Ensemble dispersion spectra and the estimation of error statistics for a limited area model analysis

Simona Stefanescu, NIMH Romania
Scientific supervisor: Loik Berre, Meteo-France

29th September 2003

ABSTRACT

This study presents the derivation of the forecast error covariances using the ensemble approach. As an alternative to the NMC method, a randomization technique is used to estimate the forecast error covariances. Using perturbed observations as input to the assimilation scheme, an ensemble of perturbed analysis is generated. The differences between two perturbed analysis, which are a measure of the uncertainties on the analysis, are propagated during the 6h forecast in the form of uncertainties in the first guess. Differences between analysis or first guess fields from the ARPEGE coupling model, valid for the same dates and times, but for different members of the ensemble, are used to produce the forecast error statistics for limited area model ALADIN. The computation of forecast error statistics using the 6h forecasts of ALADIN, starting from ARPEGE perturbed initial and boundary data, is also included in this study. A complete description of forecast error statistics should account also for the model error. The statistics computed from differences between the ALADIN 6h forecast and the ARPEGE analysis or first guess, for the same member of the ensemble, gives some information about the model errors. Finaly, the statistics computed from differences between different models and for different members of the ensemble is supposed to give a better estimation of the forecast errors. A comparison with the statistics derived through the NMC method is presented too.

It was found that the ensemble approach produces reliable good estimation of the forecast errors.

Contents

1 Introduction **3**

2 Forecast error statistics derived through the NMC method **5**
2.1 Description of the datasets 5
2.2 Results obtained with standard and lagged NMC methods 5

3 Estimation of the error statistics using the ensemble prediction approach **6**
3.1 Description of the datasets 6
3.2 Contributions to the evolution of dispersion spectra in a perfect-model framework 7
3.2.1 *Effect of ARPEGE analysis* 7
3.2.2 *Effect of DFI applied to the ARPEGE fields* 8
3.2.3 *Effect of the ALADIN 6h forecast* 8
3.2.4 *Comparison between the four basic dispersion spectra* 9
3.2.5 *Comparison with the NMC method* 9
3.3 Model error evaluation and decomposition 9
3.3.1 *Statistics computed from ARPEGE/ALADIN differences* 9
3.3.2 *Weight of the model error contribution* 10
3.3.3 *Decomposition of the ARPEGE/ALADIN differences variance* 12
3.3.4 *Comparison of ARPEGE/ALADIN forecast differences approach with the lagged NMC method* 13
3.4 Implications for the specification of the error statistics in the generalized formulation of the ALADIN 3D-Var 13
3.4.1 *Statistics obtained in the framework of a system simulation experiment* 13
3.4.2 *Implied changes in the variance spectra after the model error decomposition* 13
3.4.3 *Comparison with the NMC statistics* 14

4 Conclusions **14**

1 Introduction

To make accurate forecasts it is important to know the initial state of the atmosphere. The aim of data assimilation is to produce the best possible estimation of the initial state (called analysis), taking into account different sources of information like observations, a short range (usually 6h) forecast of the model and some balance properties of the atmosphere. Statistical features of the observations and forecast errors are needed in order to determine the weights of observations and 6h forecast (known also as first guess or background) in the analysis. The background error covariance matrix contains forecast error autocovariances, which relate errors of the same variable obtained at different locations, and cross-covariances, which relate errors of different variables, in accordance with the balance properties. This matrix is crucial to the performance of the analysis system, as it is used to filter the noise in the observations and to interpolate spatially the filtered information (Hollingsworth 1987).

Several approaches for producing the forecast error covariances to be used in the frame of a 3D-var data assimilation system in a limited area model are known. One approach consists of computing the error statistics from differences between model forecasts and observations. Hollingsworth and Lonnberg (1986) and Lonnberg and Hollingsworth (1986) used observations from high-density radiosonde station networks in order to estimate the forecast error covariances. The main shortcoming of this method is that it can not work properly for inhomogeneous observation networks. Moreover, simple covariance models are used, assuming the separability of the horizontal and vertical structures.

Another approach, known as the NMC method, was developed by Parish and Derber (1992) at the National Center for Environmental Prediction (NCEP). Through this method, the forecast error statistics is computed from a set of differences between different ranges of the model forecasts valid for the same time, but starting from different analysis times. Further, this approach was also used in other NWP centers, such as European Centre for Medium Range Weather Forecasts (ECMWF; Rabier et al. 1998) and Meteo-France (Desroziers et al. 1995). The NMC method has the advantage that provides forecast error statistics that are easy to implement in the operational data assimilation schemes. The method also produces structure function that are consistent with the model dynamics and physics (Derber and Bouttier 1999, Berre 2000). The most stingent shortcoming of the NMC method is that it does not retrieve the proper correlation structures in data-sparse areas (Van Tuyl and Daley 1999, Fisher 1999). The observations enter indirectly into the computations through the analysis increments added to the initial state used for the short-term forecast. Thus, the model spinup can lead to forecast error misspecifications in data-sparse areas. Furthermore, the overestimation of the correlation lengthscales can occur because of the 24h lag in the forecasts, while the 6h forecast error statistics used in the 3D-var assimilation scheme requires sharper structures (Van Tuyl and Daley 1999, Fisher 1999). It is important to note that, for a limited area model, the standard formulation of the NMC method will produce forecast error statistics which includes also the contribution of the error comming from the coupling model. Because the goal of the limited area 3D-var analysis is to analyse the small scales of the atmosphere, whithout reanalysing the scales already analysed in the global model, a set of mesoscale selective forecast errors is required. In order to eliminate the contribution of the error comming from the coupling model, Siroka et al. (2001) proposed an alternative version of the NMC method (refered as lagged NMC method) to be used in the mesoscale 3D-var data assimilation. The lagged formulation uses differences between two different forecast ranges valid for the same time, which start from the same ARPEGE analysis and are coupled with the same lateral boundary forcing.

The most recently tried approach consists in computing the forecasts error statistics from a representative ensemble of perturbed assimilation systems. This method was implemented at Atmospheric Environment Service, Canada. Houtekamer et al. (1996) combined two complementary approaches into a system simulation experiment, namely an ensemble prediction based on a perfect-model approach, for which only the observations that enter in the assimilation cycle are randomly perturbed, and model sensitivity experiments ensemble prediction based on a perfect-observation approach. Thus, in such a system simulation experiment, all elements of the forecasting system are subject to error imperfection. Houtekamer et al. (1996) performed a number of parallel assimilation cycles using representative perturbations to the observations, the surface fields and the forecast model. The perturbed observations are used as input into the data assimilation scheme, giving a perturbed analysis. This analysis is integrated for 6h with a perturbed version of the model, using perturbed surface fields, to obtain a perturbed first guess for the next assimilation cycle. Finally, Houtekamer et al. (1996) validated the generated ensembles of first guess fields using two independent methods. First, they compared the spread in the ensemble of the perturbed first guess fields against the optimal interpolation statistics (OI) of the 6h forecast. The second method was to compare the difference between the first guess and the observations with the spread in the ensemble and the variance of the observational error. Both methods shown that, for the extra-tropics regions, the wind errors are underestimated in the lower atmosphere and the height errors are underestimated in the upper atmosphere. Houtekamer et al. (1996) concluded that the underestimation is a real problem of the system simulation experiment method and that some sources of error have not been properly represented.

The analysis ensemble approach was also implemented at the ECMWF. The method currently used at ECMWF to estimate the background error statistics is to run an ensemble of independent analysis experiments, each experiment using observations perturbed by adding random errors which are consistent with the assumed statistics of observation errors (Fisher, 1999). The model perturbations were included through the use of stochastic physics (Buiza et al., 1999). The background error statistics is calculated from differences between background fields for the same dates and times, but for different analysis experiments. Comparing to the NMC method, the ensemble method produces sharper vertical correlations at the lower wavenumbers. Also, the ensemble approach yields significantly smaller standard deviations for wind and slightly larger standard deviations for temperature in the troposphere, excepting the planetary boundary layer where the standard deviation for temperature is smaller (Fisher, 1999).

The same approach was tried by Belo Pereira (2002) in the global model of Meteo-France ARPEGE and the results were in accordance with that one obtained at ECMWF.

One of the advantages of the ensemble method is that it produces a good estimation of the forecast errors also in the data sparse areas. As a shortcoming, the feedback effect can be mentioned, namelly the fact that a noisy analysis system will produce an unbalanced statistics which will contribute further to a noisier analysis (M. Fisher, seminar 2003).

In the present study, the ensemble approach is used to sample the forecast error covariances to be used in a 3D variational data assimilation for a limited area model. Using perturbed observations as input to the assimilation scheme, an ensemble of perturbed analysis is generated. The differences between two perturbed analysis, which are a measure of the uncertainties on the analysis, are propagated during the 6h forecast in the form of uncertainties in the first guess. Differences between perturbed analysis or first guess fields from the coupling model (i. e. the global model ARPEGE), valid for the same dates and times, but for different members of the ensemble, are used to produce the forecast error statistics for limited area model ALADIN. The computation of forecast error statistics using the 6h forecasts of ALADIN, starting from ARPEGE perturbed initial and boundary data, is also included in this study. A complete description of forecast error statistics should account also for the model error. The statistics computed from differences between the ALADIN 6h forecast and the ARPEGE analysis or first guess, for the same member of the ensemble, gives some information about the model errors. Finaly, the statistics computed from differences between ALADIN and ARPEGE fields, for different members of the ensemble, is supposed to give a more complete description of the forecast errors.

The paper is organized as follows. First, section 2 gives a short description of the forecast error statistics derived through the NMC method, as further we want to compare our ensemble results with the NMC derived statistics. Both standard and lagged versions of the NMC method are used in this study. In section 3, the forecast error statistics derived through the ensemble prediction method is presented. First, the perfect model and perfect observations approaches are investigated and then the results obtained combining this two approaches in a system simulation experiment are shown. A disscussion related to the model error evaluation and decomposition is also included in section 3. The comparison between the error statistics computed with the ensemble prediction and NMC approaches has been carried out too. In section 4, the conclusions and possibilities for further investigations are summarized.

2 Forecast error statistics derived through the NMC method

2.1 Description of the datasets

The NMC statistics is computed from differences between 36h and 12h forecasts of the limited area NWP model ALADIN. The ALADIN fields, valid at 12 UTC, were taken for a 48 days period (04.02 - 23.03.2002), the same as in the case of ensemble statistics computation. The error correlations are assumed to be homogeneous and isotropic. The adaptation of the NMC formalism to the limited area model ALADIN is presented in Berre (2000).

The standard formulation of the NMC method uses differences of forecasts valid for the same time, but starting from different analysis times. The 12h forecast uses new initial and lateral boundary conditions from ARPEGE data assimilation cycle and thus it benefits from new information. In this way, the error of the limited area model ALADIN containes the forecast error of the ARPEGE model, which enters through the initial and lateral boundary data. The use of such forecast error statistics will force the LAM 3D-var assimilation scheme to analyse also the scales that are already analysed in the global data assimilation cycle of ARPEGE model. Because the goal of the limited area 3D-var analysis is to analyse the small scales of the atmosphere, whithout reanalysing the scales already analysed in the global model, a set of mesoscale selective forecast errors is required. In order to eliminate the contribution of the error comming from the coupling model, Siroka et al. (2001) proposed an alternative version of the NMC method (refered as lagged NMC method) to be used in the mesoscale 3D-var data assimilation. The lagged formulation uses differences between 36h and 12h ALADIN forecasts valid for the same time, which start from the same ARPEGE analysis and are coupled with the same lateral boundary forcing.

In our experiments, the 36h ALADIN forecast is performed using the ARPEGE analysis as initial conditions and ARPEGE analysis plus ARPEGE forecasts with ranges from 3h to 36h as lateral boundary conditions. For the standard NMC computations, the 12h forecast is performed in the same way, using the ARPEGE analysis of the next day and ARPEGE forecasts with ranges from 3h to 12h. The differences in the lagged method come from the 12h forecast of ALADIN, which is performed now using the 24h forecast of the ARPEGE model (starting from the same analysis used as initial state for ALADIN 36h forecast) as initial conditions and ARPEGE forecasts with ranges from 24h to 36h as lateral boundary conditions. The lateral boundary coupling frequency is the same for all experiments and corresponds to the operational frequency of 3h.

2.2 Results obtained with standard and lagged NMC methods

Comparing to the standard NMC variance spectra, the lagged NMC variance spectra are smaller for all variables and all levels, especially in the large scales (see for example the variance spectra for divergence presented in figure 1).

The vertical profiles of standard deviation for divergence, vorticity, temperature and specific humidity are presented in figure 2. It can be observed that the standard NMC method produces larger values for all variables comparing to the lagged NMC method.

In order to evaluate the effect of the 12h integration of the ALADIN model on the lagged NMC statistics, the statistics computed from differences between initialized ARPEGE 24h forecast (interpolated into the ALADIN grid) and ALADIN 24h forecast has been investigated. In fact, this is what we have in the lagged NMC method, before the 12h integration of the ALADIN model. Figure 3 shows the standard deviation profiles for this case, in comparison with the results obtained in case of lagged NMC method. It can be seen that the 12h forecast of ALADIN is reducing the standard deviation, with an exception for vorticity in the mid-troposphere.

The length scales are smaller for the lagged NMC case, in accordance with what the variance spectra shown (see figure 4). The differences between standard and lagged values are larger for temperature, specific humidity and surface pressure than for vorticity and divergence. The explanation is that this difference comes from large scales and then it influence more temperature, specific humidity and surface pressure, as these variables are more representative for the large scales than divergence and vorticity.

3 Estimation of the error statistics using the ensemble prediction approach

3.1 Description of the datasets

In this study we are interested to compute forecast error statistics for a 3D variational data assimilation in the limited area NWP model ALADIN (more references about the model can be found in Bubnova et al. 1993 and Radnoti et al. 1995). The cycle CY25t1 of the ALADIN model was used for all experiments. Table 1 gives a short description of the main geometrical characteristics of the ALADIN France integration domain.

Table 1

Parameter	$J = K$	$M = N$	$L_x = L_y$	$\delta x = \delta y$	$Nflev$
Value	300	149	2852.686	9.508	41

The ensemble statistics was computed over a 48 days period, from 04.02 until 23.03.2002, using fields valid at 12 UTC. The ARPEGE perturbed analysis and first guess fields were taken from independent assimilation experiments performed by Belo Pereira and Berre (2002). For each experiment (called also member of the ensemble), all observations were perturbed by adding random numbers, which have a gaussian distribution with mean zero and variance equal to the assumed variance of observation errors. These perturbations in observations generate differences in the analyses from different experiments. The differences in the analyses are a measure of the uncertainties in the initial conditions, that are further propagated during the 6h integration of the model in form of uncertainties in the backgrounds. A period of 3 days of assimilation (skipped in the present study) was needed to echilibrate the statistics derived from differences between background fields for pairs of members of the ensemble (Belo Pereira, 2002). The differences between fields from pairs of members P31-P32 and P36-P37 were used in the present study. It is worthwhile to note that the two statistics computed from differences between the ensemble members P31-P32 and P36-P37 shown similar results and this fact proves that the statistics is stable. For this reason, in the next sections we present only results obtained by using the ensemble members P31-P32.

In a first stage, the ARPEGE analysis and first guess coupling fields were horizontaly interpolated into the ALADIN grid. The error statistics was computed from both differences between pairs of ARPEGE analysis and first guess fields, for two different members of the ensemble. The effect of ARPEGE analysis has been investigated in this way.

After performing the horizontal interpolation of the ARPEGE fields into the ALADIN grid, some artificial structures have been observed in the small scales, knowing that there is no information in the ARPEGE fields at scales bellow approximativelly 19 Km (the spatial resolution of ARPEGE model over France integration domain). To remove this noise, a digital filter initialization (DFI) was applied to the ARPEGE fields. The statistics computed from differences between initialized ARPEGE fields was also investigated in this study.

The computation of forecast error statistics from 6h forecasts of the ALADIN model, starting from ARPEGE perturbed initial and boundary data, has been carried out too.

In order to describe the model errors, the statistics obtained from differences between ALADIN 6h forecasts and ARPEGE analysis or first guess fields, for the same member of the ensemble, was investigated. Finally, the statistics computed from differences between ALADIN and ARPEGE fields, but for different members of the ensemble, is supposed to describe both initial conditions and model errors.

For all experiments, the variance spectra for 5 variables (divergence, vorticity, temperature, specific humidity and surface pressure) at 4 levels (13, 22, 29 and 41) have been plotted. Also, the vertical profiles of standard deviation and correlation length scale for the same variables have been investigated.

3.2 Contributions to the evolution of dispersion spectra in a perfect-model framework

Ensemble experiments based on the perfect-model assumption are presented in this section. The statistics are computed from fields provided by the same model, but from different members of the ensemble. The errors due to the uncertainties in the initial conditions are taken into account by perturbing the observations used in the assimilation cycles of different members of the ensemble.

3.2.1 *Effect of ARPEGE analysis*

One can observe a reduction of the variances by the ARPEGE analysis comparing to the ARPEGE first guess, especially in the large scales (see figure 5, which presents the variance spectra for vorticity). The reduction is about 30 percents for the wavenumber 1. The ARPEGE analysis does not reduce so much the standard deviation for temperature in the planetary boundary layer (see figure 36). For specific humidity, the effect of ARPEGE analysis produces a very small reduction of variance in the high troposphere and an increase of variance in the low troposphere, especially in the large scales (figure 6).

The effect of ARPEGE analysis produces large reduction of surface pressure variance in the large scales (about 50 percents for wavenumber 1), but it can be observed an increase of variance in the very small scales (corresponding to horizontal scales of about 8 km), comparing to the ARPEGE first guess variance (see figure 7).

Figure 8 shows a small reduction of the length scale for divergence and vorticity in the planetary boundary layer due to the effect of ARPEGE analysis. For temperature and specific humidity, the ARPEGE analysis shows a reduction of the length scale in the high troposphere, but an increase in the lower troposphere. The effect of the ARPEGE analysis is to reduce the length scale of surface pressure.

All figures previouly presented shown the values of variance averaged over the 48 days period. In order to see the evolution of the ARPEGE analysis effect for each day of this period, we computed the variance for each date (for each day of the period, at times 0, 6, 12 and 18 UTC), averaged over all wavenumbers and all model levels. The results obtained for averaged variance were confirmed. Figure 9 shows a reduction of divergence variance by the ARPEGE analysis comparing to the ARPEGE first guess, for almost all dates. There are some exceptions for few dates, where the effect of ARPEGE analysis is to increase the error variance. This is valid also for vorticity and surface pressure variances. The reduction of the error variance by the ARPEGE analysis is evidently also looking to the error variance accumulated in time plotted in figure 10.

As we saw different behaviours for temperature and specific humidity averaged standard deviations for planetary boundary layer and high atmosphere (figure 36), we calculated the variances for each date averaged over the model levels 1-30 (the high atmosphere part) and the model levels 31-41 (the boundary layer part). From figure 11 it can be observed that the reduction of the temperature error variance by the ARPEGE analysis is not so obvious in the planetary boundary layer (for some dates there is even an increse of variance by the ARPEGE analysis), as it is for the other levels of the model. For specific humidity, in general, the analysis increases the error variance in the planetary boundary layer (see figure 12). This is consistent with the unexpected results we have obtained for standard deviation profile averaged over the 48 days period.

3.2.2 *Effect of DFI applied to the ARPEGE fields*

After applying DFI, we can observe a reduction of the error variance for ARPEGE analysis and first guess in the mid troposphere, especially when going to the small scales (see for example figure 13, which shows the vorticity error variance spectra). This can be explained by the fact that DFI removes some unbalanced components of the error variance, namely the structures artificialy created by the horizontal interpolation in the small scales. This reduction can be observed for wavenumbers smaller then 40 (corresponding to a horizontal distance of about 18-19 Km), in accordance with the fact that the ARPEGE horizontal resolution over France integration domain is about 19 Km and for this reason we can not have any information below these scales. Such unrealistic structures created by the horizontal interpolation have been pointed out also in Wafaa (2000) for the Morocco ALADIN integration domain.

It was observed that DFI add some energy in the error variance spectra between wavenumbers 20 and 50, but not for all levels and all variables. Near the surface, there is also an addition of small scales variance, possibly thanks to the dynamical adaptation of the ALADIN model. The surface pressure spectrum shown no addition of variace in the small scales (not shown).

There is a large reduction of the length scale for divergence and vorticity due to the DFI in the planetary boundary layer. For all variables, DFI increase the length scale above planetary boundary layer (see figure 14). This fact is consistent with the reduction of the variance energy in the small scales when DFI is removing the spurious small scales features created during the horizontal interpolation of the ARPEGE fields into the ALADIN grid.

3.2.3 *Effect of the ALADIN 6h forecast*

Comparing the statistics computed from differences between ALADIN 6h forecasts with the statistics obtained from differences between ARPEGE fields, we can note a large increase of standard deviation for divergence and vorticity in the lower and mid troposphere (see figure 15).

It was observed a reduction of the error variance computed from ALADIN 6h forecast differences in the large scales, especially when comparing with ARPEGE first guess variance. This can be seen for divergence variance in figure 16. When comparing with DF initialized ARPEGE first guess variance, the reduction is very small or there is an addition of energy. We can explaine this by the fact that some gravity waves are probably removed by DFI, or some large scales flows become imposible after the dynamical adaptation, knowing what is happening in the small scales.

The effect of ALADIN 6h forecast is to increase the error variance in the small scales. The ALADIN 6h and ARPEGE first guess error variance curves start to depart significantly at wavenumbers greather than 13-17, i.e. corresponding to length scales smaller than 220-170

Km (see figure 16). The means is that the limited area model build up its own structures below these wavenumbers (in Wafaa 2000 critical wavenumbers are 7-10, corresponding to length scales of 400-300 Km; we have higher wavenumbers probably due to the higher resolution of ARPEGE model over France comparing to the Morocco integration domain).

The statistics computed from ALADIN 6h forecast differences show sharper correlations than the statistics obtained from differences between ARPEGE first guess fields, for all variables and all levels, except the cases of ARPEGE first guess for divergence in the mid and high troposphere and for vorticity in the high troposphere (see figure 17).

3.2.4 *Comparison between the four basic dispersion spectra*

To summarize all results previously presented, the four basic dispersion spectra (ARPEGE analysis, DF initialized ARPEGE analysis, ARPEGE first guess and ALADIN 6h forecast) were plotted on the same figure (see for example figure 18, which shows the vorticity variance spectra). The reduction of the large scales errors by the ARPEGE analysis and the removal of spurious small scales features are obvious, but the lesser ARPEGE errors in the small scales, comparing to the ALADIN 6h errors looks like a paradox, as we will expect that the higher resolution model (ALADIN) should give a better solution in the small scales.

3.2.5 *Comparison with the NMC method*

The ensemble method tends to provide results that are intermediate between standard and lagged NMC. In the large scales (wavenumbers greather than 10), the variances computed with the ensemble method are smaller than the standard NMC variances, and greather than the lagged NMC variances (see for example figure 19, which presents the variance spectra of divergence), with an exception for surface pressure variance (figure 20), for which the ensemble method using differences between different members of ARPEGE analysis produces smaller variances than the NMC lagged method in the very large scales (wavenumbers greather than 2-3).

The ALADIN 6h forecast ensemble statistics shows less large scales variance than the standard NMC statistics (see for example figure 21, which presents the vorticity variance spectra). This fact can be explained by the shorter forecast ranges used in case of ensemble method and by the representation of the ARPEGE analysis effect. The representation of the initial conditions and lateral boundary conditions uncertainties in the ALADIN 6h forecast ensemble statistics explaine the larger variance obtained through the ensemble method, in the large scales, comparing to the lagged NMC method. From the variance spectra for specific humidity presented in figure 22, it can be seen that in the large scales the ALADIN 6H forecast ensemble statistics produces variace spectra closer to the standard NMC variance spectra than to the lagged NMC variance spectra.

For the ensemble statistics computed from differences between the 6h forecasts of ALADIN, the standard deviation curve is located between standard NMC and lagged NMC standard deviation curves, except for divergence in the high troposphere and for vorticity in the mid and high troposphere, for which the ensemble standard deviations have the smallest values (see figure 23).

3.3 Model error evaluation and decomposition

3.3.1 *Statistics computed from ARPEGE/ALADIN differences*

A complete description of forecast error statistics should account also for the model error. The statistics computed from differences between the ALADIN 6h forecast and the ARPEGE analysis or first guess, for the same member of the ensemble, gives us informations about the model errors.

The standard deviation and length scale computed from differences between ALADIN 6h forecast and ARPEGE analysis are larger then the corresponding quantities obtained from differences between ALADIN 6h forecast and ARPEGE first guess fields, especially for temperature and specific humidity, but no so much for divergence and vorticity (see figure 24). This difference comes from the large scales (wavenumbers larger than 10-20). The explanation is that the ARPEGE analysis improves the temperature and specific humidity, which are more large scale fields, but not so much the divergence and vorticity, which are more important in the mezoscales. Moreover, from the error variance spectra it can be seen that temperature and specific humidity variances have maximum in the large scales, while for divergence and vorticity the maximum is located in the small scales (not shown). The length scale obtained from differences between ALADIN 6h forecast and ARPEGE analysis is larger than that one computed from differences between ALADIN 6h forecast and ARPEGE first guess (not shown) and this fact is consistent with the variance spectra, which show more energy in the large scales for the ALADIN 6h minus ARPEGE analysis curve (see for example figure 25, which presents the variance spectra for temperature).

It is important to note that the difference in error variance obtained from ALADIN 6h forecast minus ARPEGE analysis and from ALADIN 6h forecast minus ARPEGE first guess can result not only from the ARPEGE analysis increments, but also from the uncertaneities in the ARPEGE analysis due to the last 6h perturbed observations, knowing that if we want to have analysis and first guess valid at the same time, the analysis field will include also the perturbed observations from the last 6 hours of assimilation cycle.

The standard deviation computed from differences between ALADIN 6h forecast and DF initialized ARPEGE fields is smaller than the standard deviation obtained from differences between ALADIN 6h forecast and noninitialized ARPEGE fields, in accordance with the fact that DFI is removing some energy of the variance spectra (i.e. some unrealistic structures artificialy created during the horizontal interpolation procedure) in the ARPEGE coupling fields, in the small scales.

3.3.2 *Weight of the model error contribution*

Suppose we run two ensemble system simulation experiments, denoted by subscripts i and j in the following expresions. The perturbation of observations is accounted by using different members of the ensemble and the model errors are included by using two different models, ARPEGE and ALADIN. The difference between fields from these two models and from different members of the ensemble can be written as follows:

$$\delta x = x_i^{arp} - x_j^{ald} = (x_i^{arp} - x^t) - (x_j^{ald} - x^t) = \epsilon_i^{arp} - \epsilon_j^{ald} =$$

$$(\epsilon_i^{o,arp} + \epsilon_i^{m,arp}) - (\epsilon_j^{o,ald} + \epsilon_j^{m,ald}) = (\epsilon_i^{o,arp} - \epsilon_j^{o,ald}) + (\epsilon_i^{m,arp} - \epsilon_j^{m,ald}) \tag{1}$$

Here, x_i^{arp} and x_j^{ald} denote the ARPEGE and respectivelly the ALADIN fields, while x^t represents the true state of the atmosphere. ϵ_i^{arp} and ϵ_j^{ald} are the errors of the ARPEGE and ALADIN fields, which can be decomposed in errors in the initial conditions due to the perturbation of observations ($\epsilon_i^{o,arp}$ and $\epsilon_j^{o,ald}$) and errors due to the fact that the models are not perfect ($\epsilon_i^{m,arp}$ and $\epsilon_j^{m,ald}$).

The variance calculated from these differences reads:

$$\overline{\delta x \delta x^T} = \overline{(\epsilon_i^{o,arp} - \epsilon_j^{o,ald})(\epsilon_i^{o,arp} - \epsilon_j^{o,ald})^T} + \underbrace{\overline{(\epsilon_i^{o,arp} - \epsilon_j^{o,ald})(\epsilon_i^{m,arp} - \epsilon_j^{m,ald})^T}}_{=0} +$$

$$\underbrace{\overline{(\epsilon_i^{m,arp} - \epsilon_j^{m,ald})(\epsilon_i^{o,arp} - \epsilon_j^{o,ald})^T}}_{=0} + \overline{(\epsilon_i^{m,arp} - \epsilon_j^{m,ald})(\epsilon_i^{m,arp} - \epsilon_j^{m,ald})^T} \tag{2}$$

It was assumed that there is no correlation between the errors in the initial conditions due to the use of perturbed observations and the model errors and therefore terms 2 and 3 from the right hand side of the previous equation are set to zero. Further, the expression of variance can be written as:

$$\overline{\delta x \delta x^T} = \overline{\epsilon_i^{o,arp} \epsilon_i^{o,arpT}} + \overline{\epsilon_j^{o,ald} \epsilon_j^{o,aldT}} - (\overline{\epsilon_i^{o,arp} \epsilon_j^{o,aldT}} + \overline{\epsilon_j^{o,ald} \epsilon_i^{o,arpT}}) +$$
$$\overline{\epsilon_i^{m,arp} \epsilon_i^{m,arpT}} + \overline{\epsilon_j^{m,ald} \epsilon_j^{m,aldT}} - (\overline{\epsilon_i^{m,arp} \epsilon_j^{m,aldT}} + \overline{\epsilon_j^{m,ald} \epsilon_i^{m,arpT}}) \tag{3}$$

Writing the equation (3) for the particular case of perfect-model assumption and taking into account the facts that members i and j are independent and that the observations were perturbed by adding random numbers which have the variance equal to the prescribed variance of the observation error, than we obtain the same expresion of variance as is presented in Fisher 1999:

$$\overline{\delta x \delta x^T} = 2\overline{\epsilon^o \epsilon^{oT}} \tag{4}$$

The means of equation (4) is that the differences between fields will have the same correlations as the errors and twice the variance.

On the other hand, assuming perfect observations by computing differences between fields from different models, but for the same member of the ensemble (i=j in our case and thus we will skip this index in the following derivations), equation (3) becomes:

$$\overline{\delta x \delta x^T} = \overline{\epsilon^{m,arp} \epsilon^{m,arpT}} + \overline{\epsilon^{m,ald} \epsilon^{m,aldT}} - (\overline{\epsilon^{m,arp} \epsilon^{m,aldT}} + \overline{\epsilon^{m,ald} \epsilon^{m,arpT}}) \tag{5}$$

It is worthwhile to note that, in the previous equation terms 1 and 2 are in fact the square of standard deviation model error for ARPEGE, respectivelly ALADIN, while the terms included in the parantheses represent the correlations between ARPEGE and ALADIN model errors. Despite the fact that ALADIN and ARPEGE physical and dynamical setups can differ due to the different scales represented in the two models, the correlations between the two model errors can not be neglected, as ARPEGE and ALADIN models are based on the same set of equations. But, if we think in terms of ARPEGE model errors comparatively to the ALADIN errors (assuming that the ALADIN model gives a perfect solution for mesoscales), we can write:

$$\delta x = x^{arp} - x^{ald} = (x^{arp} - x^t) - (x^{ald} - x^t) = \epsilon^{arp} - \epsilon^{ald} = \epsilon^{m,arp} - \epsilon^{m,ald} = \epsilon^{m,arp/ald} \tag{6}$$

where with $\epsilon^{m,arp/ald}$ we denote the relative model error of ARPEGE comparing to ALADIN model error.

In this case, the variance computed from differences between ALADIN and ARPEGE fields becomes:

$$\overline{\delta x \delta x^T} = \overline{\epsilon^{m,arp/ald}\epsilon^{m,arp/aldT}} \tag{7}$$

But, because in the variance computed from differences between ALADIN and ARPEGE fields for different members of the ensemble, we have twice the variance of the errors generated by the uncertainties in the initial state due to the observations perturbation, we would like to represent also twice the variance of the model error. Adding twice the model error to the errors due to the use of perturbed observations, we expect to increase the total error variance in the small scales, which is resonable if we take into account the fact that the ARPEGE model error relative to the ALADIN model should be located mainly in the small scales.

Indeed, it was observed that, in the case of twice the model error, the variance for divergence and vorticity increases more in the small scales (see for example figure 26 for divergence variance spectra). This is not so much obvious for temperature, specific humidity and surface pressure (see figure 27 for specific humidity variance spectra), which are more large scale variables and probably the relative error of ARPEGE comparing to ALADIN model is not so large in the small scales.

3.3.3 *Decomposition of the ARPEGE/ALADIN differences variance*

From figures 26 and 27, it can be observed also an increase of the variance in the large scales (especially in the low levels), in the case of using twice the model error variance. As we can not relate this large scales increase of variance with the ARPEGE dynamics errors relative to the ALADIN model, it can be explained as some large scales errors due to the coupling formulation inaccuracies in the ALADIN model. Therefore, it would be usefull to separate this two contributions in the model errors computed from differences between ALADIN and ARPEGE fields. The difference between the variance spectra of ALADIN 6h forecast and initialized ARPEGE first guess (computed in the ensemble prediction perfect-model framework) represents the part of the model error related to the ARPEGE dynamics errors. The initialized ARPEGE first guess was chosen instead of not initialized field because in this case we do not have the spurious small scales features created by the horizontal interpolation of the ARPEGE guess into the ALADIN grid. Figure 28 presents the ALADIN 6h variance compared to the ARPEGE initialized first guess variance for divergence. The two curves presented in this figure start to depart significantly for wavenumbers smaller than 10, so that the difference between the two curves can be interpreted as the small scales part of the model errors due to the fact that the ARPEGE dynamics is not so well represented at this scales. Using these two variances, we can define a parameter α to be used in the separation of ARPEGE dynamics errors from the coupling deficiency errors:

$$\alpha = (variance(ald06) - variance(dfiarpguess))/variance(ald06) \tag{8}$$

Further, the model error variance ($\overline{\epsilon^{m,arp/ald}\epsilon^{m,arp/aldT}}$ previously described) can be decomposed as follows:

$$variance(\epsilon^{m,arp/ald}) = \alpha * variance(\epsilon^{ss}) + (1 - \alpha) * variance(\epsilon^{ls}) \tag{9}$$

where $variance(\epsilon^{ss})$ denotes the part of the error variance due to the small scales ARPEGE dynamics deficiencies, while $variance(\epsilon^{ls})$ denotes the part of the error variance due to the large scales ALADIN coupling innacuracies.

The spectral variation of the parameter α shown values close to 0 in the large scales and close to 1 in the small scales, as it is required for the decomposition of the model error variance (see for example figure 29 which present this parameter for divergence).

3.3.4 *Comparison of ARPEGE/ALADIN forecast differences approach with the lagged NMC method*

Figure 30 presents the variance spectra of vorticity for differences between the 24h forecast of ALADIN and initialized 24h forecast of ARPEGE , which is exactly what is represented in the lagged NMC method before the 12h integration of ALADIN (the solid curve), and for differences between the 6h forecast of ALADIN and initialized ARPEGE first guess, from the same member of the ensemble (the long dashed curve). We can see that the difference between these two variance curves is similar to the difference between the lagged and standard NMC variance curves. It is obvious that the lagged NMC approach is close to an approach based on ARPEGE/ALADIN forecast differences.

3.4 Implications for the specification of the error statistics in the generalized formulation of the ALADIN 3D-Var

3.4.1 *Statistics obtained in the framework of a system simulation experiment*

The statistics computed from differences between different models (ARPEGE and ALADIN) and for different members of the ensemble is presented in this section.

For divergence and vorticity, the standard deviation and length scale computed from differences between ALADIN 6h forecast and ARPEGE analysis are very close to that ones obtained from differences between ALADIN 6h forecast and ARPEGE first guess (see figure 31 which presents the standard deviation vertical profiles). This fact is consistent with the variance spectra, which show the maximum of error variance for divergence and vorticity in the small scales and for that these fields are not affected so much by the ARPEGE analysis. For temperature, specific humidity and surface pressure, the standard deviation and length scales computed from differences between ALADIN 6h and ARPEGE analysis are larger than that ones obtained from differences between ALADIN 6h forecast and ARPEGE first guess. The difference comes from the large scales (wavenumbers smaller than 4-5) (see for example figure 32 for specific humidity variance spectra).

We have to note here that the difference in error variance obtained from ALADIN 6h forecast minus ARPEGE analysis and from ALADIN 6h forecast minus ARPEGE first guess is a result of the ARPEGE analysis increments combined with the use of perturbed observations for two different members of the ensemble. Moreover, there is an effect of the uncertaneities in the ARPEGE analysis due to the last 6h perturbed observations.

3.4.2 *Implied changes in the variance spectra after the model error decomposition*

By adding the large scales error variance due to the ALADIN coupling formulation innacuracies to the ALADIN 6h error variance due to the initial conditions uncertainties (the long-dashed curve presented in figure 38 for divergence variance sectra), we expect to describe better the total ALADIN forecast errors (the dotted curve in figure 38).

On the other hand, by adding the small scales error variance due to the ARPEGE dynamics deficiencies to the ARPEGE initialized analysis error variance due to the initial conditions uncertainties (the solid curve in figure 37), we would expect to describe better the total ARPEGE analysis errors (the dot-dashed curve in figure 37).

It can be observed that, finally, the ARPEGE analysis and ALADIN 6h total variances look very well, as the ALADIN 6h variance is larger than the ARPEGE analysis variance in the large scales and the ARPEGE analysis variance is larger than the ALADIN 6h variance in the small scales (see figure 34 for divergence variance spectra). From figure 33 it can be seen that the unexpected results obtained for ALADIN 6h forecast and ARPEGE analysis variances, computed in the framework of the perfect-model assumption (i. e. the fact that the ALADIN 6h variance is larger than the ARPEGE analysis variance in the small scales), are changed now into realistic behaviours for the two considered variances.

3.4.3 *Comparison with the NMC statistics*

The comparison with the NMC method shown that the ensemble method tends to provide results that are intermediate between standard and lagged NMC.

The total variance for ALADIN 6h forecast (the sum of the variance due to the initial conditions uncertainties with the variance due to the large scales ALADIN coupling deficiencies) has values between the standard and lagged NMC variances, with an exception for specific humidity, for which is closer to the standard NMC variance (see for example the figure 35 for divergence variance spectra, in which the dotted curve represent the total ALADIN 6h total variance). The lesser large scales variance comparing to the NMC standard variance can be explained by the use of the shorter forecast ranges and by the representation of the ARPEGE analysis effect, while the larger variance comparing to the lagged NMC variance can be explained by the representation of the initial and lateral boundary conditions uncertainties, and also by including the coupling formulation innacuracies.

The total variance for ARPEGE analysis (the dot-dashed curve in figure 35) has also intermediate values between standard and lagged NMC variances, with some exceptions at low levels, where the influence of the ARPEGE dynamics errors leed to an increase of the ARPEGE analysis variance in the small scales, and then the total ARPEGE analysis variance becomes larger than the standard NMC variance.

4 Conclusions

The ARPEGE analysis effect is to reduce the variance comparing to the ARPEGE first guess, especially in the large scales. An exception was found for specific humidity, namely the increase of the variance in the low troposphere by the ARPEGE analysis.

The DFI applied to the ARPEGE fields removes some unbalanced components of the error variance, namely the structures artificially created during the horizontal interpolation of the ARPEGE fields into the ALADIN grid.

It was observed a reduction of the error variance computed from ALADIN 6h forecast differences in the large scales, especially when comparing to the ARPEGE first guess variance. As a paradox, the effect of ALADIN 6h forecast is to increase the variance in the small scales, in case of a perfect-model framework. Later on, by adding the model error variance to the error variance due to the uncertainties in the initial conditions, we found good behaviours for ALADIN 6h and ARPEGE analysis variance spectra.

Adding twice the model error to the error due to the use of perturbed observations, the total error variance increases in the small scales, which is resonable if we take into account the fact that the ARPEGE model error comparatively to the ALADIN model should be located mainly in the small scales.

It was observed also an increase of the variance in the large scales (especially in the low levels), in the case of using twice the model error variance. As we can not relate this large scales increase of variance with the ARPEGE dynamics errors relative to the ALADIN model, it can be explained as some large scales errors due to the coupling formulation inaccuracies in the ALADIN model. A parameter α was defined in order to decompose the model errors into a part related to the ARPEGE small scales dynamics errors and another part related to the ALADIN large scales coupling deficiencies errors.

Afterwards, we added the large scales error variance due to the ALADIN coupling formulation innacuracies to the ALADIN 6h error variance due to the initial conditions uncertainties, in order to describe better the total ALADIN forecast errors. By adding the small scales error variance due to the ARPEGE dynamics deficiencies to the ARPEGE initialized analysis error variance due to the initial conditions uncertainties, we expect to describe better the total ARPEGE analysis errors. Finally, the ARPEGE analysis and ALADIN 6h total variances look resonable, as the ALADIN 6h variance is larger than the ARPEGE analysis variance in the large scales and the ARPEGE analysis variance is larger than the ALADIN 6h variance in the small scales.

The comparison with the NMC method shown that the ensemble approach tends to provide results that are intermediate between standard and lagged NMC. The ALADIN 6h forecast ensemble statistics shows less large scales variance than the standard NMC statistics. This fact can be explained by the shorter forecast ranges used in case of ensemble method and by the representation of the ARPEGE analysis effect. The representation of the initial conditions and lateral boundary conditions uncertainties in the ALADIN 6h forecast ensemble statistics explaine the larger variance obtained through the ensemble method, in the large scales, comparing to the lagged NMC method.

At the end, we can conclude that the ensemble approach produces reliable good estimation of the forecast errors.

Furter investigations can include experiments with an ensemble of ALADIN 3D-var perturbed analysis. Also, the model error representation can be improved by identifying more sources of the errors.

Acknowledgments.

The present study has been carried out during a stay in CNRM/GMAP, possible thanks to Meteo-France fundings. The author would like to express her thanks for many advises and interesting disscusions to Loik Berre, who has been her scientific supervizor. The author is grateful to Margarida Belo Pereira, who has tried first the ensemble approach in the global model ARPEGE, making possible the use of ARPEGE perturbed fields in the present study.

REFERENCES

Belo Pereira M., 2002: Improving the assimilation of water in a NWP model. Report for ALATNET Newsletter 4.

Belo Pereira M., 2002: Improving the assimilation of water in a NWP model. Report for ALATNET Newsletter 5.

Berre L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. Mon. Wea. Rev., 128, 644-667.

Bubnova R., Horanyi A., Malardel S., 1993: International project ARPEGE-ALADIN. EWGLAM Newsletter, No. 22, Institut Royal Meteorologique de Belgique, 117-130.

Buiza R., Miller M., Palmer T. N., 1999: Stochastic representation of model uncertainties in the ECMWF EPS, O.J.R. Meteorol. Soc., 125, 2887-2908.

Derber J., Bouttier F., 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus 51A, 195-221.

Desroziers G., Mathiot V., Orain F., 1995: A study of ARPEGE forecast error covariances. Proc. WMO Second Int. Symp. on Assimilation of Observations in Meteorology and Oceanography, Vol. 1, Tokyo, Japan, WMO, 263-268.

Fisher M., 1999: Background error statistics derived from an ensemble of Analyses. ECMWF Research Department Technical Memorandum No. 79, September 1999, 12pp.

Fisher M., 2003: Background error covariance modelling. Presentation to the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8 to 12 September 2003.

Hollingsworth A., 1987: Objective analysis for numerical weather prediction. Short- and medium-range numerical weather prediction. Proc. WMO/IUGG NWP Symp., Tokyo, Japan, Meteorological Society of Japan, 11-19.

Hollingsworth A. and Lonnberg P., 1986: The statistical structure of short-range forecast errors determined from radiosonde data. Part I: The wind field. Tellus 38A, 111-136.

Lonnberg P. and Hollingsworth A., 1986: The statistical structure of short-range forecast errors determined from radiosonde data. Part II: The covariance of height and wind errors. Tellus, 38A, 137-161.

Houtekamer et al., 1996: A system simulation approach to ensemble prediction. Mon. Wea. Rev., 124. 1225-1242.

Parrish D. F., Derber J. C., 1992: The National Meteorological Center's spectral statistical interpolation analysis system. Mon. Wea. rev., 120, 1747-1763.

Rabier et al., 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. O.J.R. Meteorol. Soc., 124, 1809-1829.

Radnoti G. et al., 1995: The spectral limited area model ARPEGE-ALADIN. PWPR Rep. Series, No. 7, WMO TD No. 699, 111-118.

Sadiki W., Fisher C., Geleyn J-F., 2000: Mesoscale background error covariances: recent results obtained with the limited area model ALADIN over Morocco. Mon. Wea. Rev., 128, 3927-3935.

Siroka M. et al., 2001: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. Meteorol. Atmos. Phys., 82, 227-244.

Van Tuyl A. H., Daley R., 1999: Estimation of forecast error covariances over an oceanic region using NORPEX data. Preprints, 13th Conf. on Numerical Weather Prediction. Denver, CO, Amer. Meteor. Soc., 15-19.