Steluta ALEXANDRU: "Scientific strategy for the implementation of a 3D-VAR data assimilation scheme for a double-nested limited area model"

The purpose of the work is to find the best scientific strategy for the implementation of the 3D-VAR data assimilation scheme for a double-nested limited area model. ALADIN/HU is a double-nested limited area model, which is coupled with ALADIN/LACE (which is coupled with ARPEGE). The boundary conditions are obtained from ALADIN/LACE integration, at 12 km resolution. The resolution of the ALADIN/HU model is about 8 km.

To produce a good forecast, a good description of the initial conditions is necessary. The objective of data assimilation is to define the best initial state of the model taking into account all the possible information. 3D-VAR consists in minimizing of a cost function, in order to get the best fit to the available information sources. The cost function is the sum of three terms:

$$J(x) = J_b(x) + J_o(x) + J_c(x)$$

where $J_b(x)$ represents the departures from the first-guess, $J_o(x)$ represents the departures from the observations, and $J_c(x)$ controls the amplitude of gravity waves in the analysis. At the moment for ALADIN model only the first two terms are defined.

Most of the time of the first months was devoted to study of the existing informatic environment with special emphasis on the 3D-VAR scripts. In 2000, the 3D-VAR data assimilation scheme was ported to Budapest for the Origin 2000 machine of the Hungarian Meteorological Service. The AL13 version of ALADIN model is used for the 3D-VAR, taking into account classical NMC statistics for the background term and only SYNOP and TEMP observations for the observational term of the cost function. The first-guess is the 6h forecast from an earlier model run. It contains information on the small scales (the scales of the model), but being a forecast it is not fully precise. The 3D-VAR script is running four times per day (to have the first-guess), but the 48h integration is made twice per day.

The steps performed in the 3D-VAR script are: first, the transformation of the observational data into ascii file, using Mandalay, then the program OBSORT redistributes the observational data across the available processors, followed by the running of the SCREENING (nconf=002), for calculating the high resolution departures and for screening of observations (i.e. determines which observations to be passed for use in the analysis), then the incremental variational analysis (nconf=131) is performed, using the first-guess (6h forecast from the previous model run), the observation file prepared by the screening, and the classical background error statistics file. Because the 3D-VAR data assimilation scheme is acting for the upper-air meteorological fields, after the variational analysis step, the optimal interpolation (CANARI) analysis is applied for the surface variables, and finally the model integration is realized to obtain the first-guess.

The time-consistency coupling is used for the time being to provide the information about the large scales. Time-consistency means that the lateral boundary data are coming from the same run of the model (LBC0=00h analysis of the ALADIN/LACE, LBC1=6h forecast of the ALADIN/LACE). Thus the information is

consistent in time. On the other hand, space-consistency means that the zero coupling file is identical with the initial file (LBC0=INIT file - the ALADIN/HU analysis, LBC1=6h forecast of the ALADIN/LACE), and this coupling technique will be tested in the future.

Twice per day a verification script is running to evaluate the model performances. The statistical measures (RMSE, bias) are used as indicators of the extent at which model prediction match observations. These objective scores are calculated for the forecast of the model using 3D-VAR scheme, and also for the operational forecast (in dynamical adaptation). The compared fields are the temperature, geopotential, zonal and meridional wind, relative humidity, direction and intensity of the wind.

We observed that for example, in the period 17.11.2001 - 24.12.2001, at the 850 hPa level, the scores are better for the wind and geopotential, for the first 6h integration, using 3D-VAR scheme. For geopotential, after 6h integration, the scores start to be almost the same. The worst scores are for relative humidity, where the BIAS is around 5%, in the first 6h integration of the model with 3D-VAR scheme, and around 1.5% for the operational one, and the RMSE is 17% using 3D-VAR scheme, and 13.5% for the operational. After 24h integration, we observed that the scores start to be close one each other. Also the model using 3D-VAR scheme underestimated the temperature, in the first 18h integration, but less than the operational model. Then the scores are almost the same.

After the understanding of the working environment of the 3D-VAR data assimilation scheme, we decided to focus our attention on the next aspects:

- ◇ Case studies.

  For the beginning we chose two cases, from 11.06.2001 and 28.08.2001. First case is from the period that G. Bölöni made the tests with blendvar and varblend in Prague, and the idea was to see the results for a double-nested limited area model. The cases were chosen because we considered them as interesting meteorological situations, with a front passing through the domain. On these two cases we will make the experiments with different background error statistics (standard, lagged statistics), first-guesses (6h forecast of the previous run, blending of the ALADIN forecast with ARPEGE analysis), initialization methods, together with their posteriori evaluation, etc.

- ◇ We want to establish which one of the coupling techniques is better, time-consistency or space-consistency.

- ◇ Then we plan to run 4 times per day the blending procedure for the first-guess. The necessary files, which contains the large scales are provided from Prague. These are the first two historical files from the ARPEGE assimilation cycle, the final blended analysis in the ALADIN/LACE resolution, and the 6h forecast of the ALADIN/LACE model, started from the previous file, with no digital filter initialization.

- ◇ Other experiments will be performed using the blendvar and varblend combinations, with lagged statistics.

Future work

The next months will be devoted to the experiments with different background error statistics, first-guesses, initialization methods and posteriori evaluation. So at the end of this first stay, we would like to get an idea about the best possible version of 3D-VAR data assimilation scheme for a double-nested limited area model.