

A REGIONAL ENSEMBLE FORECAST SYSTEM FOR SOUTHEASTERN SOUTH AMERICA: PRELIMINARY ASSESSMENT.

Juan José Ruiz^{1,2*}, A. Celeste Saulo^{1,2} and Eugenia Kalnay³

¹Centro de Investigaciones del Mar y la Atmósfera CONICET-UBA

²Departamento de Ciencias de la Atmósfera y los Océanos. FCEyN UBA

³Department of Meteorology, University of Maryland

Abstract

The main objective of this study is to assess the skill of a short range regional ensemble forecast system over Southeastern South America. The proposed regional ensemble forecast system consists of 9 members: 4 pairs of perturbed forecasts plus a control forecast. The initial and boundary conditions are perturbed using scaled lagged forecasts (SLAF) from the Global Forecast System operational runs. These perturbed initial and boundary conditions are used to run a 9-regional forecast ensemble, using WRF V2.0 mesoscale model.

The skill of the ensemble mean is compared against the skill of the control forecast. Also, in order to investigate if the ensemble dispersion could be used to forecast the magnitude and distribution of the control forecast errors, the relationship between these variables is measured over the region.

The root mean squared error between the forecasted fields and the analysis is computed as a function of forecast times in order to assess the skill of both the ensemble and the control forecasts. When available, upper air observations are also used in a similar way.

Results suggest the benefit of adopting an ensemble technique to improve regional forecasts. The verification of other ensemble derived products as the probabilistic quantitative precipitation forecast over this region, and implementation of other methods to generate initial condition perturbations are important issues to be addressed in future work.

1. Introduction

Ensemble forecast has proved to improve medium-range weather forecast (Toth and Kalnay 1993 among others) and also short-range higher resolution weather forecast (Hou et al. 2001, Hamill and Colucci 1997, among others). Ensemble forecasting intends to account for initial conditions (IC) uncertainty and also model deficiencies, estimating the probability distribution of future states of the atmosphere. IC uncertainties are reproduced by adding some kind of perturbation to the analyzed field and to the boundary conditions. Over South America the lack of observations is a critical issue (as can be seen for example in the perturbation mask implemented in Toth and Kalnay 1997) so IC uncertainties could be considered an important source of error in the short and medium range weather forecasts. This lack of data is even more critical in the smaller scales, which can nowadays be resolved in operational numerical weather forecasts. This

motivates the implementation of short-range ensemble techniques over the Southeastern South America (SESA) region.

Another important issue related with numerical weather prediction is the availability of computational resources. As stated by Hamill and Colucci (1997) a 5 km single resolution run roughly needs as much computational power as 16 members 10 km resolution ensemble. Some IC perturbation generation methods also need the performance of extra global models runs to provide consistent boundary conditions to each ensemble member. In this work the generation of IC is performed without any increase in the computational cost.

To perform a preliminary assessment of a short range ensemble forecast over SESA the general characteristics of the ensemble performance are analyzed and compared with similar studies for other regions and with theoretical concepts. The methodology implemented for IC perturbation and ensemble verification is described in section 2. Section 3 shows the general characteristics of the ensemble performance and the verification of the ensemble mean and the control mean for different variables. Conclusions are presented in section 4.

-
- Corresponding author address: Juan Ruiz, Centro de Investigaciones del Mar y de la Atmósfera (CONICET-UBA).
E-mail: jruiz@cima.fcen.uba.ar

2. Methodology

2.1 Model configuration and ensemble description.

In this experiment the Weather Research and Forecasting model (WRF) version 2.0 has been used (Skamarock et. al. 2005) to perform regional ensemble forecasting over SESA. General model settings –described below- were the same for the control forecast and for each of the ensemble members. The domain is centered in the La Plata basin (Figure 1), and the grid follows a Mercator projection with a horizontal resolution of approximately 50 km and 31 vertical sigma levels. The model is run in the non-hydrostatic mode. The microphysics scheme utilized is the Eta Grid-scale Cloud and Precipitation (<http://www.emc.ncep.noaa.gov/mmb/mmbpll/eta12tpb/>). Convection is parameterized using Grell scheme (Grell 1993), RRTM scheme (Mlawer et. al. 1997) and Dudhia (1989) scheme are used to represent radiative fluxes. Mellor and Yamada, (1982) treatment is used in the boundary layer processes parameterization and a four layer surface model is used to represent surface processes (Chen and Dudhia 2000).

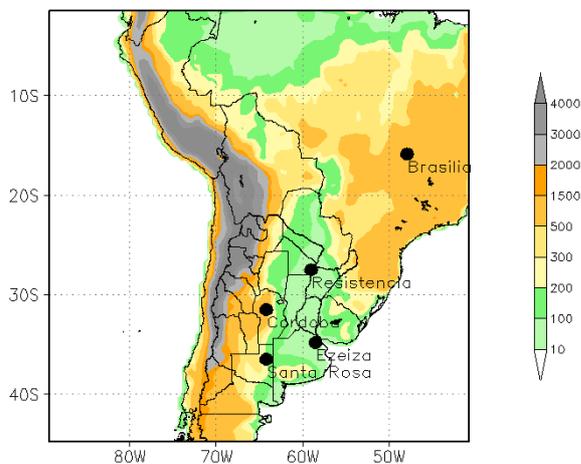


Figure 1: Model domain and topography (shaded, in meters). Location of the upper air stations selected for model verification.

The experimental Short Range Ensemble Forecast (SREF) for SESA consists of 9 members (8 perturbed forecasts plus a control forecast). Different model runs are obtained through the perturbation of initial and boundary conditions using the SLAF technique (Ebisuzaki and Kalnay 1991). This technique was for the first time applied to a regional ensemble forecast during the Storm

and Mesoscale Ensemble Experiment (SAMEX) (Hou et. al. 2001).

In the SLAF method, the perturbations to the boundary and initial conditions are the difference between previous forecasts verifying at the same time and the analysis corresponding to that time (Global Forecasting System (GFS) control run forecasts are used to generate perturbed boundary conditions, instead of the analysis). The perturbations are rescaled because perturbations associated with “older” forecasts are usually larger than the ones associated with “younger” ones. Originally the SLAF technique reduces the amplitude of the perturbation assuming a linear growth of the forecast errors with time. In this experiment the amplitude of the perturbations is reduced so they all have near the same amplitude measured by the mean amplitude in the geopotential height at 200 hPa. Finally, the perturbations are added to and subtracted from the analysis and the GFS control run to obtain the perturbed initial and boundary conditions respectively. This procedure is applied to all atmospheric variables and also to the ground temperature and moisture content.

As described by Hou et. al. (2001), SLAF technique seeks to create dynamically growing perturbations which can be applied both to initial and boundary conditions. These perturbations include “errors of the day”, they get larger amplitudes where the errors in the previous forecasts grew faster due to atmospheric instabilities. One of the major advantages of this technique is that no extra computational cost is required in the generation of the perturbations: only the control run of a global model is needed. In the present case the GFS control runs (with 1°x1° resolution) at 12 and 00 UTC are used to generate the perturbations and to provide the initial and boundary conditions to the control run. This data is available through the National Oceanic and Atmospheric Administration (NOAA) ftp site - <ftp://ftp.ncep.noaa.gov>-. 12 and 00 UTC global forecasts started up to 48 hours before each analysis time are employed to generate the perturbations so each ensemble is composed by four pairs of perturbed members (generated using the global forecasts started 12, 24, 36 and 48 hours before the analysis time).

The experimental SREF for SESA is initialized once a day at 12 UTC. This time is chosen because more observations are available over the region. Nevertheless, it should be mentioned that global forecasts started at 00 UTC also extend their influence over the ensemble through the perturbations generated from them. The perturbed runs have a length of 48 hours while the control run has a total length of 96 hours.

These runs have been performed between October 22nd and November 27th 2005.

2. 2 Methodology for verification of the experimental SREF for Southeastern South America.

The first question to be answered is if the behavior of the different members of the ensemble can represent the forecast error growth. To answer this, the evolution of the ensemble dispersion for different variables has been calculated as a function of forecast time. The relationship between ensemble dispersion and forecast error has been also investigated through the computation of spatial and temporal correlations.

The second question to be answered is if the ensemble mean has a better skill than any single member of the ensemble over the region of interest. To answer this question, the RMSE and the correlation coefficient between the forecast and the GFS analysis have been computed for some selected variables. Also, at 5 sounding stations (Ezeiza 34° 49' S 58° 32' W, Córdoba 31° 25' S 64° 11' W, Santa Rosa 36° 37' S 64° 17' W, Resistencia 27° 27' S 58° 59' W and Brasilia 16° S 48° W) the RMSE and correlation between the forecasts and the observations have been computed.

As the domain covers zones with different precipitation and circulation regimes, it has been divided into two large regions for verification purposes: the Northern region (NR), between 25° S and 0° S, and the Southern region (SR), between 45°S and 20°S. This subdivision roughly accounts for the more baroclinic flow within the westerlies to the south of the domain and the more barotropic environment near tropical areas.

3. Results

3.1 Ensemble spread evolution and its relationship with forecast error.

As in Hou et. al. 2001, the standard deviation or spread of the ensemble forecast members about the ensemble mean is used to measure the amplitude of the perturbation during the forecast length. The ensemble dispersion should grow over the forecast period at a rate similar to the forecast error so the “truth” can appear as a plausible member of the ensemble (Toth and Kalnay 1993).

Figure 2 shows the ensemble dispersion for the geopotential height at 200 and 500 hPa. as

a function of the forecast length averaged over the complete period. The difference between the NR and the SR is clear in the behavior of the ensemble dispersion. Over the NR, the initial perturbations in the geopotential height decreases with forecast length, showing some growth at the end of the forecast period. This behavior suggests that the perturbations generated through the SLAF technique are not well balanced over tropical regions. For the Southern Region (SR) the perturbations exhibit nearly constant growth during the forecast period, similarly to Hou et. al. (2001) results for different single model ensembles. The increase in the growing rate with height and the growing rates are also similar to those reported in that paper. As a result of the different growing rates between SR and NR, the total domain perturbation growing rate is nearly nil throughout the forecast period.

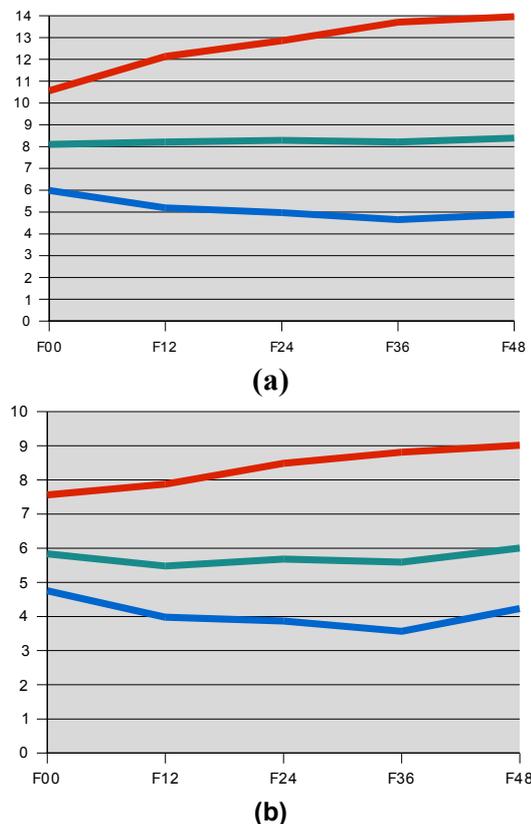


Figure 2: Ensemble dispersion vs. forecast length for (a) Geopotential Height at 200 hPa, and (b) Geopotential Height at 500 hPa. NR in blue, SR in red, and total domain in green.

Significant differences in the evolution of the initial perturbations can be seen in the moisture content at lower levels (Figure 3). Moisture as well as zonal and meridional wind

components (not shown), exhibit perturbation growth in both regions during all forecast periods. Dispersion of moisture at 850 hPa. over the NR, however, also shows an important diurnal cycle which becomes evident because all forecasts start at 12 UTC. During late morning and afternoon (between 12 and 00 UTC) there is a rapid increase of ensemble dispersion in specific humidity at 850 hPa., with some decay of the ensemble dispersion during the night (between 00 and 12 UTC). This behavior can be associated with the strong diurnal cycle of convection over the NR and particularly over continental areas. Forecasted convection generates perturbation growing rates which are faster than in situations where no convection or precipitation processes are taking place. Examples of this sensitivity can be found in Zhang et. al. (2003) and Tan et. al. (2004). As convection is more frequent over the NR, the kind of perturbation growth described by Tan et. al. (2004) could be more active over this region than in the SR where convection is more sporadic. It would be desirable to perform error growth experiments over this particular region in order to find out how individual perturbations grow and how initial perturbations can be improved.

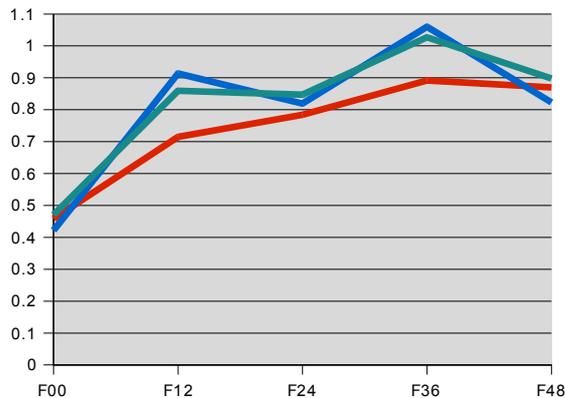


Figure 3: Ensemble dispersion vs forecast time for specific humidity at 850 hPa. NR in blue, SR in red, and total domain in green.

In order to see if the perturbations grow fast enough so that the truth can be a plausible ensemble member, the ensemble mean-square error is computed for the same variables at 24 and 48 hours. The verifying truth in this experiment is the GFS initial condition interpolated to the WRF grid using the WRF Standard Initialization. As stated in Hou et. al. (2001), the spread of the ensemble should be similar to the forecast error of the ensemble mean -for a perfect ensemble-. In the present experiment, the mean squared error (MSE) grows faster than the ensemble dispersion.

In order to analyze what causes MSE growth, the MSE has been decomposed into the square of the domain bias and the square of the standard deviation of the error (SDE) as in Hou et. al. (2001) -Equation 1-. It should be noted that before RMSE computation, the BIAS has been removed from the different fields, producing lower RMSE values when compared with other verifications.

$$MSE = \overline{(f_{i,j} - v_{i,j})^2}^{i,j} \quad \text{(Equation 1)}$$

$$SDE^2 = \overline{(f'_{i,j} - v'_{i,j})^2}^{i,j}$$

Where $v_{i,j}$ is the analysis value for the position i,j , and $f_{i,j}$ is the forecasted value which verifies at the same time. Also $f'_{i,j} = f_{i,j} - \bar{f}^{i,j}$ and $v'_{i,j} = v_{i,j} - \bar{v}^{i,j}$, the over bar denotes spatial average.

Figure 4 shows an example of the relation between the error growth and the dispersion growth for the temperature at 850 hPa. in the SR. As for other variables, the ensemble dispersion grows between 24 and 48 hours, the dispersion is smaller than the SDE of the ensemble mean and the error of the ensemble mean grows faster. This is confirmed by the rank histograms of the variables considered (not shown) which shows maximum frequencies for the first and last categories indicating that the verifying value is often outside the range encompassed by the ensemble members.

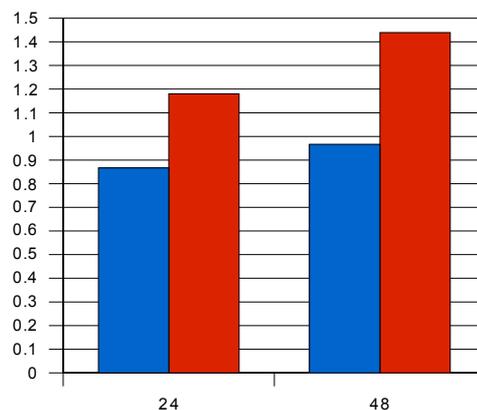


Figure 4: Ensemble dispersion (blue) and SDE (red) of the ensemble mean at 24 and 48 hours for temperature at 850 hPa.

Another important issue is to analyze if there is a relationship between the SDE of the ensemble mean and the ensemble dispersion. It would be desirable that the ensemble dispersion would grow faster in situations where the forecast

errors of the ensemble mean grow faster. If this were the case, the ensemble dispersion would grow faster or slower according to the intrinsic predictability associated with a particular weather situation and the ensemble dispersion could be used as a predictor of the occurrence of larger errors in the forecast. For the present experiment the relationship between SDE and forecast errors is investigated through spatial correlations on a daily basis and through a temporal correlation between the average SDE and forecast errors over SR, NR and the total domain.

The correlation coefficients obtained through the computation of the spatial correlation between the error of the ensemble mean and its dispersion are low -below 0.22, not shown- indicating less than 4% of the total variance explained by this relationship. This suggests that it is not possible to predict the position and/or the shape of the regions that have the largest forecast errors, at least using this ensemble configuration. This also means that the areas where the dispersion of the ensemble for certain variables is large do not necessarily match the areas where the forecast errors are large. Hou et. al. (2001) also reported small pattern correlation between ensemble spread and forecast error for most of the single model ensembles analyzed there. They also show that a superensemble where different ensemble techniques and models were put together reach spatial correlation coefficients around 0.4. They state that this low pattern correlation is because a large spread indicates potential for large errors, but actual errors may be small.

More promising results are found analyzing the relationship between the SDE of the ensemble mean and the spread of the ensemble averaged over a large region. In this case the temporal correlation is performed over the NR, SR and total domain during the complete period. Greater correlation coefficients are found over the SR for the components of the wind at different altitudes (Figure 5).

The results showed in Figure 5 indicate that the ensemble increases its dispersion as the intrinsic predictability of the atmosphere decreases over the SR so an increase of forecast error is associated with larger ensemble dispersion over the region. This is more evident for the wind than for the geopotential height or the temperature. According to Figure 5, 25-36 % of the variance of the forecast error in the U component of the wind can be explained through its relationship with the ensemble dispersion over the SR. For the NR no relationship becomes evident from the results

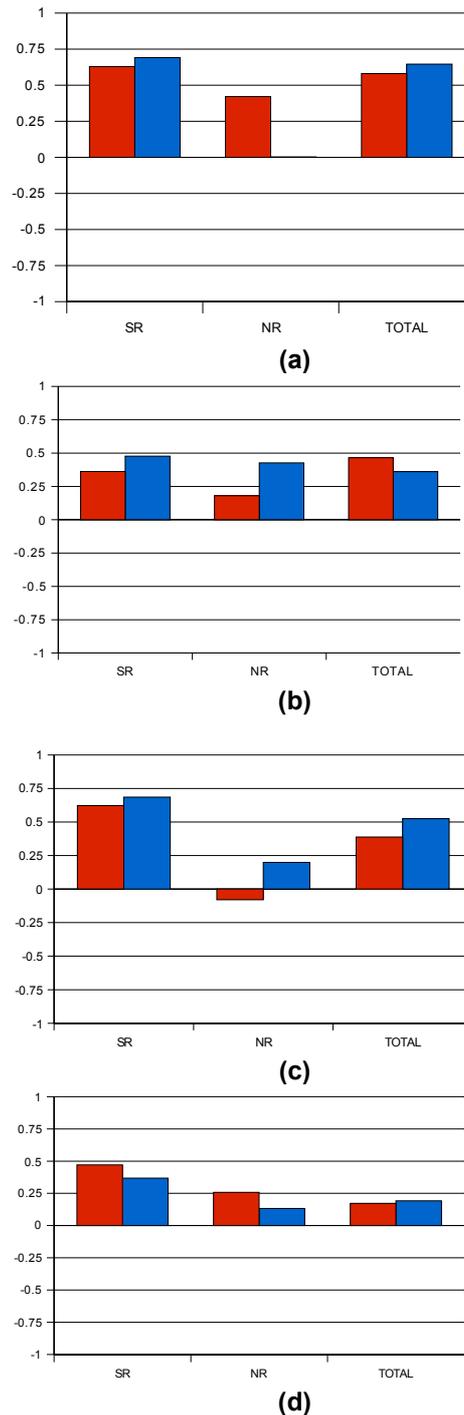


Figure 5: Temporal correlation between SDE of the ensemble mean and the ensemble dispersion averaged over the NR, SR and total domain. 24 hour forecast (red) and 48 hour forecast (blue) for (a) U-wind at 500 hPa. (b) Geopotential Height at 500 hPa., (c) U-wind at 850 hPa., (d) specific humidity at 850 hPa.

3.2 Ensemble and control verification.

showed in Figure 5. In this case the averaged ensemble dispersion can not be used as a predictor of the forecast error over the region.

In this section the RMSE is used to perform forecast verification of the ensemble mean and the control forecast. The verification is done against the GFS analysis interpolated to the

WRF grid and against upper air observations at 5 locations as indicated in the previous section. In all cases the BIAS has been previously removed from the selected variables. Different variables RMSE for the ensemble mean and the control forecast can be seen in Table I.

Variable	Time	Ens. Mean		Ctrl. Forecast	
		SR	NR	SR	NR
u-wind (m/s) 850 hPa.	24 hs.	2.31	2.34	2.69	2.75
	48 hs.	2.78	2.8	3.24	3.22
v-wind (m/s) 850 hPa.	24 hs.	2.53	2.33	2.91	2.7
	48 hs.	3.02	2.68	3.43	3.19
geopotential (m) 500 hPa.	24 hs.	9.83	4.95	10.24	5.59
	48 hs.	14.86	5.51	15.37	6.2
temperature (K) 850 hPa.	24 hs.	1.13	0.73	1.25	0.9
	48 hs.	1.47	0.83	1.61	1
q(g/Kg) 850 hPa.	24 hs.	1.1	0.9	1.3	1.2
	48 hs.	1.3	1.1	1.7	1.4

Table I: Ensemble mean and control forecast RMSE calculated after bias removal for different variables.

Variable	Time	Ensemble Mean					Control Forecast				
		SAEZ	SACO	SARE	SAZR	SBBR	SAEZ	SACO	SARE	SAZR	SBBR
q (g/Kg) 850 hPa.	24 hr	1.68	2.08	2.55	1.69	2	1.9	2.26	2.79	1.82	2.17
	48 hr	1.89	2.4	2.68	1.73	1.82	2.28	2.82	3.07	2.06	2.15
temperature (K) 850 hPa.	24 hr	1.38	2.17	1.35	1.58	1.45	1.88	2.38	1.36	1.81	1.68
	48 hr	1.83	2.69	1.76	1.82	1.45	1.99	2.85	2.04	1.91	1.5
geopotential (m) 500 hPa.	24 hs	15.94	9.64	10.31	10.73	6.39	17	10.98	11.31	11.39	6.97
	48 hs	20.25	13.97	13.92	18.72	6.63	20.29	13.87	14.9	19.32	7
v-wind (m/s) 850 hPa.	24 hs	3.07	5.14	3.65	4.34	3.36	3.23	5.09	3.85	5.06	3.97
	48 hs	3.91	6.1	4.17	4.45	3.34	4.21	6.65	4.54	4.98	3.67

Table II: Ensemble mean and control forecast RMSE -calculated after bias removal- for 5 sounding stations located within the ensemble domain.

The results show that the ensemble mean performs better than the control forecast over both regions. The advantage is greater for the wind components and for the specific humidity where differences between the root mean square of the control run and the ensemble mean are similar in magnitude to the difference between the 48 and

24 hours control forecast, thus indicating that the ensemble improves the forecast quality having as good scores for 48 hr ensemble forecasts as for 24 hr control run forecasts. However, RMSE might not be the best way to compare the skill of an ensemble. The resolution used in this study allows the model to resolve part of the mesoscale.

At this scales, small errors can grow fast as shown by Zhang et. al. (2003), and can increase the magnitude of the RMSE. In the ensemble mean, these smaller and less predictable features are smoothed. This can lower the RMSE of the ensemble mean but not necessarily improve the utility of the forecast. The spectral analysis of the control forecast and the ensemble mean could allow a quantitative measure of this effect and is part of a work in progress.

Five sounding stations have also been used to compare the skill of the control forecast against the ensemble forecast. The results obtained comparing against upper air observations confirm the previous ones: greater improvements can be found in the specific moisture at 850 hPa., on the other hand, the 500 hPa. geopotential height shows less improvement. Comparison against observations and analysis shows that the ensemble mean diminishes RMSE and also that error growth is slower for the ensemble mean than for the control forecast. Values obtained in this work are lower than those obtained in other verifications like that performed by Hamill and Colucci (1997). This difference can be due to the BIAS removal performed before the RMSE computation.

4. Conclusions:

Regional ensemble forecasting is a promising tool for improving weather forecasts over South America. In particular SLAF is one of the less computer demanding techniques although it can generate perturbations for the initial and boundary condition that exhibit adequate growth during the forecast.

For the selected domain it could be identified quite different behaviors between the northern, more barotropic, region and the southern, more baroclinic, region. Over the SR perturbation growth is similar to that found in previous studies like Hou et. al. (2001) but over the NR domain, perturbation growth seems to be more influenced by convection and exhibits a strong diurnal cycle. It is not clear to what extent the growing rate in the NR is controlled by the initial perturbation or by error growth in lower scales associated with convection like those described in Zhang et. al. 2003. In other words: Is the background growth of the dispersion over the NR associated with the up-scaling of the rapid growing mesoscale perturbations or with the slower growth of the larger scale perturbation generated with SLAF?.

Also, over the NR, the correlation between the forecast error and the ensemble dispersion is not significant. In this region the MSE shows less

temporal variability than over the SR and the errors over the NR could be more dominated by smaller scale features than over the SR. This suggests that the ensemble technique might not be as useful for the NR as for the SR or that the generation of perturbations should be improved for this region. To understand the behavior of error growth over the NR perturbation growth experiments such as those performed by Zhang et. al. 2003 should be carried out.

The RMSE computed for different variables shows that the ensemble mean verifies better than the control run. However, the RMSE computed for the ensemble mean may not be a good estimation of the ensemble forecast quality: average over ensemble members could produce smoothing of mesoscale features over the forecast domain. This is because these particular features are smaller in scale and are highly intrinsically unpredictable so phase shifts of the order of half wave-length or more are easily found between the different members of the ensemble.

Although the ensemble mean shows lower RMSE values, the mean of the ensemble could not be a good tool for forecasters since mesoscale features can be drastically smoothed. The information provided through ensemble forecasting should be synthesized in terms of probability of occurrence of a particular kind of phenomena, as is done in the probabilistic quantitative precipitation forecast. For this particular experiment probabilistic quantitative precipitation forecasts are going to be verified using raingauge data and remote sensing precipitation estimates.

Also the implementation of other techniques for generating initial and boundary conditions should be performed to compare with the results obtained with SLAF.

Acknowledgements: *This research is sponsored by the Research Grants UBACyt X155 PIP 5417 and IAI-CRN055*

References:

Dudhia J., 1989: **Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model.** *Journal of the Atmospheric Sciences*: Vol. 46, pp. 3077–3107.

Ebisuzaki, W., and E. Kalnay, 1991: **Ensemble experiments with a new lagged average forecasting scheme.** *WMO Research Activities in Atmospheric and Oceanic Modeling*: Rep. 15, 308 pp.

- Chen, S.-H., and J. Dudhia, 2000: **Annual report: WRF physics**, Air Force Weather Agency, 38pp.
- Grell G. 1993: **Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations**. *Monthly Weather Review*: Vol. 121, pp. 764–787.
- Hamill T. M. and Stephen J. Colucci. 1997: **Verification of Eta-RSM Short-Range Ensemble Forecasts**. *Monthly Weather Review*: Vol. 125, pp. 1312–1327.
- Hou D., Eugenia Kalnay and Kelvin Droegemeier: **Objective Verification of the SAMEX' 98 Ensemble Forecast**. *Monthly Weather Review*: Vol. 129, pp 73-91.
- Mellor G. and Yamada T.: **Development of a turbulence closure model for geophysical fluid problems**. *Reviews of Geophysics and space Physics*; 20 (4), 851-875.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: **Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long-wave**. *J. Geophys. Res.*: 102(D14), 16663-16682.
- F. Zhang, Chris Snyder and Richard Rotunno. 2003: **Effects of Moist Convection on Mesoscale Predictability**. *Journal of the Atmospheric Sciences*: Vol. 60, pp. 1173–1185.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: **A description of the Advanced Research WRF Version 2**. *NCAR Tech Notes-468+STR*
- Tan Z., Fuqing Zhang, Richard Rotunno and Chris Snyder. 2004: **Mesoscale Predictability of Moist Baroclinic Waves: Experiments with Parameterized Convection**. *Journal of the Atmospheric Sciences*: Vol. 61, pp. 1794–1804.
- Toth Z. and Eugenia Kalnay. 1993: **Ensemble Forecasting at NMC: The Generation of Perturbations**. *Bulletin of the American Meteorological Society*: Vol. 74, pp. 2317–2330.
- Toth Z. and Eugenia Kalnay. 1997: **Ensemble Forecasting at NCEP and the Breeding Method**. *Monthly Weather Review*: Vol. 125, pp. 3297–3319.