecast

Figure 2: Global Spread of the Temperature.

5 The Breeding Method

The breeding method has been used since 1992 to generate perturbation for ensemble forecasting at the NCEP, where the method simulates the development of growing errors in the analysis cycle [7]. This method, developed by Toth and Kalnay, has an important property that all random perturbation assume the structure of the Leading Local Lyapunov Vectors (LLLV) after a transient period (approx. 3 days).

This growing errors, associated with the evolving state of the atmosphere, dominate the subsequent forecast errors growth [7].

The Breeding technique consists in the integration of the same non-linear model twice, beginning with different initial conditions. The steps, described in [7], are the following (see fig. 3):

1. Add a small, arbitrary perturbation to the analysis initial state.

2. Integrate both model (control and perturbed) with its initial conditions for a short period ΔT (time step, normally 6 hours)

3. Subtract one forecast from the other. This difference in a variable state is called Bred Vectors.

4. Scale down the difference so that it has the same size (choose a norm, e.g. RMS amplitude) as the initial perturbation.

5. The perturbation resized is added to the analysis corresponding to the following time step and repeat the cycle forward in time.

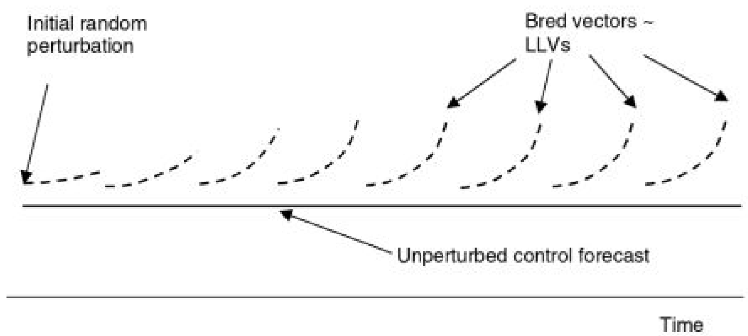


Figure 3: Schematic of the Breeding Method. Source:[3]

Since the construction of the Bred Vector is related to the Lyapunov Vectors, the perturbation in space and time are similars after the transient period. This property is used to compare with the errors showed in figure 2 by the ensemble forecast.

6 Experiments and results

It was performed a simulation of the breeding technique with an initial uniform perturbation in the Virtual Temperature Analysis increased by 0.1% in all the globe. The Initial date was the 02 of December of 2014 at 00 UTC.

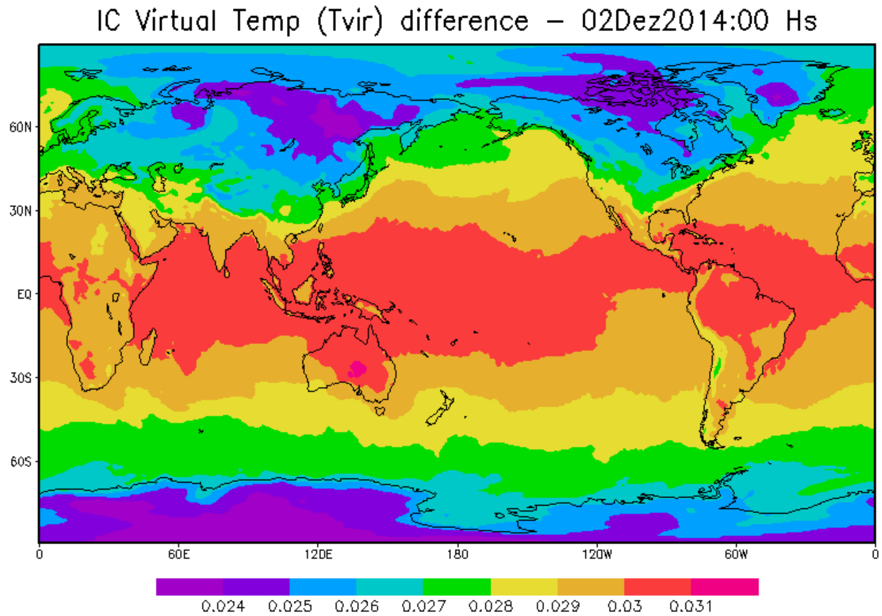
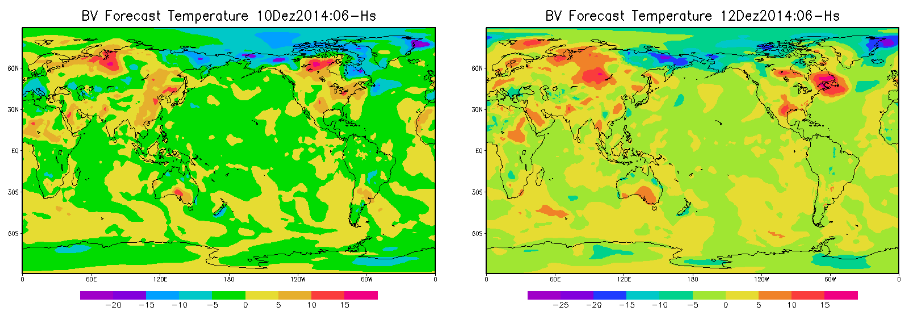


Figure 4: Initial Condition (I.C.) difference of control and perturbed model in the Virtual Temperature Analysis at the surface.

The chosen norm was the rms of the difference of temperature at the entire grid:

$$Norm = \frac{\sum_{i=1,j=1}^{i=nlat,j=nlon}[Temp - Ctrl_{i,j} - Temp - Pert_{i,j}]^2}{nlat * nlon},$$

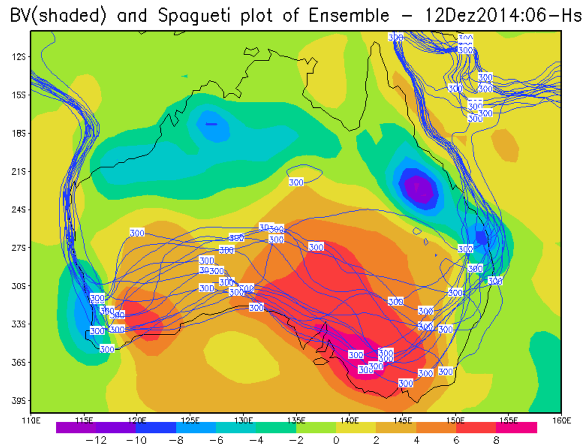
where $nlat$, $nlon$ are the total discretization points of latitude and longitude coordinates. The size of the initial perturbation was 0.0288016.



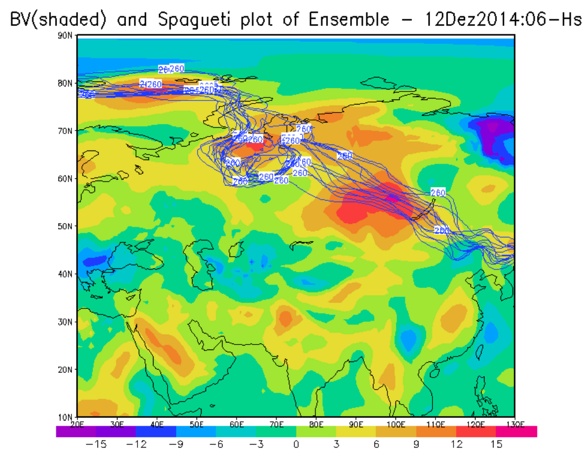
(a) 8 days forecast (b) 10 days forecast

Figure 5: Bred Vectors of the Surface Temperature. Bred Vectors are the difference of the Temperature Forecast of both models, perturbed and unperturbed.

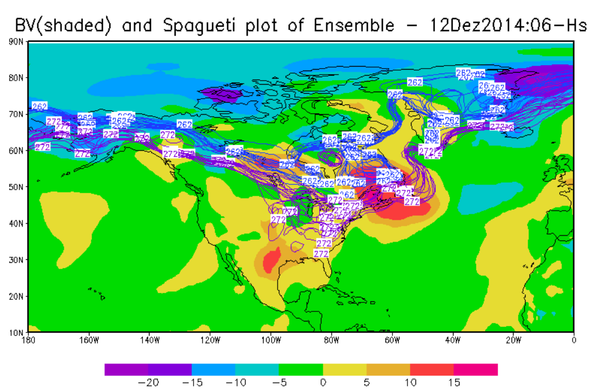
According to Figure 2, the regions of major spread are in Australia, North of Europe and North of America which is also in agreement with the size of major Bred Vectors showed by Figure 5.



(a) Region of Australia



(b) Region of North-Europe



(c) Region of North-America

Figure 6: Bred Vectors of the Surface Temperature and spaghetti plot of the members of the Ensemble Forecasting..

In Figure 6 are shown the results of Bred Vectors of the surface Temperature in shaded colours, also the same figure shows some levels of temperature, in spaghetti plot of the ensemble members, for different regions of the globe.

Using the spaghetti plot method it can be seen the spread of the different members of the ensemble, and can be inferred that in different regions, there is great spread as well as great amplitude of the Bred Vector, so, this consistency shows that Bred Vectors can be used as a tool to measured the reliability of the forecast, as well as the spread of the ensemble forecasting.

7 Conclusion

The breeding technique was applied to the AGCM-CPTEC model generating Bred Vectors that indicate regions with high and low predictability. BV showed that are in good agreement with the spread of the ensemble members and can be used as an estimation of the reliability of the forecast taking into account its simplicity and less cost of generation.

ACKNOWLEDGMENTS

Authors would like to thanks CNPQ for the financial support. LFSR would like to thanks, Dr. Paulo Kubota from the DMD-CPTEC department and Dra. Rosngela Cintra from LAC-INPE, for the support in the initial configurations of the AGCM Model.

References

[1] Scientific Weather Prediction: A Newspaper Article on Weather Forecasting by V. Bjerknes. Consulted in September, 2016. Available in `http://folk.uib.no/ngbnk/Bjerknes_150/short-paper-Bjerknes-1904-Rev.pdf`

[2] The History of Numerical Weather Prediction. Consulted in September, 2016. Available in `http://celebrating200years.noaa.gov/foundations/numerical_wx_pred/welcome.html`

[3] E. Kalnay (2003). Atmospheric Modeling, Data Assimilation and predictability. Cambridge University Press.

[4] Warner, T. T. (2010), Numerical Weather and Climate Prediction, Cambridge University Press, ISBN=9781139494311.

[5] LORENZ, E. N (1963). Deterministic non-periodic flow. J. Atmos. Sci., v. 20, p.130141.

[6] Aravquia J.A., de Quadro M. (2003) ASPECTOS GERAIS DA PREVISO NUMRICA DE TEMPO E CLIMA. INPE. Consulted in september 2016, available in http://mtc-m16b.sid.inpe.br/col/sid.inpe.br/jeferson/2003/12.23.10.21/doc/publicacao.pdf

[7] Z. Toth, E. Kalnay. "Ensemble Forecasting at NCEP and the Breeding Method". Monthly Weather Review 1997 125:12, 3297-3319