RC LACE internal report
July and August, 2003

Further blending experiments
Studying of LH-Statistics

Helga Tóth
toth.h@met.hu

Supervisors: Radmila Brožková
radmila.brozkova@chmi.cz
Gergely Bölöni
boloni.g@met.hu
Claude Fischer
claude.fischer@meteo.fr

1 Introduction

Two main scientific studies will be carried out in this report, both of them are connected to 3DVAR data assimilation system developing. It's proven from the previous treatments that using blending procedure mixed with 3DVAR data assimilation system (BLENDVAR or VARBLEND), the forecasts scores are improved (Bölöni, 2001). So far this blending procedure was DFI blending and now one the target of the work in Prague is to test the analysis and the forecast created by explicit blending system.

The second half of our researching activity would be connected to that imagination that in the BLENDVAR assimilation the blending increment is too large and the goal is to decrease this component by multiplying an alpha factor. This alpha factor stands for the standard deviation of the blended state, Aladin forecast and Arpege analysis. The idea is to obtain these standard deviations by computing Lönnberg-Holligswoth estimates from innovation vectors (errors of the analysies/guesses against temp-observations). Here the main job was to develop a program which is able to calculate these covariances.

2 Explicit and DFI blending methods

One month parallel test was performed to compare the scores between the two kind of methods. First of all assimilation cycle was run. The main difference between the two assimilation cycles was that in explicit blending there is no low resolution DFI and instead of spectral blending function (sumfield()) explicit blending script-function (sumfielc()) and binary was applied. So this means surface blending is performed in both of cases.

This blending binary (`blend.e`) is in sx6 supercomputer in `/home/mma/mma167/blending` directory. Two source files are `blend.f` and `ellips.f`. The compilation script is `comp.sx6`, the result binary is `blend.e`. For to make running namelist file is needed `fort.4`. This namelist contains the name of the input files (Aladin guess and Arpege analysis) and the output file (blended), the NSMAX and NMSMAX for Aladin and the two limit wave-numbers, from 0 till NI0 Arpege analysis, from NI0 till NI1 blended of the two files and from NI1 only Aladin guess is used. These settings are the follows: NI0=0, and NI1=31.

The DFI blending settings are the same as in the operational suits (RATIO=2.5, NSMAX_NMSMAX=28.31, TAUS=18000, NSTDFI=5). After assimilation event production was made for the mentioned month from both kind of blending analysies. Incremental DFI was also used to produce the forecast. The result of the verification and a case study will be demonstrated in the next sections.

2.1 Case study

A case study was chosen to test and compare the main differences of the blending methods. On 22th of July a big cyclone stayed over Western-Europe which moved to the east. This cyclone had a long-waved, north-south direction frontal zone which was passing Check Republic at that time. Behind this frontal-zone instable air-mass arrived to the contry producing severe thunderstorms and big precipitations all over the area of Check Republic (Figure 1).

For the case of DFI blending the analysis and the couplings were brought from the archiv mashine and to produce a 48 hour forecast by using them. In the explicit blending experiment 6 days assimilation cycle was computed and after that forecast was produced.

First of all let's have a look on the analysies fields. Fig. 2. and 3. shows the process of producing analysis on the example of 500 hPa vertical velocity. Aladin guess and Arpege analysis, pure DFI and Explicit blending fields, and these last fields after IDFI are plotted. It is obvious that both of blended analysis fields contain the small scales features which appear in the Aladin guess, but they can't be found in the Arpege analysis. It is interesting that in case of DFI blending IDFI step doesn't touch the pure DFI field, the two figures are almost the same, but in the case of explicit blending IDFI acts on the sructures indeed to make it closer to the guess. The explanation may be the follow: in IDFI step the pure spectral blending analysis is supplemented by the guess-increment, which stands for guess–DFIguess. This means that in case of explicit blending is guess-increment is larger than for DFI blending, namely the difference between the guess and DFIguess is larger for explicit blending. This is obvious if we examine the preparation of guess by the two kind of blending methods. In DFI blending a lot of DFI is done over against explicit blending. This unbalanced structure can cause a large difference between DFI and noDFI results.

The result of the precipitation forecast can be seen in the Fig. 4. and 5. If we concentrate only the area of Check Republic it can be seen well that the forecast done by DFI is not too succesful, too big amount of precipitation outside of the domain and smaller inside. While explicit blending predicted more rains over the country. However it is also true that the field is very scattered in the case of explicit blending which is also not too realistic.

The two kind of vertical velocity forecast at 500 hPa can be seen in Fig. 6. The deviation between the two fields taper off in time. Some other fields were also examined (relative humidity and temperature at some levels) but the difference of the two kind of forecast are too small.
The forecasts (ICMSH files) are archived on mashine archiv:
`/sam/sx4.lace/mma167/Prod/$CNEXP` directory, where
CNEXP=EXPL in case of explicit blending
CNEXP=DFI in case of DFI blending
The PF files are also archived on voodoo: `/utemp/users/mma167/$CNEXP`
Chagal figures can be found on voodoo in: `/home/mma167/chagal` directory.

2.2 Parallel test

The results of the one month parallel test can be seen in Fig. 7-10. The curves show that the two kind of forescasts have almost the same, neutral scores. Sightly difference can be seen on humidity scores (Fig. 7 and Fig. 10.) even at high levels (500, 700 hPa) the explicit blending produced smaller rmse. In Fig. 9 the evolution of the bias-scores of the forecast are plotted for the analysies fields. The visible difference in BIAS disappears in the case of RMSE which means, that explicit blending scores are jumping on time more than DFI blending.
Fig. 11. contains the scores of the Aladin guess fields for both case. The results are almost neutral, expect for MSLP at the begining of February. The veral files are in voodoo in `/utemp/users/mma167/veral_inp` (for DFI and EXPL forecasts).

3 Lönnberg-Hollingswoth statistics

Lönnberg-Hollingswoth statistics were calculated using several kind of backgrounds (analysies, guesses). In the next sections we will show how we calculeted the statistics, then the computer settings will be presented, and finally the results will be demonstrated.

3.1 Calculation of LH-statistics

Auto-covariances were computed by using of values of fg_depars (errors of the analysies/guess against the temp-observations) for a month (2003. February). And then we collected these series of pairs of covariances into classes of distance (horizontal bins). Each bin is represented by one value which is now the mean. These values are plotted as a function of bin distances. Since the observation errors are horizontally uncorraleted (for horizontal distances > 0) so its contribution to covariance is at horizontal distance $= 0$. On other distances the covariances are represented by the spatial errors only on the background (guess or analysis) terms. Consequently a curve can be fitted to the points which can be extrapolated to the origin (horizontal distance $= 0$). From the result of the extrapolation to 0 the background error can be seen which is not equal with the variances we calculated originally.

The code system which calculates the LH-statistics is on mashine `sx6`. The main directory is:
`/home/mma/mma167/stat`
The source files are:

```
/home/mma/mma167/stat/source/main.f90
                             read_ecmadump.f90
                             read_cmadump.f90
```

calc_mean.f90
calc_cov.f90
calc_next_date.f90
calc_numPeriod.f90
distance.f90

The compiler script is:
/home/mma/mma167/stat/scr/comp
The necesarry parameter file and namelist are:
/home/mma/mma167/stat/nam/stat2.par
fort.4
The binary is: /home/mma/mma167/stat/bin/main.e
The program is able to calculate the auto- or cross-covariances between temp-stations. The INPUT is two ASCII files: CMAFOC.dump which contains the observation values (this file has a special structure, good description about it can be found in Prague). And ECMA.dump file is the result of running mandaodb and this file has to contain the errors between the guess and the observation. This file also has to be in special format, otherwise the program is aborted. If the ASCII file contsins some other characters than numbers (for example stars), the program is aborted. This is also weakness of the program-system, which has to be solved in the future. The program is able to calculate auto- and cross covaraînces for temp data, but for surface variables it still doesn't work well for the time being.
ECMA.dump files are prepared in the the following way:
First of all CMAFOC.dump ASCII type file is taken from archiv, and they are converted to ODB format with *lamflag*, *to_ odb* and *shuffle* scripts. Then *screening* was performed with some kind of initial backgrounds (Arpege analysis after low resolution DFI, Arpege analysis after low resolution DFI and e927, Aladin guess after low resolution DFI, Aladin guess after low resolution DFI and e927, DFI explicit blending analysis).
ECMA ODB-s for February and July can be found on archiv in
/sam/sx4.lace/mma167/ECMA/$CNEXP directory. Where CNEXP means
ARPR (Arpege analysis after low resolution DFI)
ARPI (Arpege analysis after low resolution DFI and e927)
ALAR (Aladin guess after low resolution DFI)
ALAI (Aladin guess after low resolution DFI and e927)
ALAS (DFI blending analysis)
EXPA (Explicit blending analysis)
ECMA dump-s are also archived only for February on archiv:
/sam/sx4.lace/mma167/ECMA.dump with the same directory structure as above.
The running scripts are in sx6:
/home/mma/mma167/scr/dfi_assim.scr for running DFI blending assimilation cycle
/home/mma/mma167/scr/expl_bl_assim.scr for running Explicit blending assimilation cycle

`/home/mma/mma167/scr/screening.scr` for running screening
`/home/mma/mma167/scr/scr_cov` for calculating the LH-statistics

3.2 Results

Fig. 12-14. show the figures what I got with running of LH-statistic calculator program for February. Auto-covariances were computed for some variables at some levels. The results of the average covariances of the errors in fuction of horizontal bins. Individual covariancies can be plotted also as a function in absolute distances, and we can get series of points. From the displayed graphics one immediatelly can see that the auto-covariance at 0 (infact this is variance) does not fitting to that curve, which can be drawen by the auto-covariance values. This behaviour is naturally because the stations are independent in space, so except 0 distance only the spatial error of the background is appearing.

In generality the deviations between the different kind of covariances are neglectable excpet for relative humidity and geopotential where the differences are perceptible. The blended analysies (DFI and explicit) have the smallest covariances (this is expected) and Aladin guesses (ALAR and ALAI) produce the largest one which is very amazing. This can be caused by several thing, for example Aladin is worse for relative humidity and for geopotential or the LAM data have more datails. The answer is not well-known.

4 Conclusions, future studyings

Explicit and DFI blending assimilation system were compared by doing case study and one month parallel test. The scores and the results are nearly neutrals which is very optimistic for us to use explicit blending procedure combined with 3DVAR in place of the more complicated DFI blending. Of course more treatments are needed to be sure of it. For instance it would be nice to see some verification scores which are produced by without IDFI.

A program system was developed which is able to calculate the LH-statistics which could be an important tool for 3DVAR developing. In this topic a lot of work is still needed. The code-system is not fully, some more expansion is necessary, for example cross- and auto-covariances for surface variables. In the near future the contruction of the alpha function is also possible. It would be useful to compare our results from LH-statistics with those profiles from B-matix used in 3DVAR.

5 References

Široká, M., 2001: Incremantal digital filter experiments in ALADIN/LACE blending production suite. *RC LACE internal report*

Bölöni, G., 2001: Futher experiments with the combination of 3DVAR and blending by DFI: test using incremental digital filter. *RC LACE internal report*

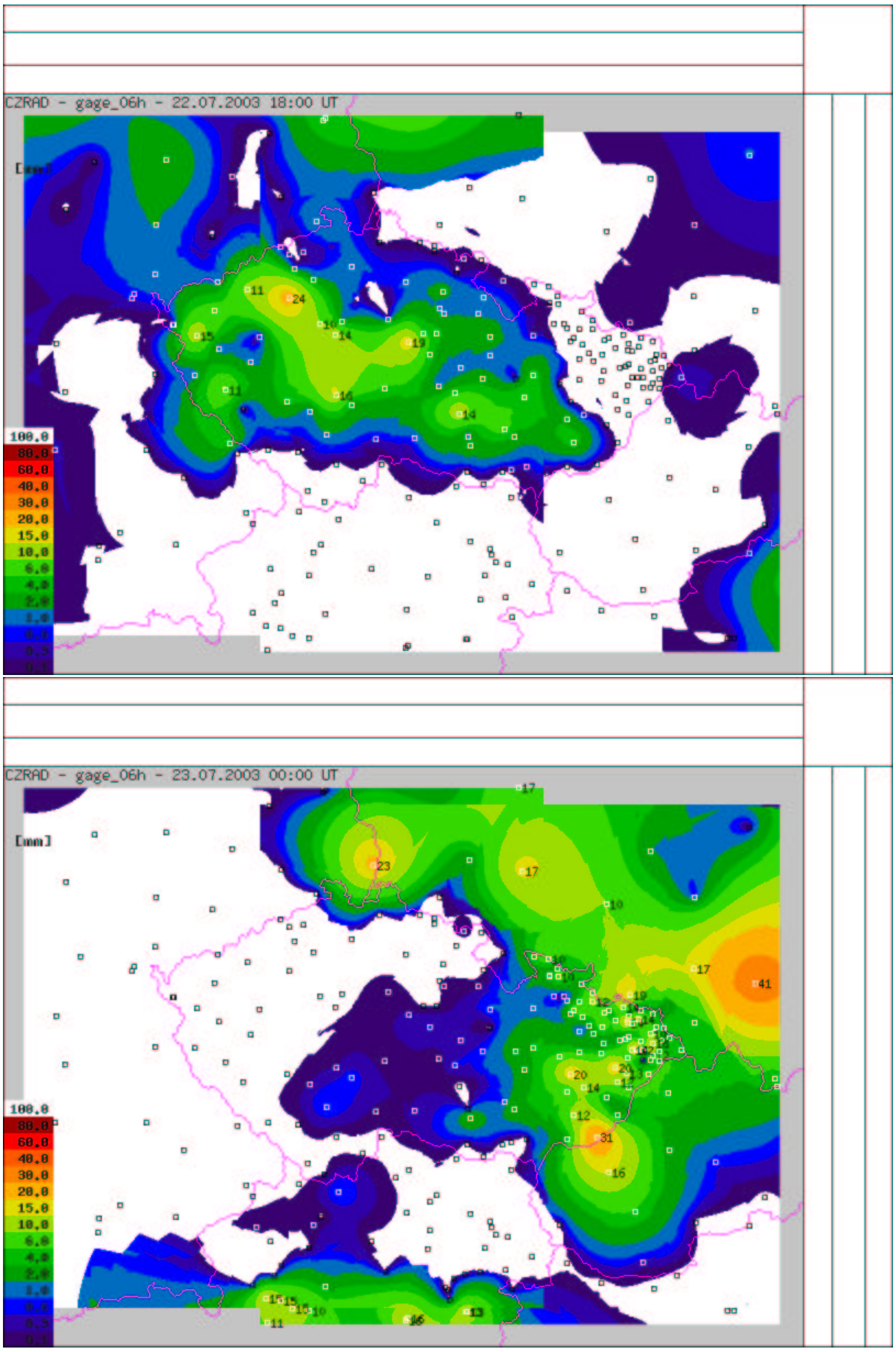


Figure 1: 6 hours precipitation measurements for the period of 12-18 UTC of 22nd of July and 18 UTC of 22nd - 00 UTC of 23rd of July

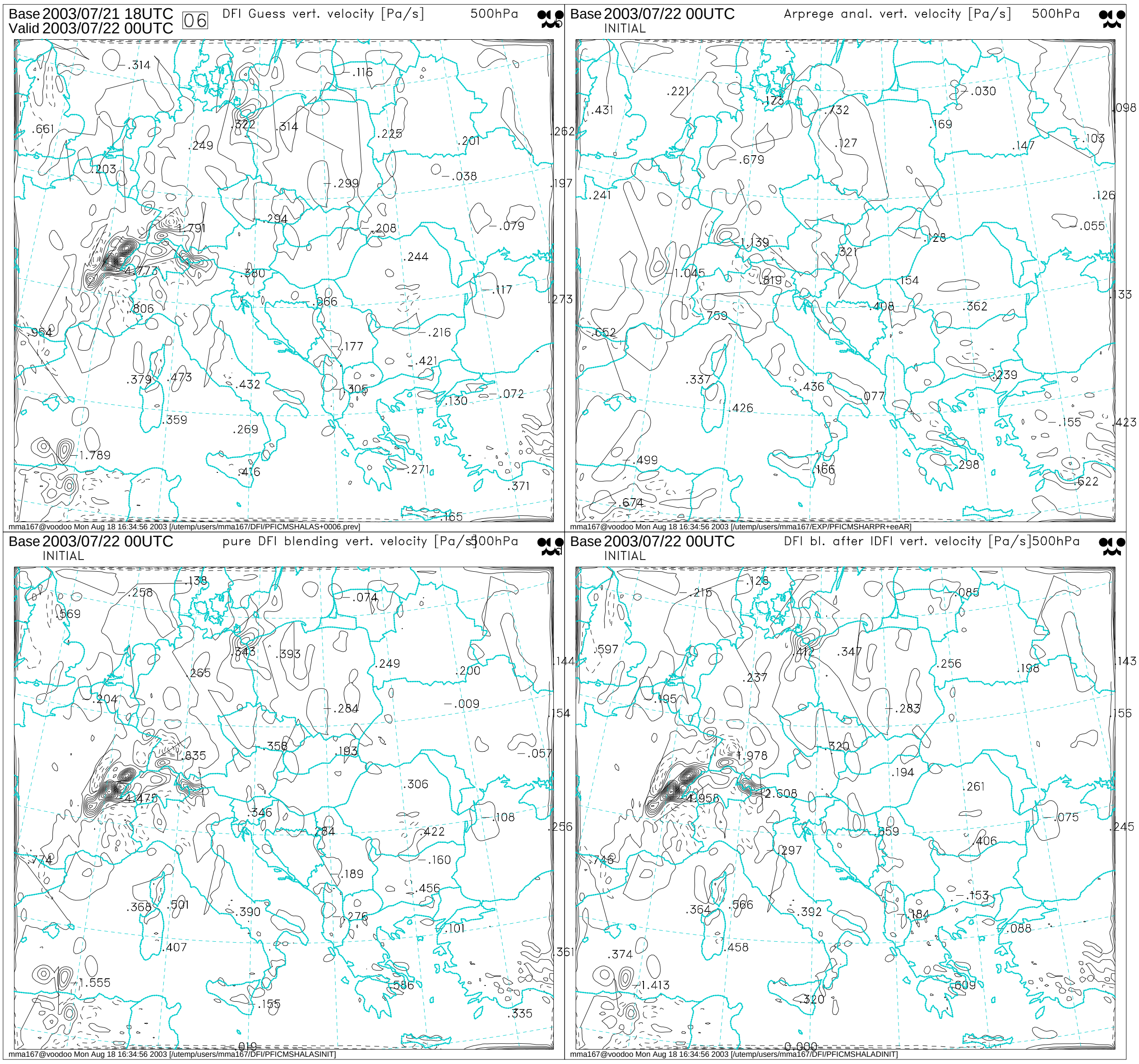


Figure 2: 500 hPa vertical velocity analysies fields (Aladin guess, Arpege analysis, pure DFI blending and DFI blending after IDFI)

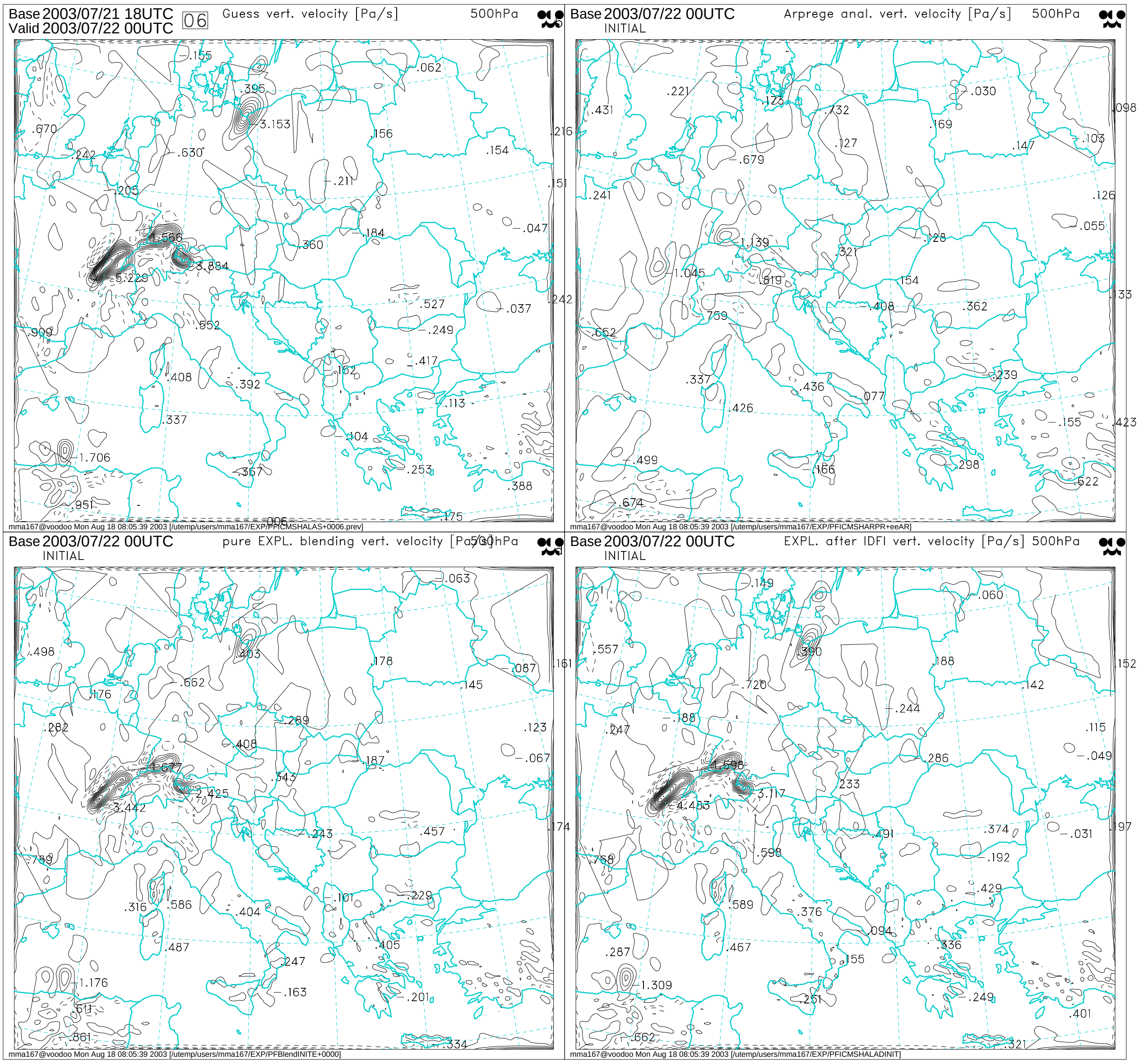


Figure 3: 500 hPa vertical velocity analyses fields (Aladin guess, Arpege analysis, pure explicit blending and explicit blending after IDFI)

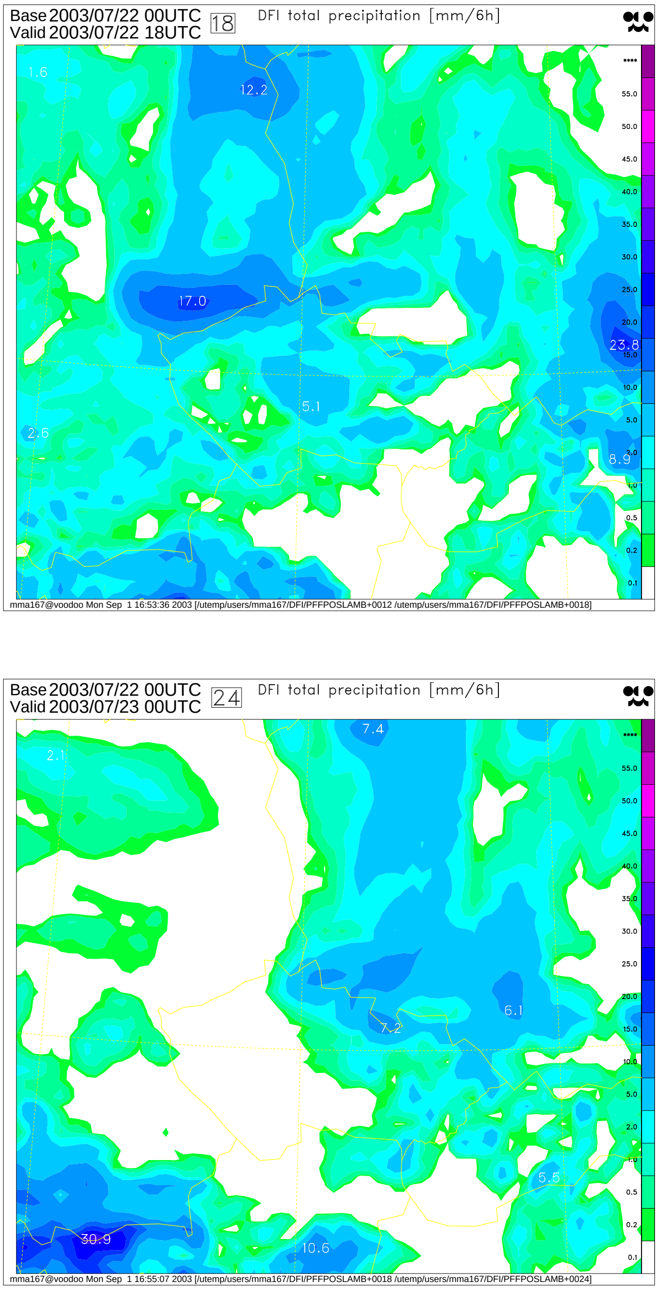


Figure 4: Precipitation fields produced by DFI blending

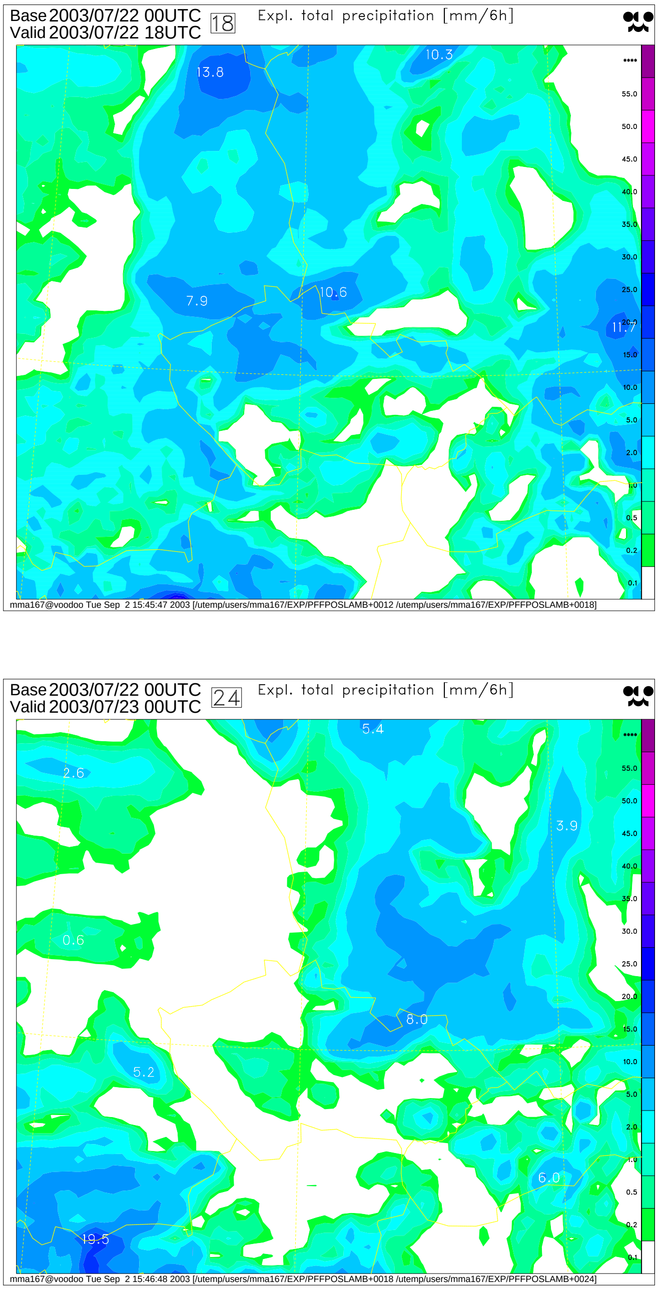


Figure 5: Precipitation fields produced by Explicit blending

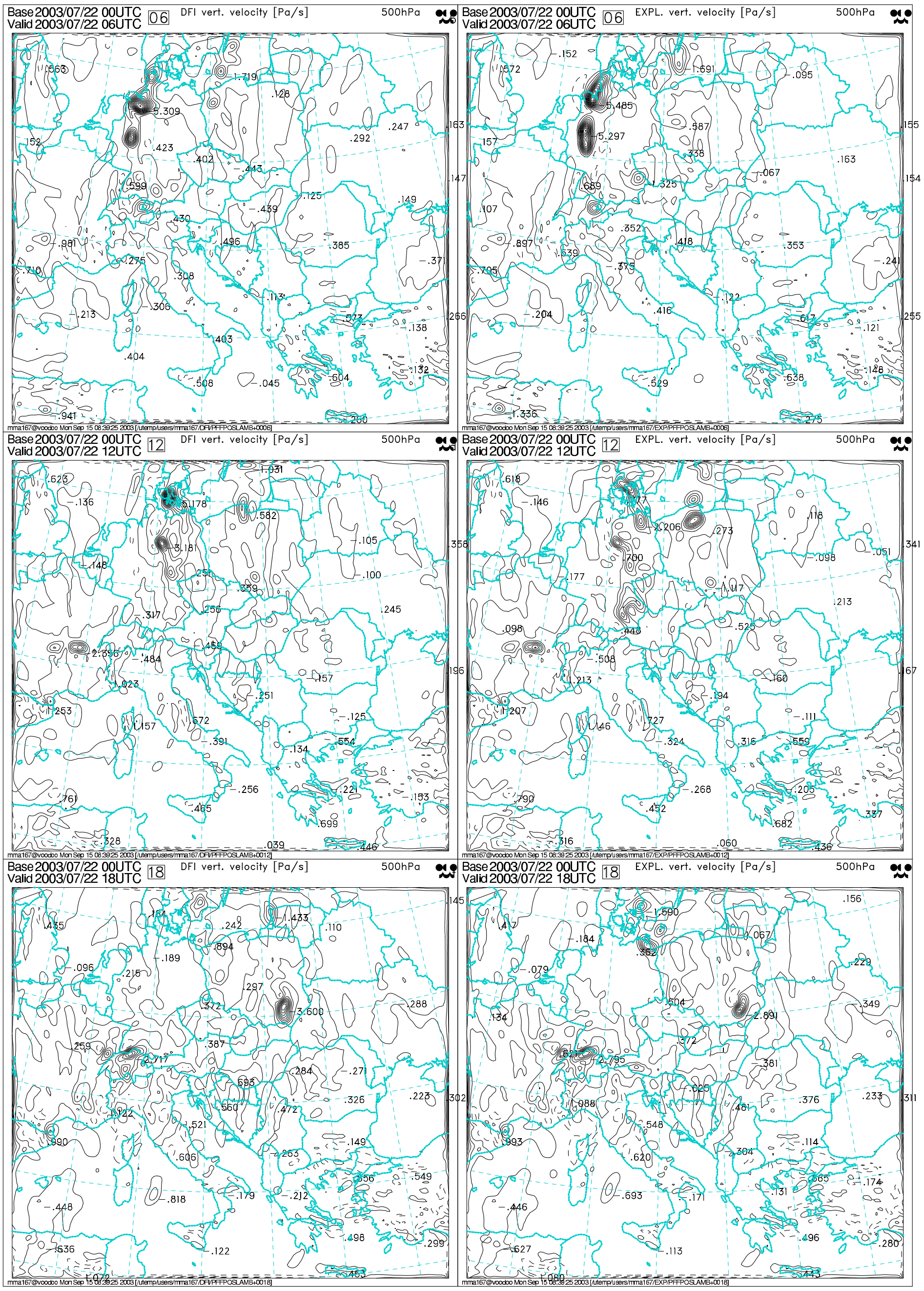


Figure 6: 500 hPa vertical velocity forecasts

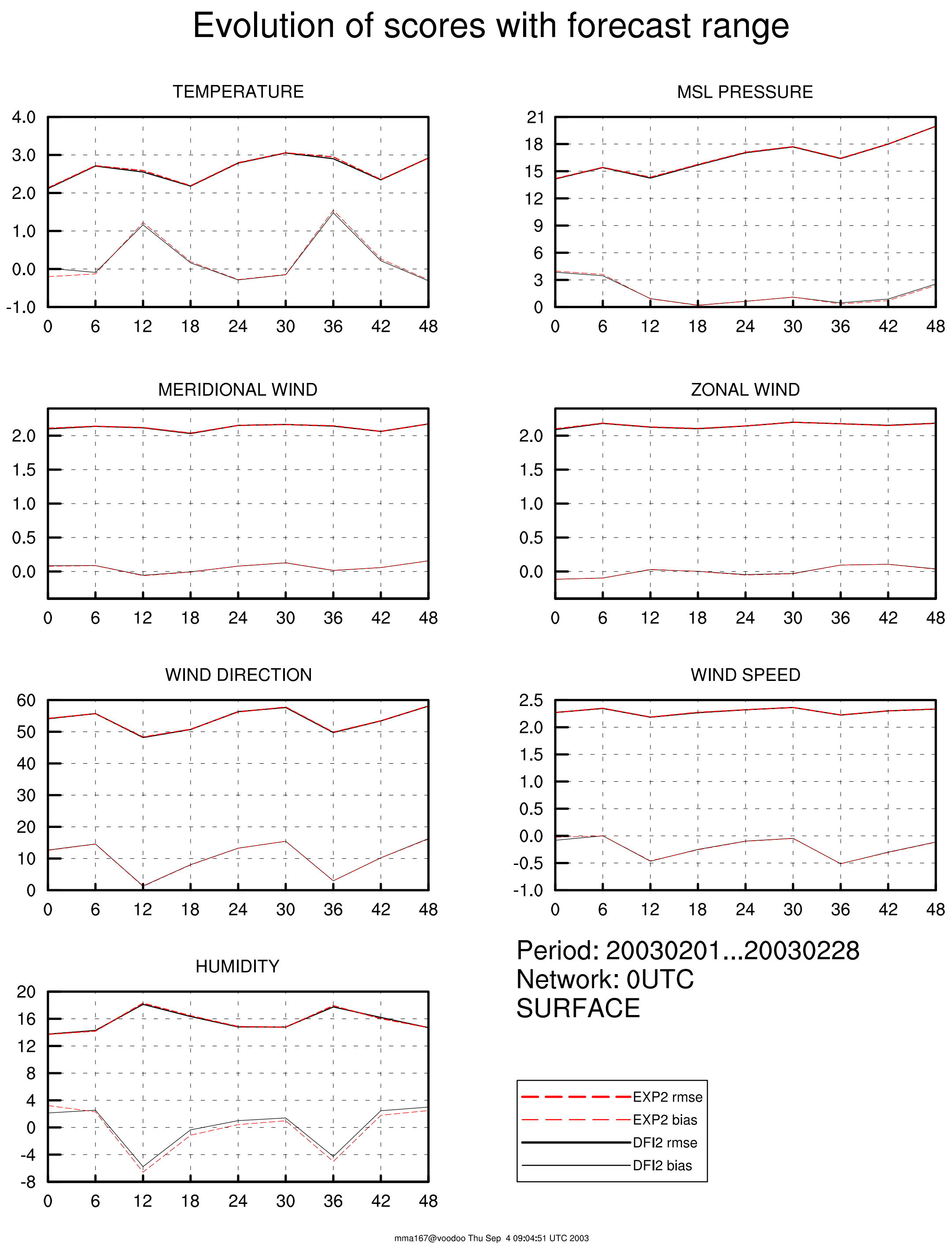


Figure 7: Result of verification of one month forecasts at the surface

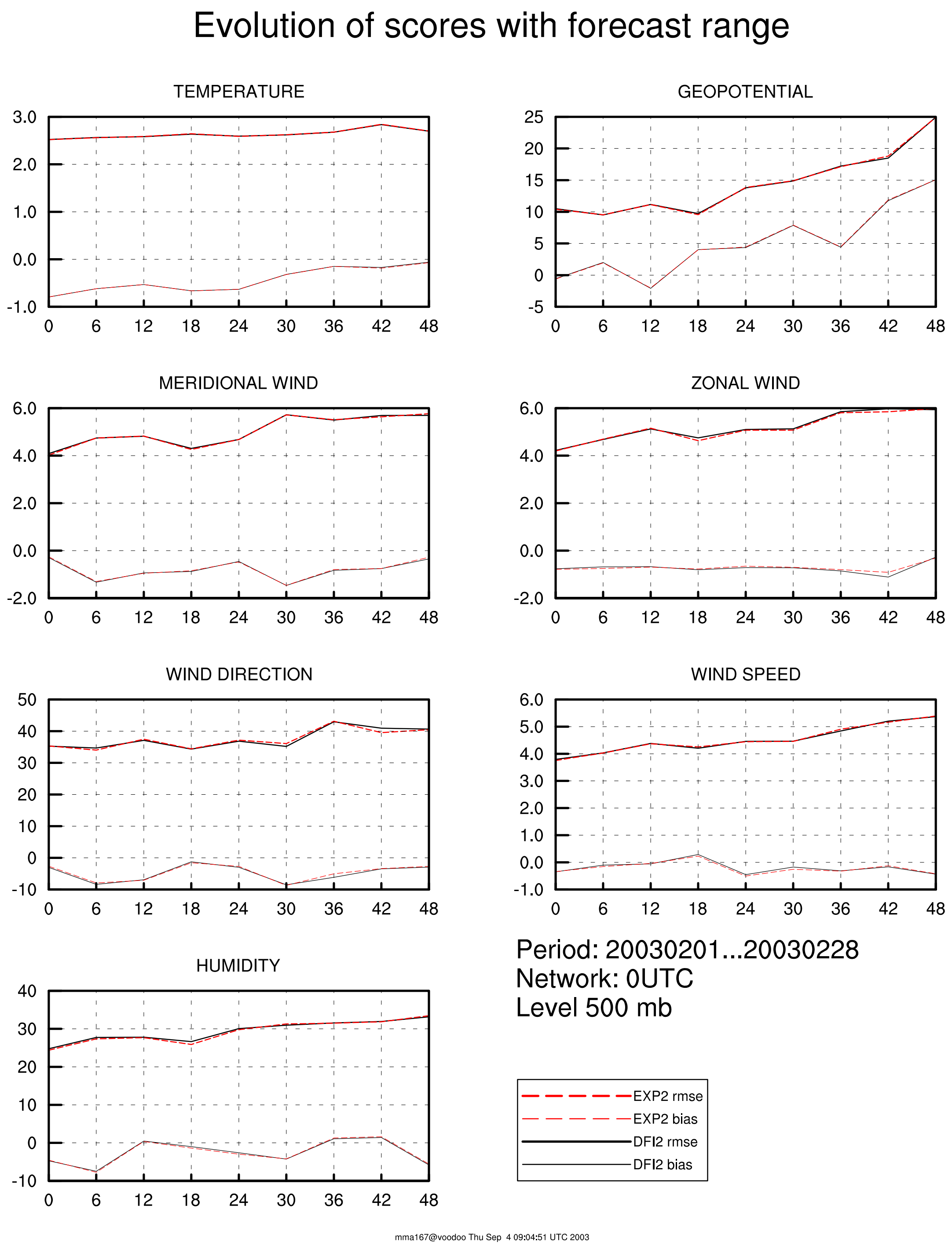


Figure 8: Result of verification of one month forecasts at 500 hPa

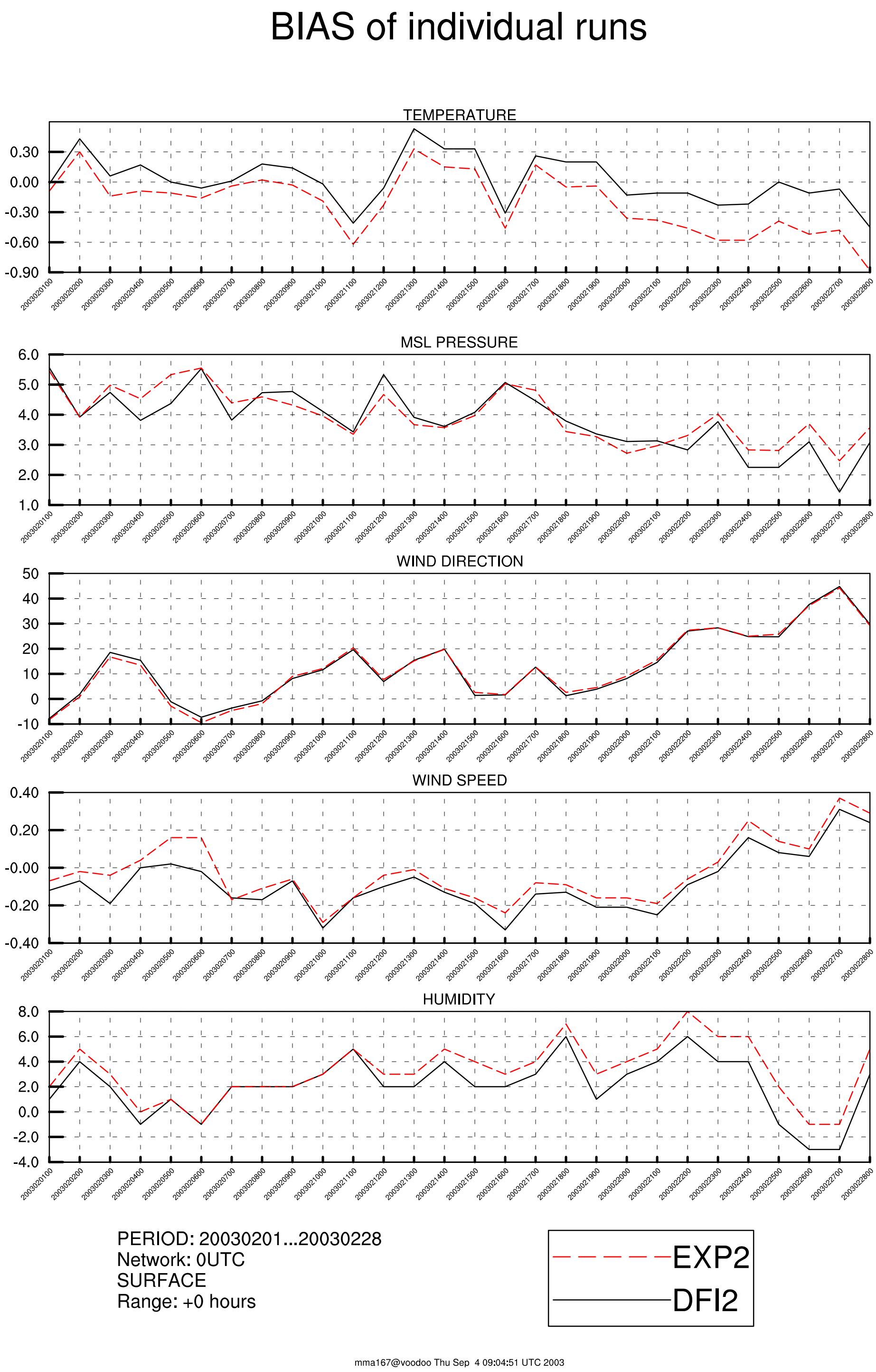


Figure 9: Evolution of the forecast scores at the surface for the beginning of the integration

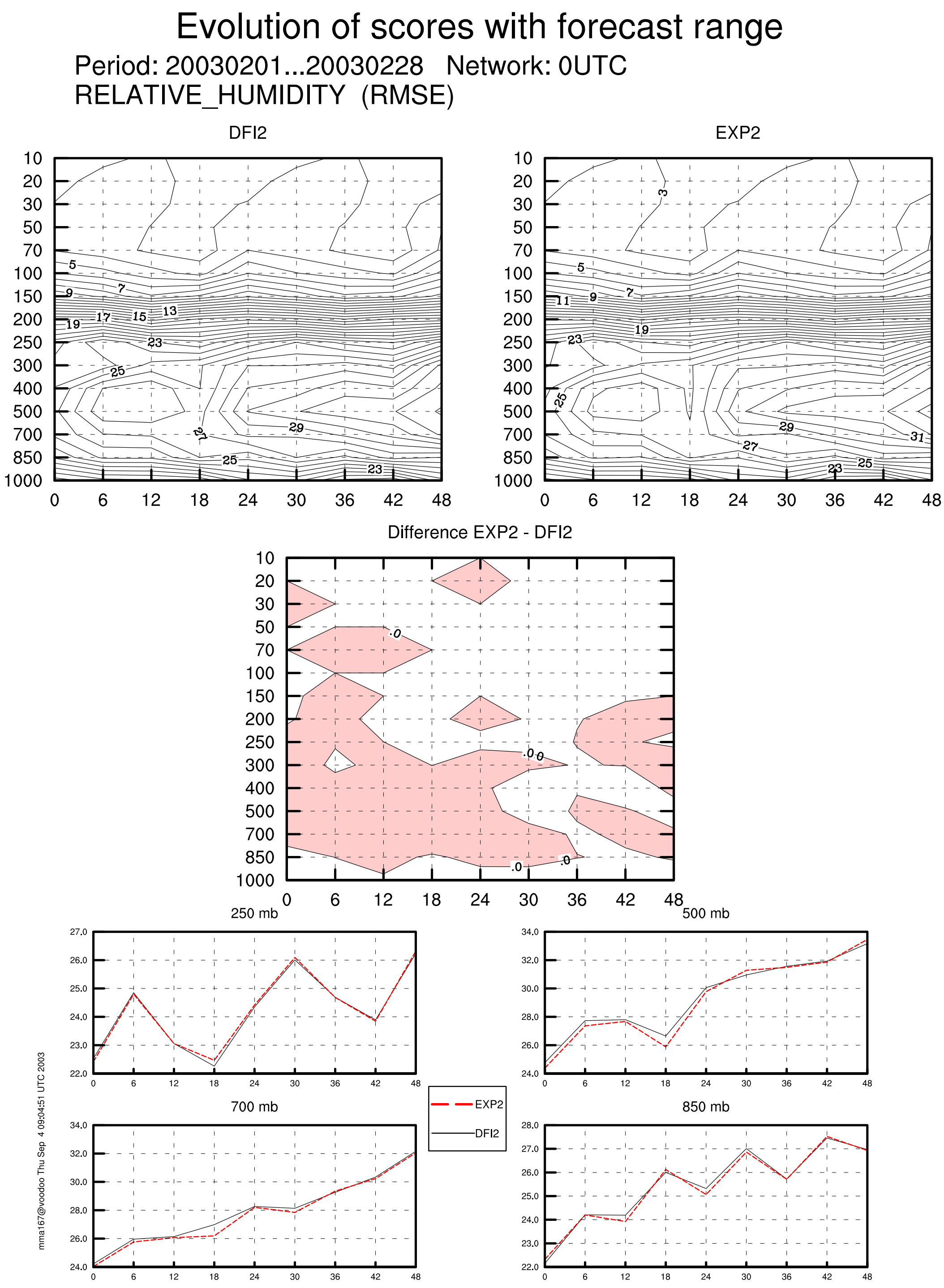


Figure 10: Relative humidity RMSE

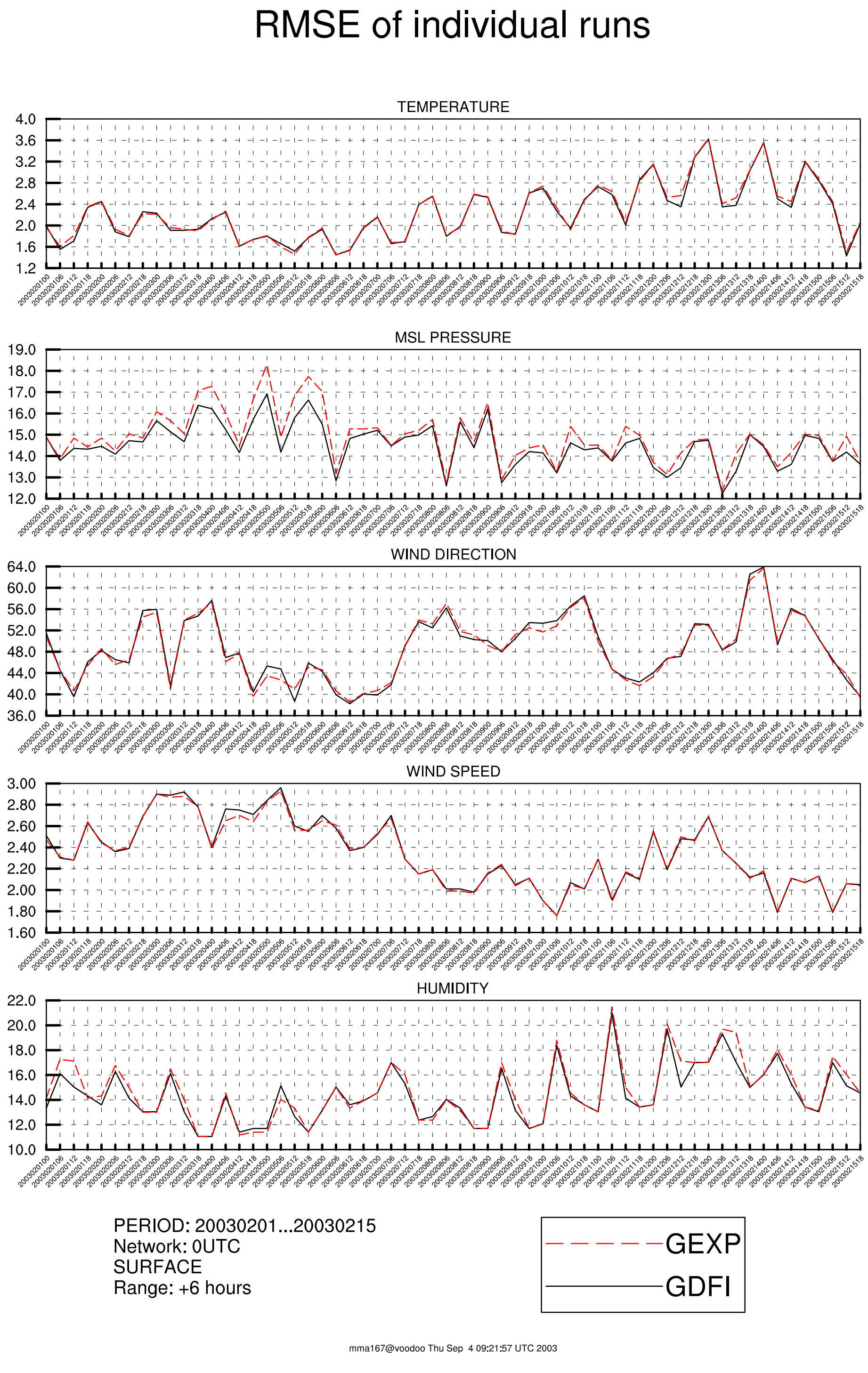


Figure 11: RMSE evolution of the two kind of blending guesses at the surface

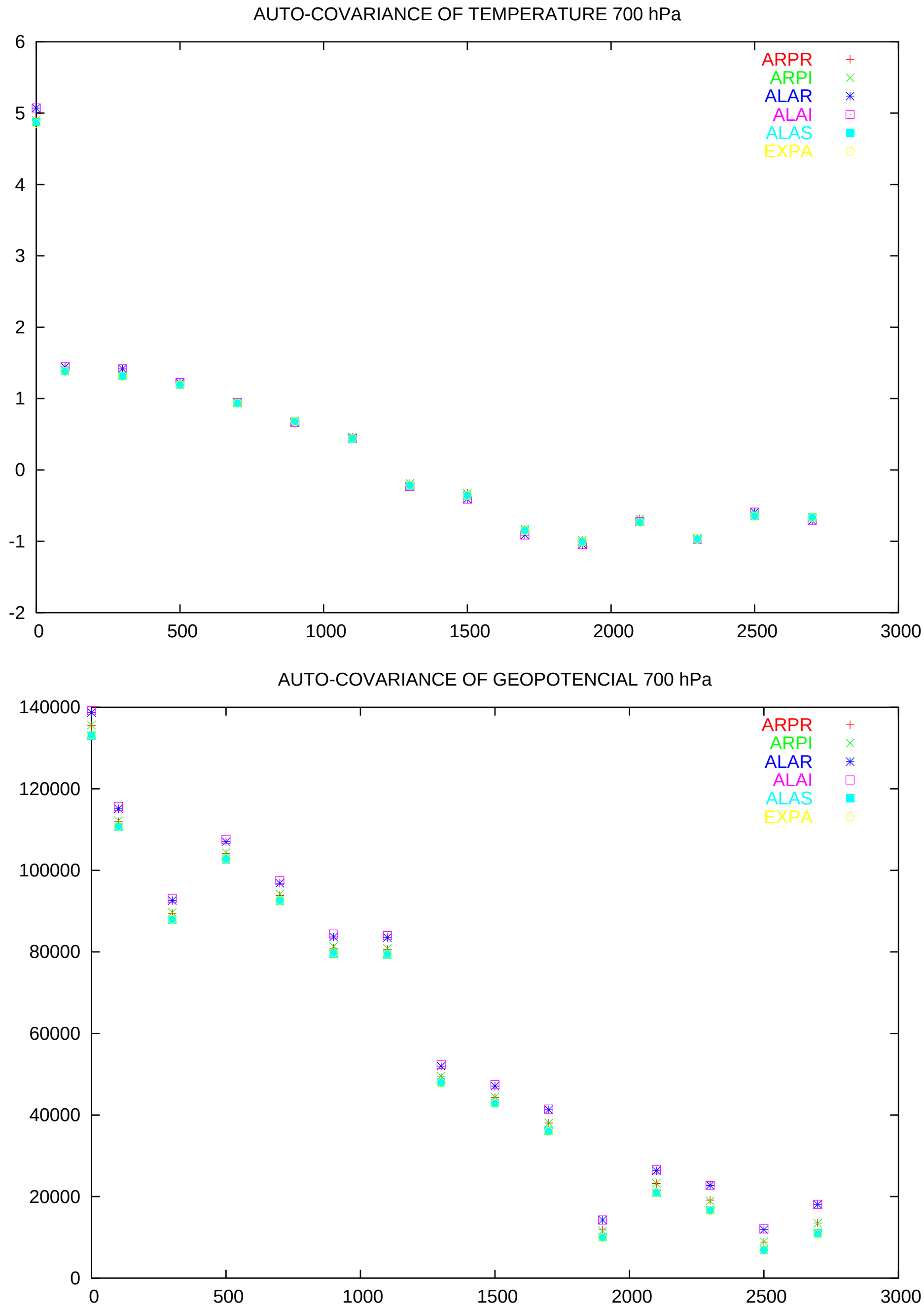
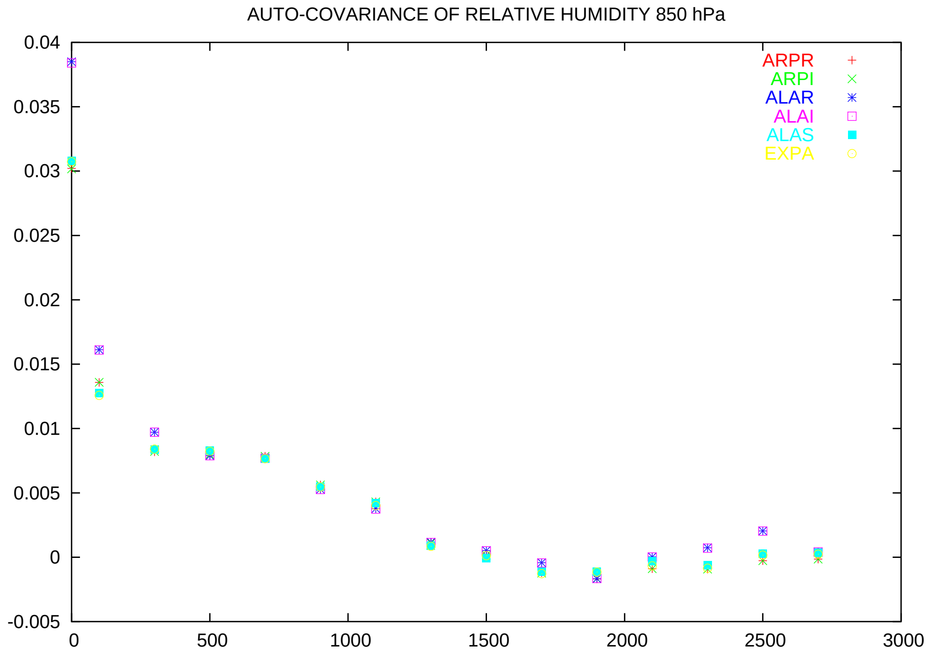


Figure 12: Result of LH-statistics 1



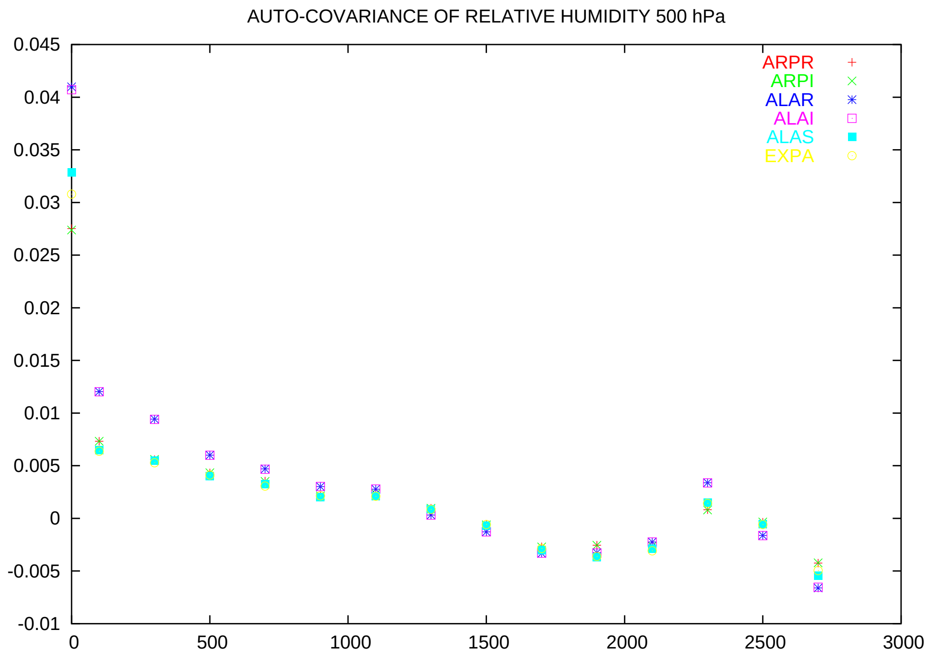
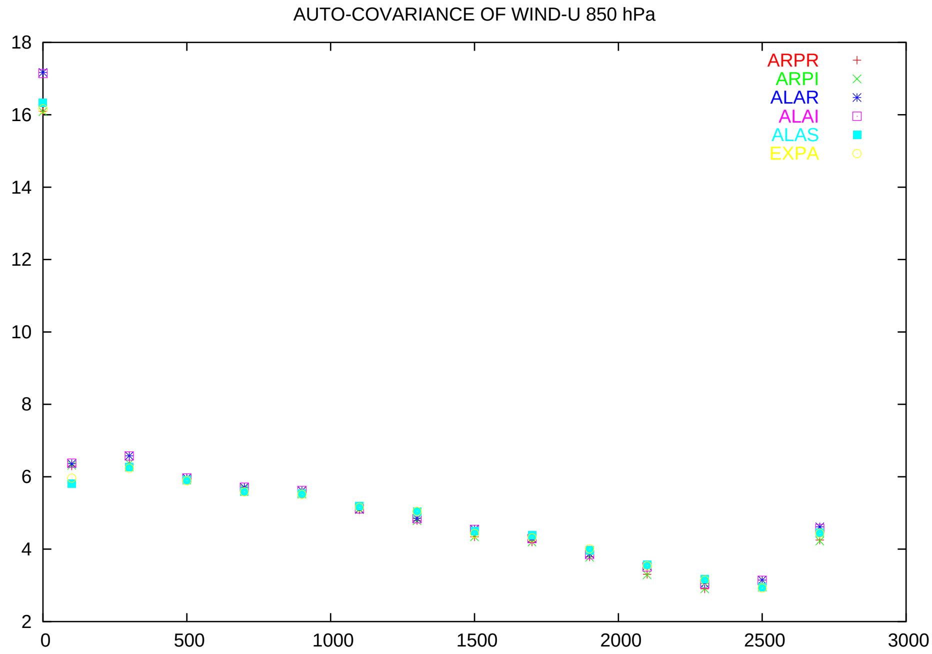


Figure 13: Result of LH-statistics 2



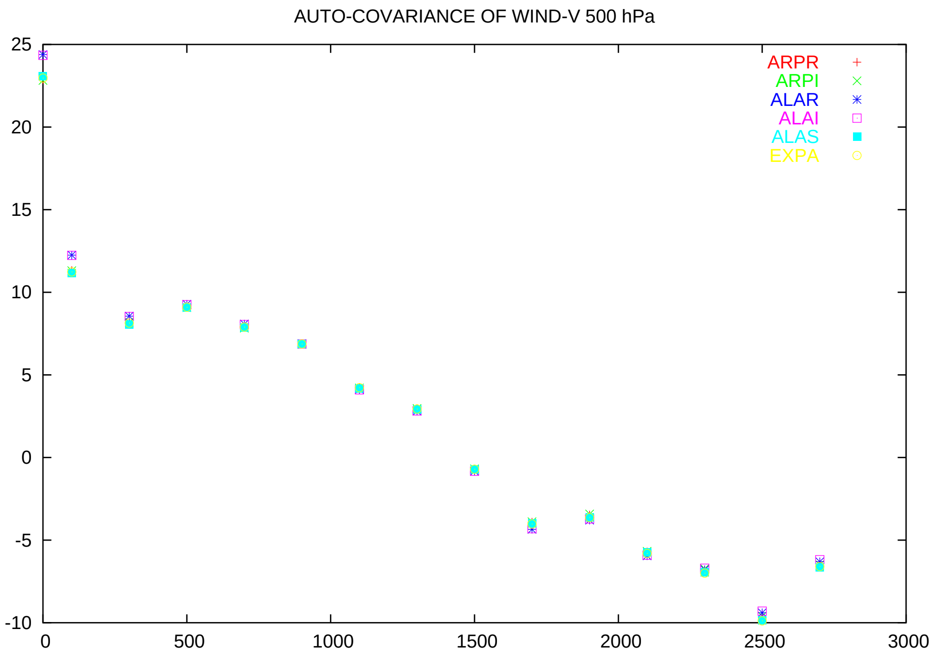


Figure 14: Result of LH-statistics 3