Report on stay at ZAMG
07/10/2013 – 15/11/2013, Vienna, Austria

Experiment with ECMWF-EPS-LBC dataset and test of an ensemble data assimilation system in an AROME-EPS framework

The fast growing uncertainty of small scale atmospheric phenomena always gives a good motivation for designing EPS on finer scale. In the world of improving numerical models and computer resources, convection-permitting EPS (based on non-hydrostatic models with resolution ~2-3km) is expected to become an affordable way (especially in limited area modeling) to represent uncertainties in the not-so-far-future. This expectation started to motivate AROME-EPS research work inside LACE community, too. At the time of being the work is still in preliminar status, and questions of used global models or perturbation methods are also opened.
The following document describes two tests connected to the two above-mentioned questions. Technical aspects are developed in an AROME-EPS framework and verification results are presented over a two-weeks winter-period. The first test was done with EPS-LBC sets with different resolution provided by ECMWF. The second test was done by a new version of AROME-EPS experimental set-up, where an ensemble data assimilation (EDA) module was implemented and switched-on.

Author:
Mihály Szűcs (OMSZ)
szucs.m@met.hu

Supervisor:
Theresa Gorgas (ZAMG)
theresa.gorgas@zamg.ac.at

Yong Wang (ZAMG)
yong.wang@zamg.ac.at

In the following document the first two parts can give an introduction over framework and agreement in connection with the exact tasks of my stay. Part 3. focuses on technical background information about my first task (ECMWF LBC test) and part 4. represents results. Part 5. is about the EDA tuned AROME-EPS, and part 6. is its pair with the results. Part 7. repeats the main conclusions and summarizes the connecter further plans how I see at the moment. Acknowledgment can be found in part 8.

1. Generalities

At the beginning of my LACE stay in Vienna first we agreed in some important technical question:

- The work was done on ECMWF's c2a machine.
- The here-represented developments are based on a shell script system which was the implementation of Francois Bouttier's system (Météo France).
- The test were done over the operational Hungarian AROME domain (320*500 gridpoint over Carpathian Basin with 2.5km resolution). There is a bigger AROME domain planned to be used in the future, which can cover more LACE countries, but at this stage of work computer resource efficiency was an important aspect.
- SBU (measure of ECMWF's computer resource) was spent from two ECMWF's special project, where Hungary is a participant: bigger part from 'spfrbout' (supervised by Francois Bouttier), minor part from 'spfrcoup' (supervised by Alex Deckmyn).

For better understanding of the results (in part 4 and 6) there are some notes about verification:

- The used verification system is a really simple one, developed and applied at Hungarian Met Service. There was an agreement that in the future common LACE verification package should be used also for convection-permitting system.
- Near-surface verification uses Metview and is based on SYNOP observation data at Hungarian SYNOP stations.
- First part of high-atmospheric verification uses also Metview and done against ECMWF's analysis. Its main advantage that provides a lot of information on this small domain as well. The main disadvantage is that its resolution can effect the results in a comparison.
- Second part of high-atmospheric verification is a simple shell script + FORTRAN routine which verifies model data against radiosonde observations.

2. AROME-EPS configuration on ECMWF's computer

Basically in this AROME-EPS script system every job produces one member of an ensemble for one case. It means, that if someone wants an N-day long test period with an M-member EPS, then N*M jobs have to be launched. Originally AROME-EPS members are the dynamical downscaling of PEARP members, so jobs can be run independently. Usually AROME-EPS contained 11 members which was the downsclaing of the first 11 members of PEARP.
In every job there are several steps. Every jobstep has a well-defined function and they have very different LoadLeveler (ECMWF's job scheduler) settings. Scripts of the jobsteps are technically also separated in different files.
The original configuration contained the following steps with the captured functionality:

1.	stget	It finds a free working directory, makes the cleaning and copies necessary files (e.g clim files) there.
2.	starp	As far this system is based on PEARP downscaling, there is a function which rerun ARPEGE, if it is necessary and if global IC files are available.
3.	stcpl	e927 configuration for ARPEGE files. There is a possibility to use ALADIN files and run ee927, as well. It can be useful if someone wants to use ALADIN for the surface or limited area EPS (like LAEF) as a mother-model
4.	stble	It is a function just for a special case when we want to use local AROME assimilation with downscaled PEARP perturbations.
5.	staro	001 configuration.
6.	stsav	Saving files.

Table 1.: Jobsteps in the basic AROME-EPS configuration

Every job has a special name which contains the following parts:

- Date and time
- Number of member
- Experiment ID

The first two part is obvious and the third needs some explanation. ExpID is six character long, where the first character is H and the following five character defines the following:

2nd character	AROME settings (e.g. SPPT active or not?)
3rd character	Global model to IC (e.g. PEARP or LAEF or HUNEPS)
4th character	Way of IC usage (e.g. simple downscaling or single AROME-DA centralized)
5th character	Global model to LBC (e.g. PEARP or LAEF or HUNEPS)
6th character	Way of LBC usage (e.g. it is possible to use control LBC for every member)

Table 2.: Experiment ID explanation

The most basic experiment is abbreviated as the following: Hgabab, where in order: g – AROME set-up without SPPT; a – PEARP as a global system for ICs; b – kth AROME-EPS member uses kth PEARP member as an IC; a – PEARP as a global system for LBCs; b – kth AROME-EPS member uses kth PEARP member as a set of LBCs.

The name of the jobs are in special job-files which are created a simple script. This script has to be run to prepare a test period, and the following things have to be defined:

- First and last date
- Time of the integrations
- Experiment ID
- Number of members in the system

Table 3. can help the better understanding of the process which was detailed in this part.

Test	Jobs	Jobsteps
N day long test, with **M members** and named with **EXP-ID**	**1st** day, **1st** member	1. stget
		...
		6. stsav
	...	...
	1st day, **Mth** member	1. stget
		...
		6. stsav
	2nd day, **1st** member	1. stget
		...
		6. stsav
	...	...
	Nth day, **Mth** member	1. stget
		...
		6. stsav
...	...	...

Table 3.: Independently from the visualized test there could be more tests as well. Jobs are also totally indepent. Jobsteps has to follow each other automatically with different LoadLeveler settings.

3. Technical remarks about usage of ECMWF-EPS-LBCs

The previous part described very briefly the AROME-EPS configuration which was available at the beginning of my LACE stay. For testing the two sets of LBCs provided by ECMWF, some script system modification was needed before jobstep 4 (AROME integration).

It has to be noted that ECMWF global model data was not stored on the conventional way, because of disk-space efficiency. Only reduced grid files were archived in MARS instead of spherical harmonics, which can be used in e927 ALADIN configuration. This technical problem motivated using GL, which is a skillful tool developed by HIRLAM. GL is able to interpolate grid represented fields from grib files to FA files on a relatively fast and cheap way.

For the above-mentioned reasons two new jobsteps were added to the original system. If ECMWF-LBCs are used, ARPEGE rerun is unnecessary of course, but technically every possible jobsteps becomes active during the run of a job. The unnecessary jobsteps finally do not do anything. For better understanding Table 4. is the upgraded version of Table 1.

Another remark is about lfi SURFEX files: While surface scheme of ECMWF is really different, at Hungarian Met Service the operational 'deterministic' AROME (which uses also ECMWF LBCs, but there is no surface assimilation) interpolates ALADIN analysis for surface files. In ALADIN

assimilation suite CANARI is active.
The upgraded configuration added the following steps to the original system:

1.	stget	Same than in Table 1.
2.	stmars	It downloads necessary files from MARS in two steps. In first step it downloads all the fields which are needed from a given timestep for a domain which is smaller than the big reduced grid used for storage. With the following MARS retrieve it downloads from the previous files the fields which are necessary in an LBC file. This two step is necessary for example because surface or orography data is not archived in MARS for every timestep but they have to exist in every LBC file (e.g. land-sea mask is archived only in +0hour files in MARS, but have to be added to +17hour coupling file).
3.	stgl	For the output of the previous step GL can be run quite easily. Of course the same climate file is used than in an e927 configuration. Initialization of non-hydrostatic fields and changing of number of vertical levels also need some attention at this part.
4.	starp	This step becomes active and does noting, if we couple to ECMWF.
5.	stcpl	e927 configuration for ARPEGE files. There is a possibility to use ALADIN files and run ee927, as well. It can be useful if someone wants to use ALADIN for the surface or limited area EPS (like LAEF) as a mother-model
6.	stble	Same than in Table 1.
7.	staro	Same than in Table 1.
8.	stsav	Same than in Table 1.

Table 4.: Jobsteps in the upgraded AROME-EPS configuration which allows to use ECMWF-EPS-LBCs

4. Verification results with ECMWF-LBCs

ECMWF provided two sets of LBCs: one with the current EPS resolution (T639~32km) and another one with the current 'deterministic' resolution (T1279~16km).
ECMWF tests were available for three periods. The so called winter-period (**26.12.2011-08.01.2012.**) was chosen from them, because this one fits the best to the topic of 'spfrcoup' special project (see part 1.).
Two runs per day (**00UTC** and **12UTC**) were available with **20+1 members.**

In the figures which are showed in this part the following naming conception is vaid:
ECHRcoup00 – Coupled to **high-resolution** LBC set started at **00UTC** (control + first 10 perturbed members)
ECHRcoup12 – Coupled to **high-resolution** LBC set started at **12UTC** (control + first 10 perturbed members)
ECLRcoup00 – Coupled to **low-resolution** LBC set started at **00UTC** (control + first 10 perturbed members)
ECLRcoup12 – Coupled to **low-resolution** LBC set started at **12UTC** (control + first 10 perturbed members)

ECLRcoup00m21 – Coupled to **low-resolution** LBC set started at **00UTC** (control + 20 perturbed members)
PEARPcoup18 – Coupled to **PEARP** at **18UTC** (~18km resolution over Hungary, control + first 10 perturbed members).
All the tests (even the PEARP downscaling) was done with SURFEX lfi files downscaled from ALADIN analysis.

The representation of the results are focusing on the following questions:
- *What is the difference between AROME-EPS tests coupled to ECMWF-EPS with low or high resolution? (high-atmosphere against ECMWF's analysis)*
- *What has bigger impact on AROME-EPS: the resolution of the global model or the number of the members? (high-atmosphere against ECMWF's analysis)*
- *How can PEARP and ECMWF-EPS based AROME-EPSs perform against each other? (high-atmosphere against radiosonde observations)*
Usually only two figures are shown for every question (a typical and an interesting one) but the availability of lot more figures is also described.

Low-resolution vs. high-resolution coupling files

This is an easy comparison, because the positive impact of HR-LBC files is obvious. Its scores are better than LR-LBC coupled versions. The difference is bigger on higher-atmosphere, especially for geopotential. The difference is smaller on low levels and near-surface.

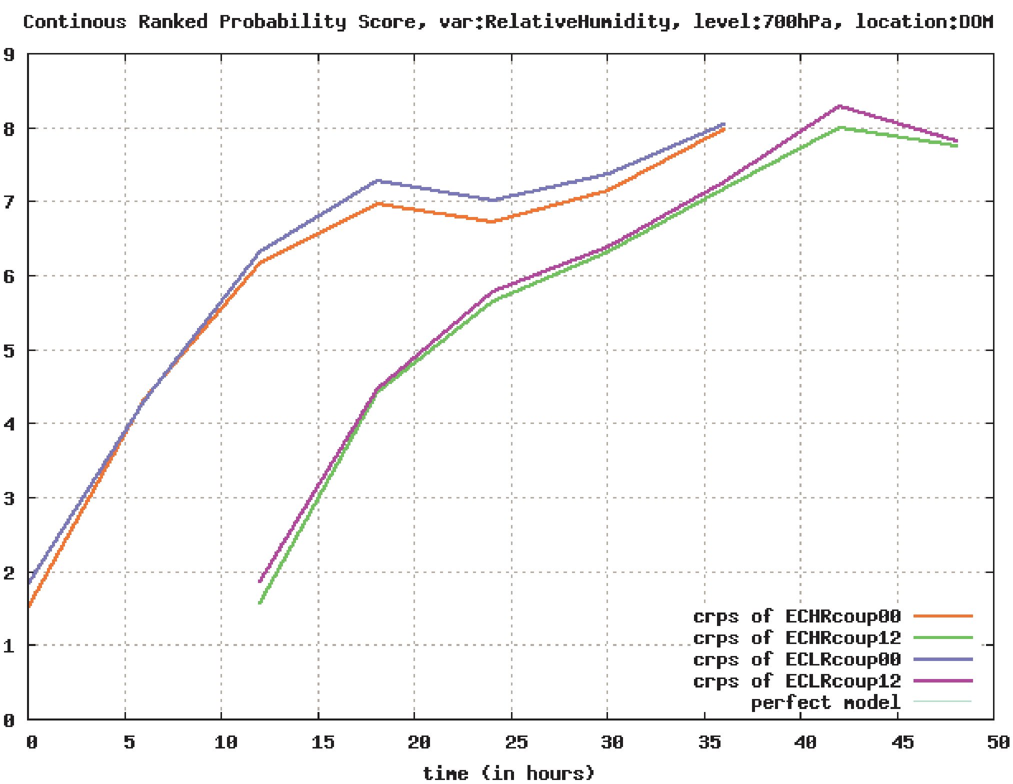

Fig. 1.: CRPS of relative humidity at 700hPa. Results with HR-LBCs are slightly but significantly better. This is valid for most of the figures from the same comparison.

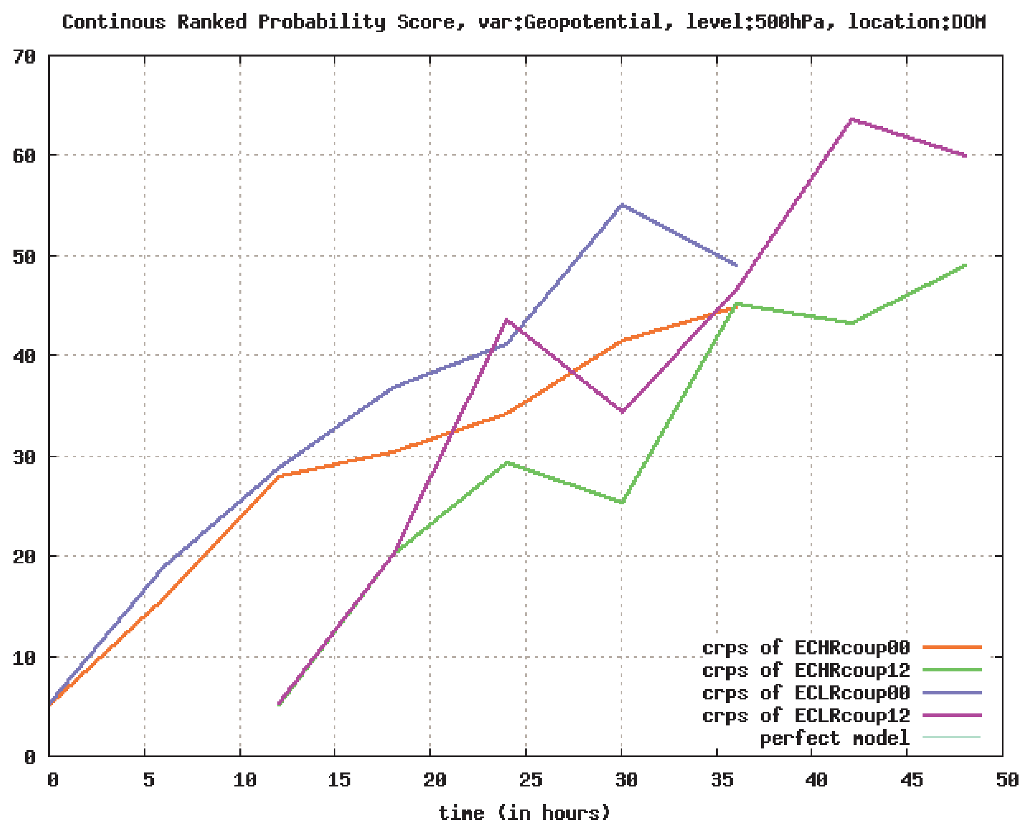

Fig. 2.: CRPS of geopotential at 500 hPa. The advantage of HR-LBCs are even bigger. While this is a verification against ECMWF analysis (which has the same resolution than HR-LBCs), probably the verification method serves also HR-LBCs more when synoptic scales are described.

Impact of increasing global model resolution vs. impact of increasing number of members

This comparison can not really decide which impact is stronger. In scores it is lead-time dependent and in high-atmosphere global model resolution looks crucial, especially for geopotential.

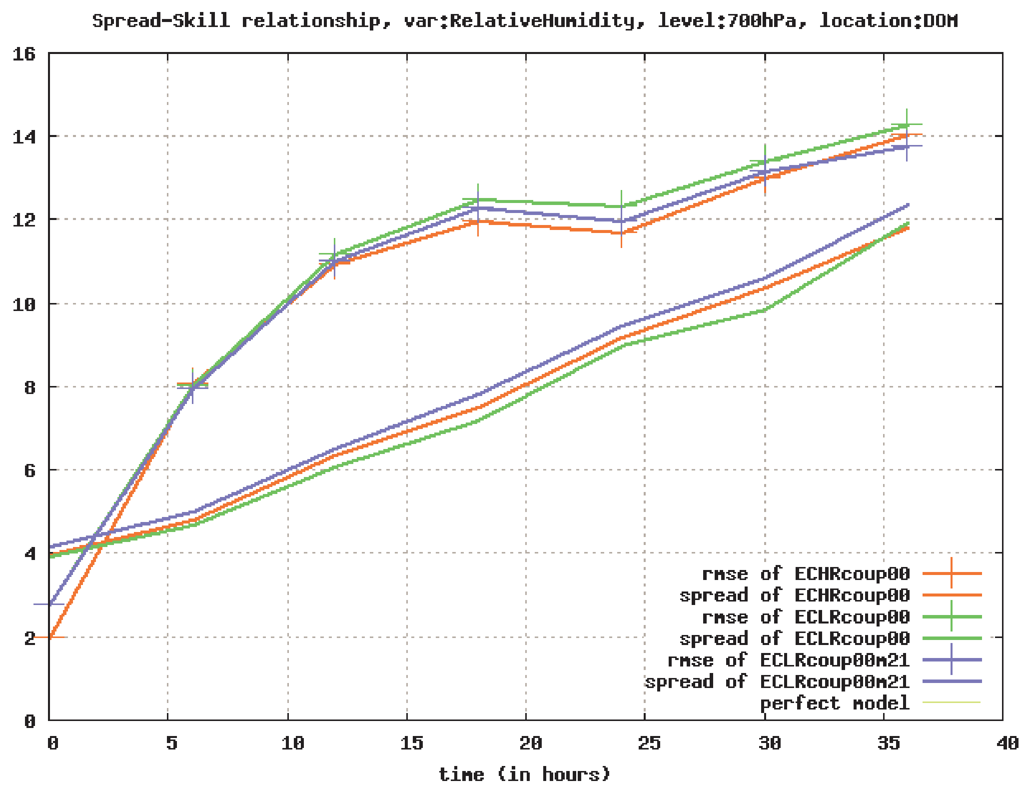

Fig. 3.: Spread-skill relationship of relative humidity at 700hPa. For most of the variables it is hard to say if doubling of members or doubling the resolution of global model has a bigger impact. 21 members could make the biggest SPREAD of course, but usually high-resolution coupling can also increase the SPREAD comparing with low-resolution coupling, especially in the second part of the forecasting time. High-resolution coupling produces the best RMSE, but at the end of the forecast, the mean of 21 members has also better quality than the mean of 11 members.

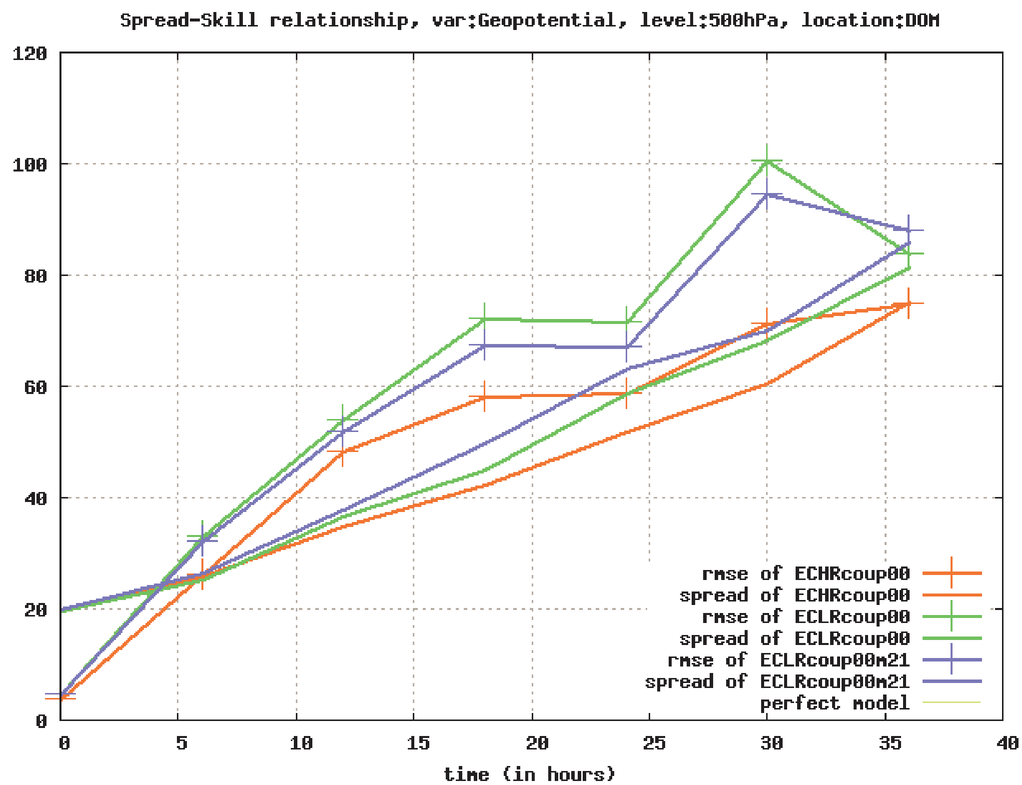

Fig. 4.: Spread-skill relationship of geopotential at 500 hPa. While impact of resolution of the global model is big on geopotential fields, it is obvios that from this point of view the impact of global model's quality outperforms the impact of number of members.

ECMW-EPS coupled AROME-EPS vs. PEARP coupled AROME-EPS

This is a hard comparison, mainly because of two reasons:

- Systems do not start from the same time.
- ECMWF analysis can not be used as independent data in verification at this comparison. This reason motivated the usage of radiosonde data, but while the domain is small, the size of the sample for verification is quite limited. Because of this reason mainly near-surface scores are in our focus.

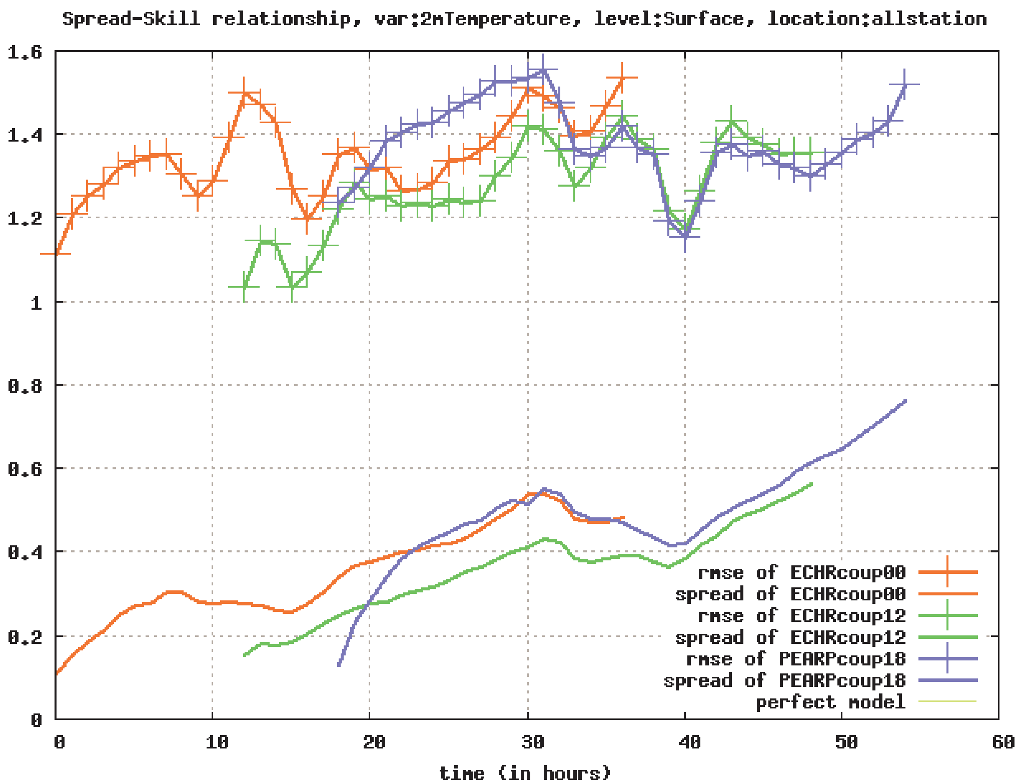

Fig. 5.: Spread-skill relationship for 2meter temperature. This is a typical results: Hard to define which version is the better, they are really similar. Probably in general PEARP coupled version has a bit bigger RMSE but bigger SPREAD as well.

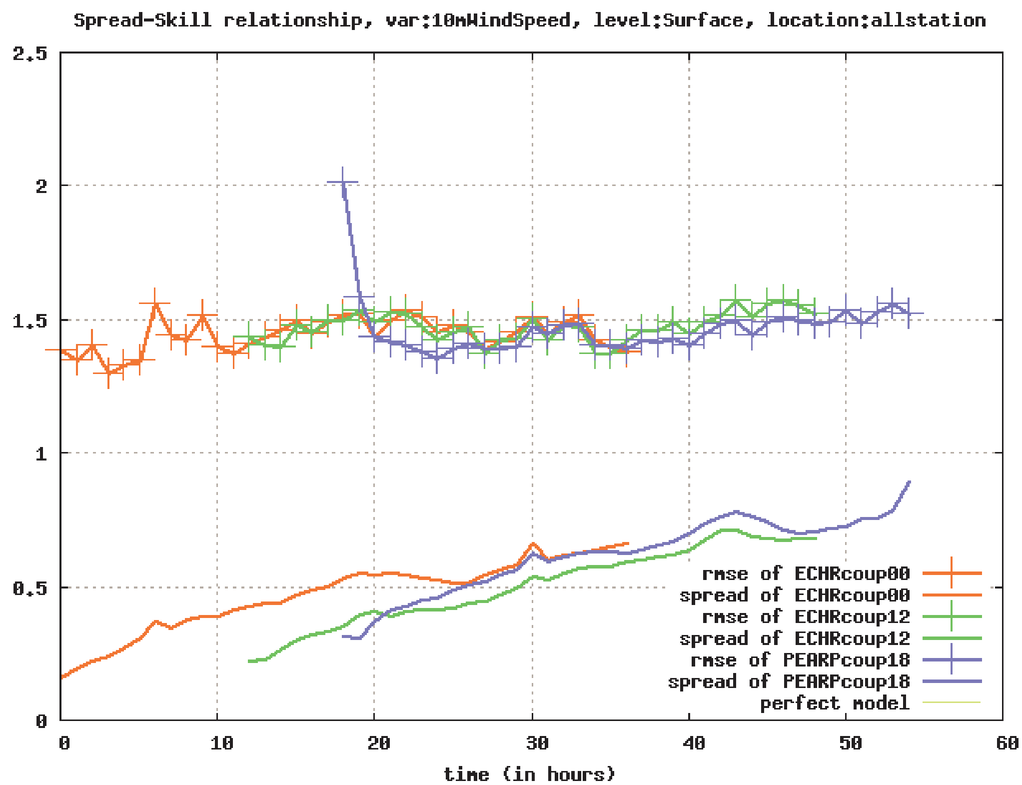

Fig. 6.: Spread-skill relationship of 10meter wind speed. Unfortunately a typical problem can be seen at the initial time of the PEARP coupled run. It is nor clear what is the reason of this big error (for BIAS see fig. 7.), because e927 is done on the same way than in operational jobs of OMSZ.

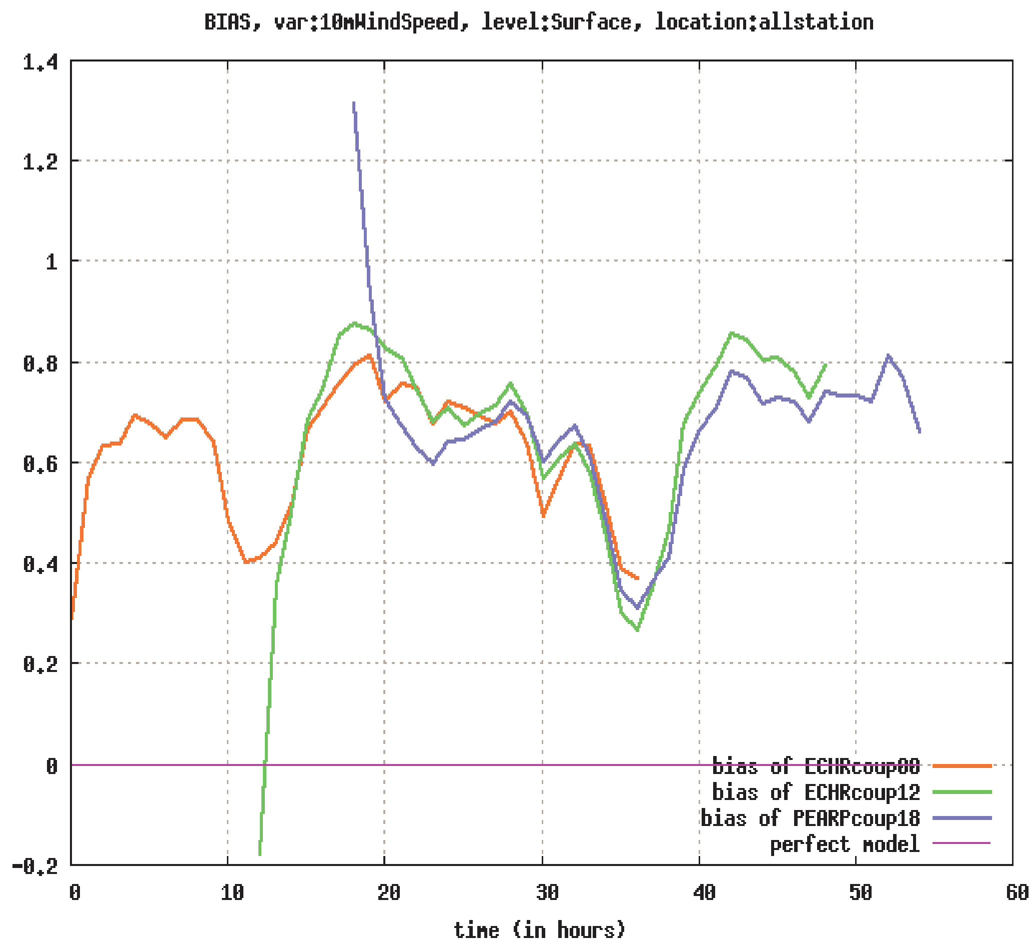

Fig. 7.: BIAS of 10meter wind speed. Annoying and strange overestimation at the beginning of PEARP coupled version. However it can be seen, that AROME always has a slight overestimation, independently from the used global model.

For more verification results, see ECMWF's ecgb:
/home/ms/hu/hu7/AROME-EPS-LBCtest1

5. **Designation of the AROME-EDA test suite**

At the beginning of my stay it was decided that AROME-EDA would be designed on the simplest possible way. It meant that in the practice, that:

- For 10+1 AROME members belonged 10+1 data assimilation,
- Only high-atmospheric 3DVAR was used,
- Only conventional data was used,
- Assimilation frequency was 3 hours,
- Cycle 36 was used,
- The whole system was coupled to PEARP.

The operational 'deterministic' AROME-DA at OMSZ is in accordance with the last two criteria and run in a similar platform than ECMWF's supercomputer, so mainly these scripts were moved and implemented to c2a.

On the level of jobs implementation of EDA meant adding a lot of new jobsteps. These steps are related to specific steps of data assimilation, or just some file manipulation.

In this EDA all the observations are perturbed to represent the uncertainty of the analysis. Technically perturbations are added after screening by an external program written by Andrea Storto (met.no). This program reads observations and their error. For all the observations it generates a random number which is from a gaussian distribution with 0 mean and 1 variance. Then observation error is multiplied by this random number and the result is added to the observation itself.

Table 5. is the extended version of table 4. with the new jobsteps which are used in the EDA.

1.	stget	Same than in Table 4.
2.	stmars	Same than in Table 4. Because of PEARP coupling, this jobstep does not do anything.
3.	stgl	Same than in Table 4. Because of PEARP coupling, this jobstep does not do anything.
4.	starp	Same than in Table 4.
5.	stcpl	Same than in Table 4.
6.	stgob	This jobstep gets the observations. At the moment this data is uploaded to ecfs from the Hungarian operational archive. For short cut-off analysis the same observations are used in this AROME-EDA than in 'deterministic' ALADIN was used.
7.	stpob	This jobstep prepares the observations.
8.	stcan	This jobstep is deserved to surface analysis. In the present configuration it does not do anything.
9.	stadd	While guess files does not contain some surface fields, they have to be added from somewhere before the assimilation. Most of these fields are needed just technically and are replaced at the beginning of the model integration from lfi files. Some of them plays role at observation operator. At the moment constant surface parameters (e.g. land-sea mask, orography) are from climate files, and soil parameters (e.g. PROFTEMPERATURE) are from the downscaled LBCs
10.	stscr	This jobstep runs screening.

11.	stperu	This jobstep runs the external program which is responsible for the observation perturbations. The binary is from an ALADIN package created by gmkpack. Gmkpack needed some trick and additional files to be enabled to create this special binary.
12.	stmin	This jobstep runs minimalization.
13.	stble	Same than in Table 4.
14.	staro	Same than in Table 4. In case of long cut-off it is just a 3-hour long integration which creates the guess. In case of short cut-off it has the normal time of a production (here 36 hours).
15.	stsav	Same than in Table 4.

Table 5.: Jobsteps in the AROME-EPS where EDA is also implemented.

On the level of tests, AROME-EDA script system needed a lot of development. This is mainly because of two new challenges:

- Jobs are not independent anymore. Any runs (after the cold start) needs a high-atmospheric guess file as an input. While ALADIN analysis is available in every 6 hour and assimilation frequency in EDA is 3 hours, for every second assimilation surface guess files are also needed from a previous run.
- Previously AROME was started from the same time when global model was available. Of course in a data assimilation suite there is a need on 8 time 3 hours long forecast to create the guess files. While global models do not run with that frequency, LBCs have to be used in lagged mode.

In accordance with the new challenges table 6. is the upgraded version of table 3., which visualizes the structure of a test. First there is a cold-start for all the members, then long cut-off analysis are done and guesses are created. Finally in the time of productions, short cut-off analysis are done and 36 hour-long model integrations are run.

Test	Jobs	Jobsteps
N day long test, with **M members** and named with **EXP-ID**	**1st** run of the period which is a **cold-start**, **1st** member	1. stget
		...
		15. stsav
	...	...
	1st run of the period which is a **cold-start**, **Mth** member	1. stget
		...
		15. stsav
	2nd run of the period with **long cut-off** analysis, **1st** member	1. stget

		...
		15. stsav
	...	...
	2nd run of the period with **long cut-off** analysis, **Mth** member	1. stget
		...
		15. stsav
	3rd run of the period with **long cut-off** analysis, **1st** member	1. stget
		...
		15. stsav
	...	...
	(N*8)th run of the period with **long cut-off** analysis, **Mth** member	1. stget
		...
		15. stsav
	1st run when production was needed, with **short cut-off** analysis, **1st** member	1. stget
		...
		15. stsav
	...	...
	Nth run when production was needed, with **short cut-off** analysis, **Mth** member	1. stget
		...
		15. stsav
...	...	...

Table 6.: More test can be run still independently. Jobs are not independent and they have to follow this order. Jobsteps also still has to follow each other automatically with different LoadLeveler settings.

Finally it is important to play attention to the question of coupling files used in lagged mode. There was only one PEARP run (at 18UTC) used during AROME-EDA tests. At most of the AROME forecasts it was obvious, that the closest PEARP forecast should be used for coupling. However 18UTC run was questionable, because the Hungarian experience is, that operationally PEARP becomes available only around 00UTC. After 00UTC if EDA is working, it does not make sense to start 18UTC run anymore, so 18UTC AROME-EPS should be also coupled to the PEARP from the previous day. However this operational aspect made comparisons a bit difficult, because reference AROME-EPS (simple downscaling of PEARP) uses a 24-hours fresher global model. This fact has to be kept in mind in part 6 as well.

There is also a 3-hour integration which uses long cut-off analysis at 18UTC and produces guess for 21UTC. From operational point of view, if there is time to wait for observations then there is time to wait for coupling files, so this integration runs with 0 hour lagging.

Time of the run and the used cut-off type	Lagged time	Length of the integration
00 UTC long cut-off	6	3
03 UTC long cut-off	9	3
06 UTC long cut-off	12	3
09 UTC long cut-off	15	3
12 UTC long cut-off	18	3
15 UTC long cut-off	21	3
18 UTC long cut-off	0	3
21 UTC long cut-off	3	3
00 UTC short cut-off	**6**	**36**
06 UTC short cut-off	12	36
12 UTC short cut-off	18	36
18 UTC short cut-off	**24**	**36**

Table 7.: The cut-off types of the different analysis, the lagged time and the length of the different possible forecasts. Two of the production times are highlighted, because their results are represented in part 6.

6. Verification results with AROME-EDA

For the test of AROME-EDA module a very similar period was chosen than for ECMWF-EPS-LBC test. The reason was that these test should stay cross-comparable and this winter-period fits to ECMWF's special project called 'spfrbout'. The ensemble of data assimilation suites started on 24th of December, 2011. There was just a relatively short spin-up time and the period of the forecasts was **26.12.2011-08.01.2012.**

In the figures which are showed in this part the following naming conception is vaid (also see the experiment ID conventions in part 2.):

PEARPcoup18 – Coupled to **PEARP** at **18UTC** (simple downscaling, used as a reference). Surface is downscaled from Hungarian ALADIN assimilation suite.

Hgfffb18 – Started from an **AROME-EDA IC at 18UTC** and coupled to a 24-hour old PEARP

Hgfffb00 – Started from an **AROME-EDA IC at 00UTC** and coupled to a 6-hour old PEARP

Near-surface comparison

First the near-surface scores are represented to deserve the positive atmosphere of this document. This verification results suggest that AROME-EPS based on EDA ICs, has a much better quality in the first 3-12 hours (depends on the variable). EDA can produce some jumpiness in quality in terms of problematic parameters (10meter wind speed and cloudiness).

10meter wind speed overestimation was referenced also in part 4. and guessed, that there are some problem with the interpolation of ARPEGE fields. However if ICs are not the simple downscaling of ARPEGE this problem simple can be solved (but still not absolutely understood) (see fig. 8-9.).

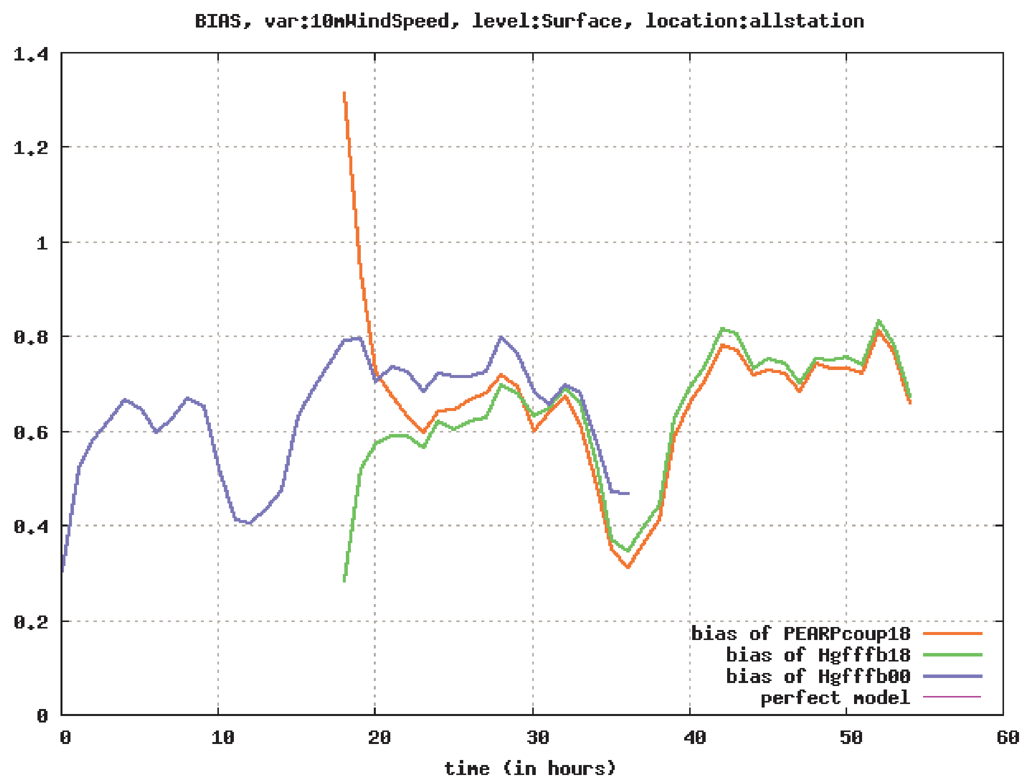

Fig. 8.: BIAS of 10meter wind speed. EDA ICs can also decreas the extreme overestimation at initial time.

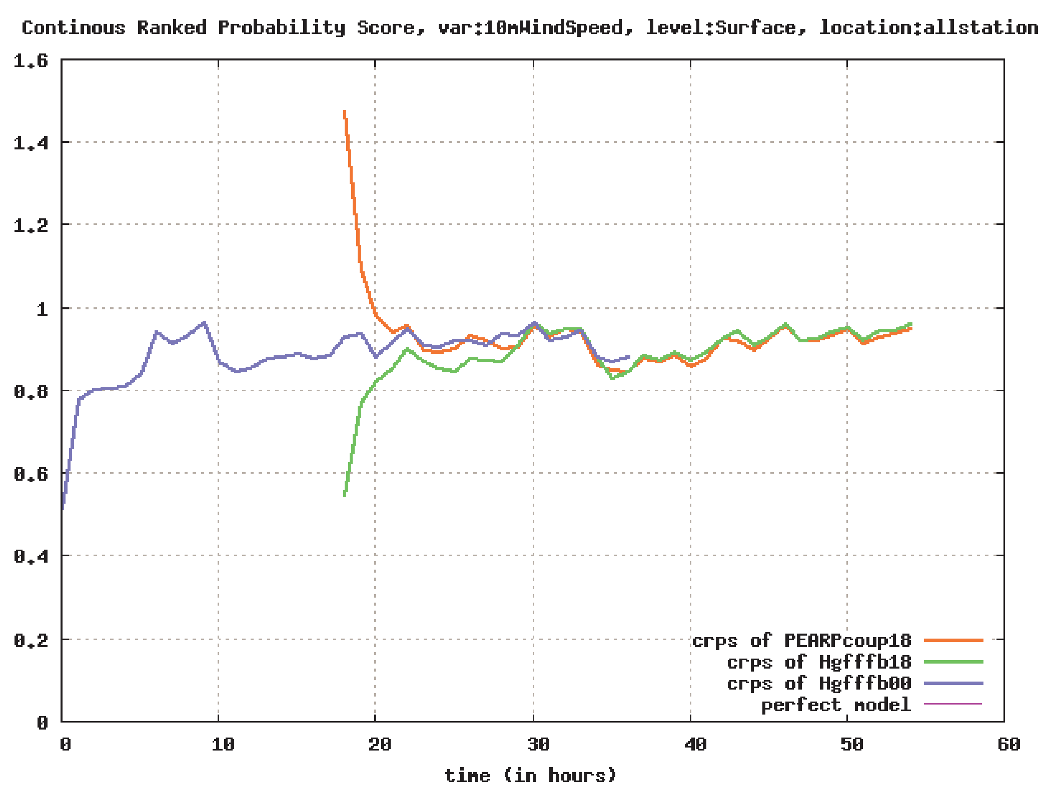

Fig. 9.: CRPS of 10meter wind speed.

Cloudiness showed big underestimation at initial time in case of simple downscaling. If an ensemble of data assimilations is running, than it is easy to initialize hydrometeors from fresh guesses and it has obviously a positive impact (see fig. 10-11.).

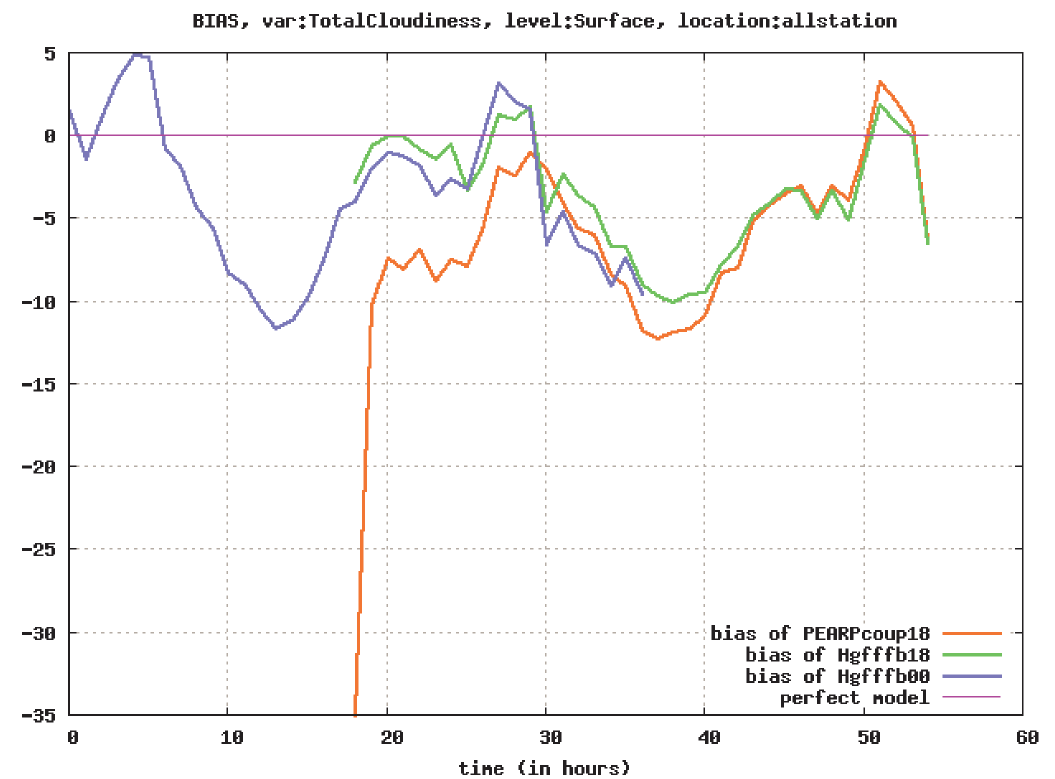

Fig. 10.: BIAS of total cloudiness. AROME-EDA made possible hydrometeor initialization which can improve cloudiness at the early hours of the forecast.

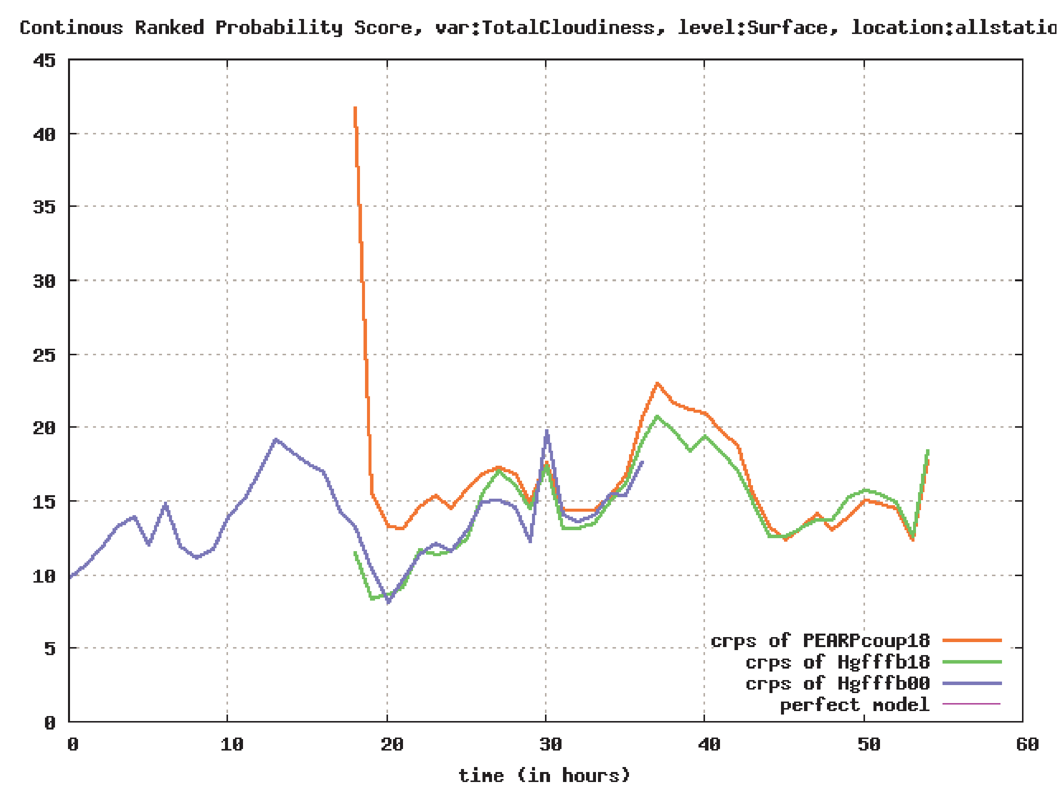

Fig. 11.: CRPS of total cloudiness.

After this two problematic variables Spread-skill relationship and CRPS of 2meter temperature and relative humidity are represented. While results look very promising it has to be noted that the additional SPREAD of EDA at the initial time decreases quite fast at the beginning, especially in case of RHU (see fig. 13.). It suggest that these perturbations near the surface can not behave like 'optimal' ones.

The classical probabilistic way of thinking with singular vectors needs small IC perturbations which evaluate fast with time. In our limited area model, near the surface, where the perturbations coming from an EDA the behavior is very different from the ideal and classical way. In my opinion it is not a problem and improvement in scores at the first hours is an important result. However it suggests that EDA (likely even with different settings) has its limitation, and anyone who wants to improve forecast skill up to 36 hours, has to think on other methods as well.

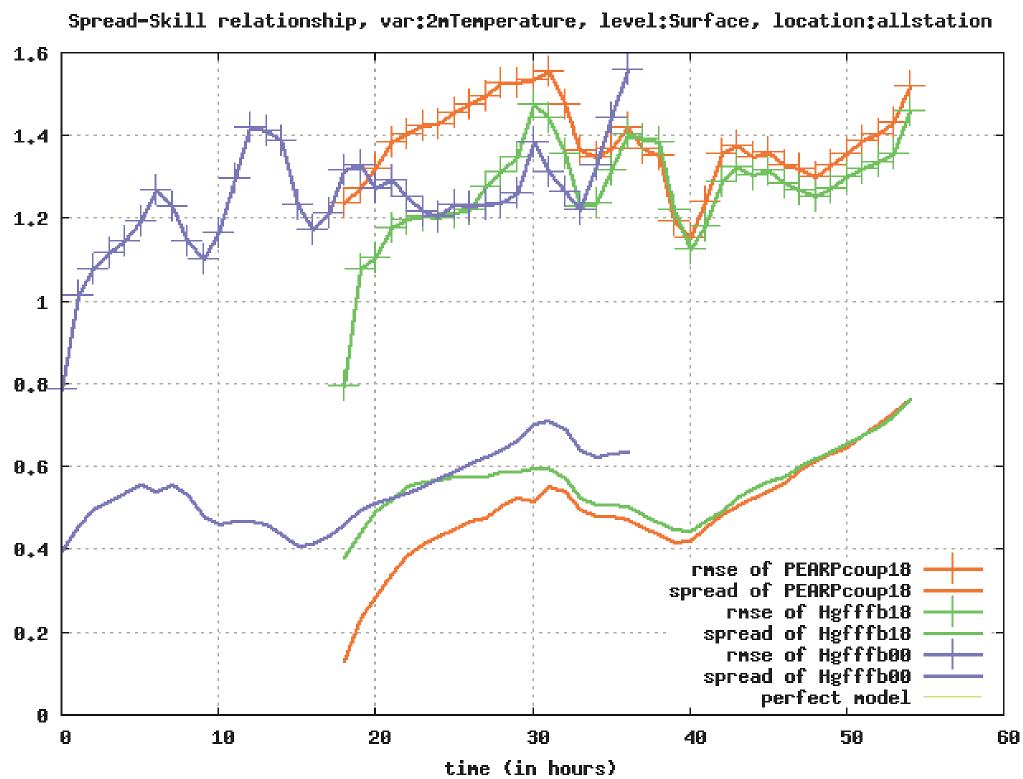

Fig. 12.: Spread-skill relationship of 2meter temperature

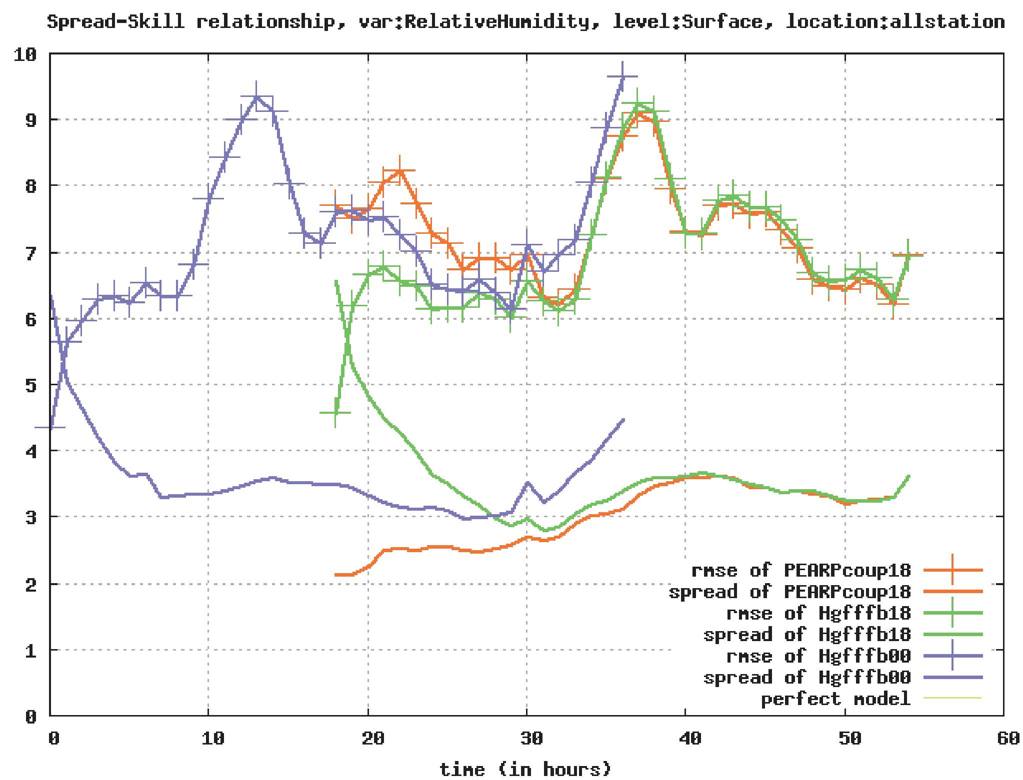

Fig. 13.: Spread-skill relationship of relative humidity

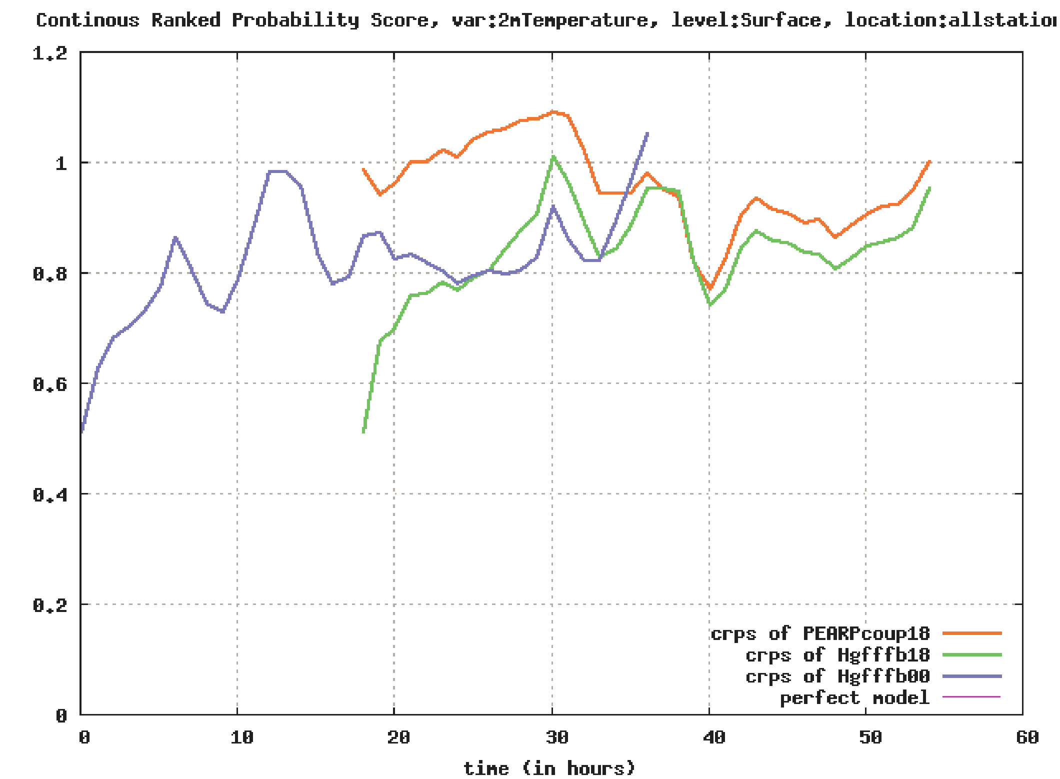

Fig. 14.: CRPS of 2meter temperature

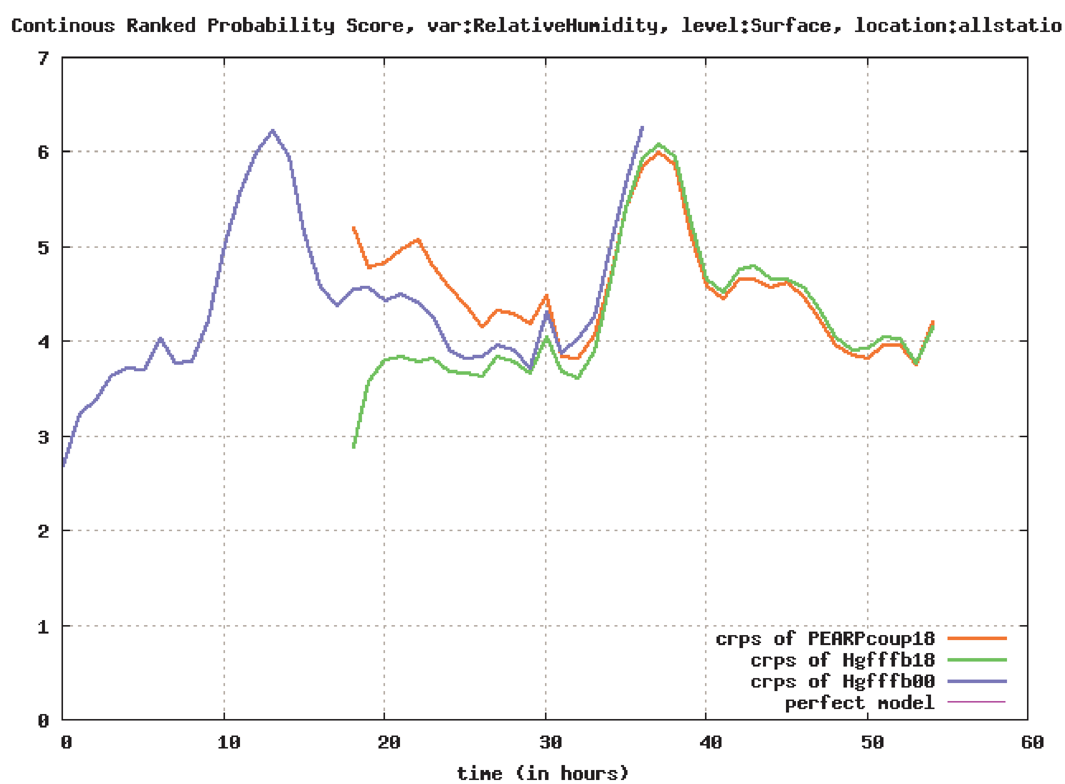

Fig. 15.: CRPS of relative humidity

High-atmospheric comparison

How it was written also in part 4., ECMWF analysis is not always thought as the best reference in a verification, especially if we are coupling to another global model. It is also guessed that with implementation of EDA there are more small scale information in ICs, which can look like noise in a lower resolution analysis. Because of the mentioned reasons radiosonde observations were chosen to verify against.

In that comparison results are not that good than in the previous one. SPREAD is usually increased, but EDA not always has an obviously positive impact on RMSE. It is important to keep in mind the question of lagged time, which was detailed in part 5. It is guessed that 18UTC forecasts from EDA experiment have the disadvantage with the older LBCs. In figures it looks, that 00UTC version has much better quality especially in higher atmosphere, where probably the impact of LBCs is bigger.

To confirm that 18UTC run quality problem is coming from the long lagged time, it is planned to run a verification on guess files with 3-hour frequency. While at generation of these files from 18UTC to 18UTC always the same global run is used, a daily trend is expected.

Fig. 16-19 represent Spread-skill relationship of some widely used variables.

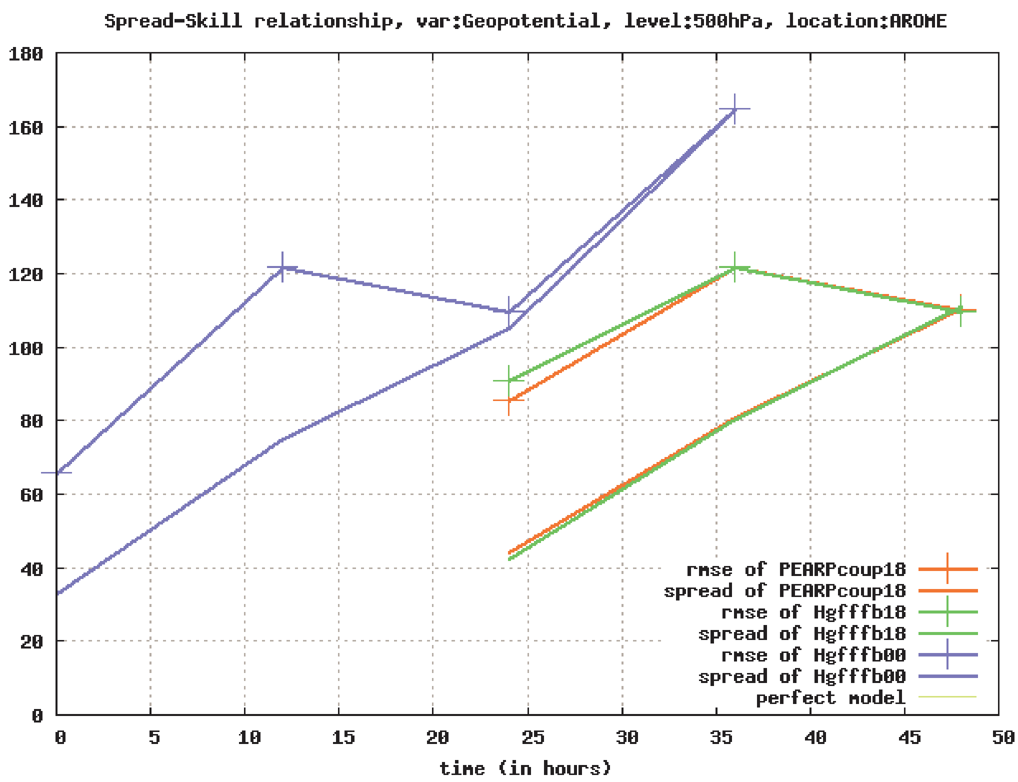

Fig. 16.: Spread-skill relationship of geopotential at 500 hPa

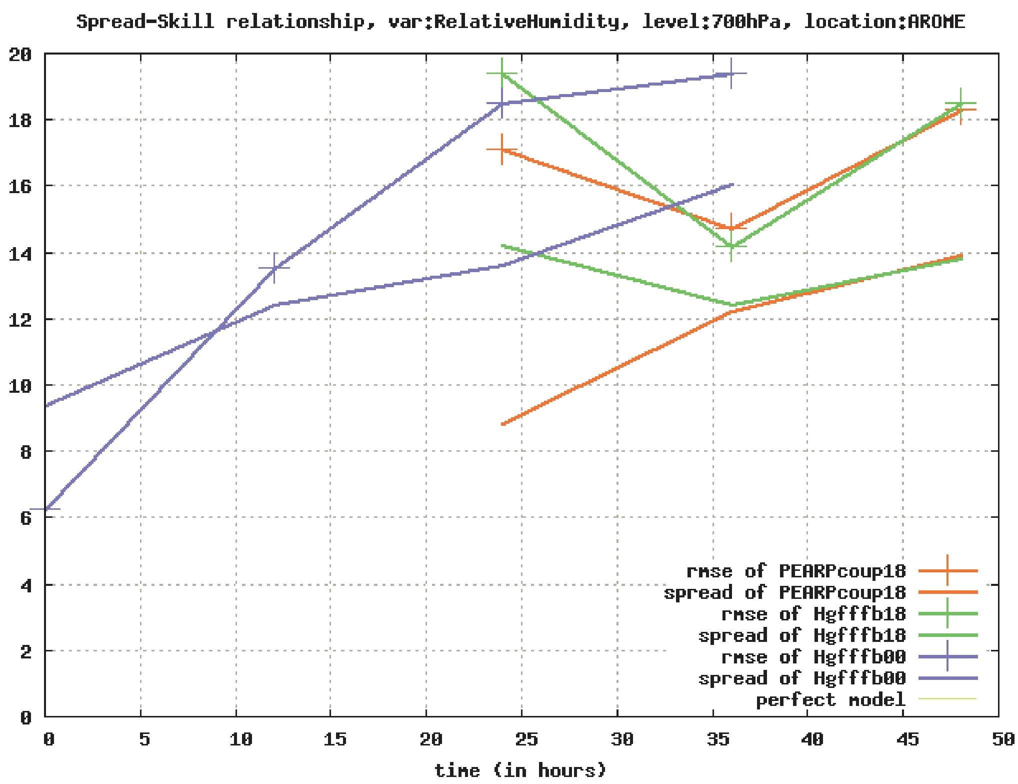

Fig. 17.: Spread-skill relationship of relative humidity at 700 hPa

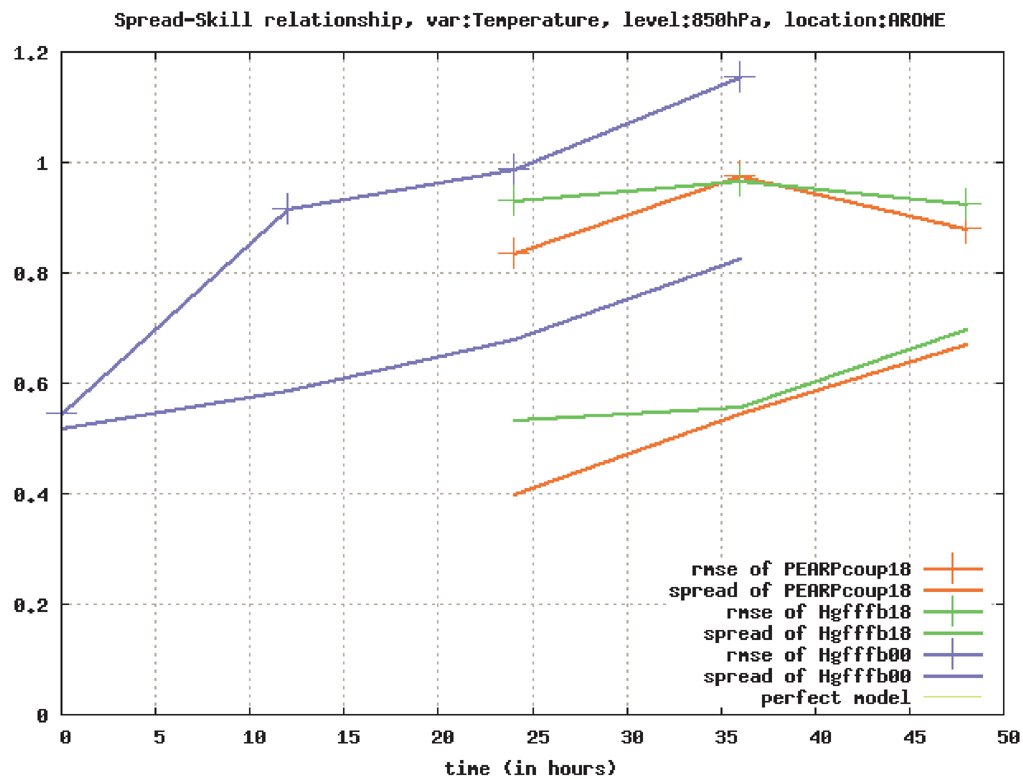

Fig. 18.: Spread-skill relationship of temperature at 850 hPa

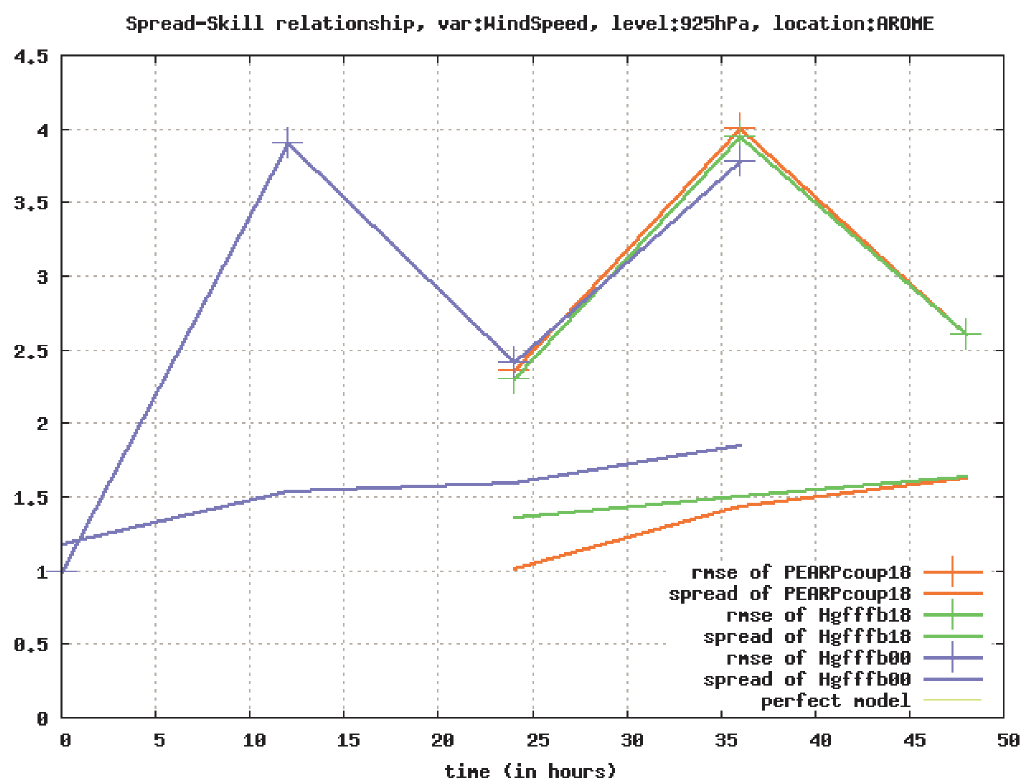

Fig. 19.: Spread-skill relationship of wind speed at 925 hPa

For more verification results, see ECMWF's ecgb:
/home/ms/hu/hu7/AROME-EPS-EDAtest1 (different subdirectories with nr.3)

7. Conclusion and motivation for further work

Part 4. represented results from ECMWF-LBC tests:

- The positive impact of the increasing resolution global model was small but significant.
- It is variable dependent, if doubling the resolution of the global model or doubling the ensemble size has a bigger impact. Probably there are more cons on the resolution increasing size.
- The impact of ECMWF and ARPEGE LBCs are hardly comparable. Probably it looks that there are a bit more problems with PEARP.

The results around the first two conclusion can help to find an ideal way of ECMWFs EPS LBC generation and usage method in the future.
The third conclusion can be also a motivation for AROME-EDA work, which can be based on PEARP or ECMWF-EPS too. With AROME-EDA some dirty interpolation noise can be eliminated at initial time and the method gives bigger flexibility to run forecasts from the same time, which makes tests more comparable.

Parts 6. represented results from AROME-EDA tests. While the impact of the method looked very promising, it can not be interpreted as a final conclusion a work, but a motivation for even more work. There are ideas about what-to-do on different time scales.
Short-term plans:

- How it was mentioned in part 5. and 6., forecasts lagging time can play an important role. It is planned to verify the guesses, to see somehow this effect.
- AROME-EPS-EDA script system still contains a bug which does not effect the results, but computer efficiency: ARPEGE is rerun too many times if PEARP is chosen as a global model.
- The visualization of forecasts and some simple case studies can help to understand better the importance of the described developments.

Medium-term plans:

- At the recent configuration every AROME-EPS member had a pair from the EDA system. It is planned to test, how can be used a 5+1 member EDA, where perturbations are centralized and this way their number is doubled. This is a method which is also used at ECMWF and Météo France.
- Earlier SPPT on this scale in 24-hour forecast could not give too much effect on ensemble quality. However it can be tested again how much improvement it gives, if it is active also in EDA and forecasts are longer (36 hours)
- Scaling of observation error is a sensitive question. Earlier in 8km tests at OMSZ we found that drastic increasing of rescaling factor can be reasonable. After the first result at 2.5km scale it does not look crucial to increase initial SPREAD more, but we should keep in mind this question.
- The represented results were evaluated from the EPS verification system of OMSZ. This is a system with many limitation (limited number of scores, no precipitation specific methods) and there is no too much manpower to develop it. It would be nice to start with common LACE EPS verification package even on convection-permitting scale.

Long-term plans:

- It was mentioned in part 6., that EDA looks a useful tool, but it has its limitation in time. On global scale it is usually used in combination with other IC perturbation method (mainly singular vectors) and model error representation (SPPT, multi-physics). Of course limited area short-term EPS gives different challenges, and probably the above-mentioned methods are not really fitting to these challenges, but there is a need to find way , how perturbations from EDA and LBCs can be further improved.

8. Acknowledgment

There are a lot of people who I am appropriate in connection with the above-presented work. First of all I would like to say thank you for modeling group of ZAMG where I found friendly atmosphere and helpful people; special thanks to Theresa Gorgas who hosted me and to Florian Meier for data assimilation related consultations and to Christoph Wittmann for helping me with technical problems.
There are also people at OMSZ who gave a lot of support, not only during this 6-weeks, especially in area of data assimilation: Mate Mile, Gergely Boloni and Roger Randriamampianina.
As usually at every AROME-EPS related work I am appropriate to Francois Bouttier whose system was a valuable kick-off for Hungarian researches and who supervises ECMWF's special project, called 'spfrbout'.
Finally I am glad for the HARMONIE and GL related help from HIRLAM's side to Andrae Ulf and Inger-Lise Frogner.