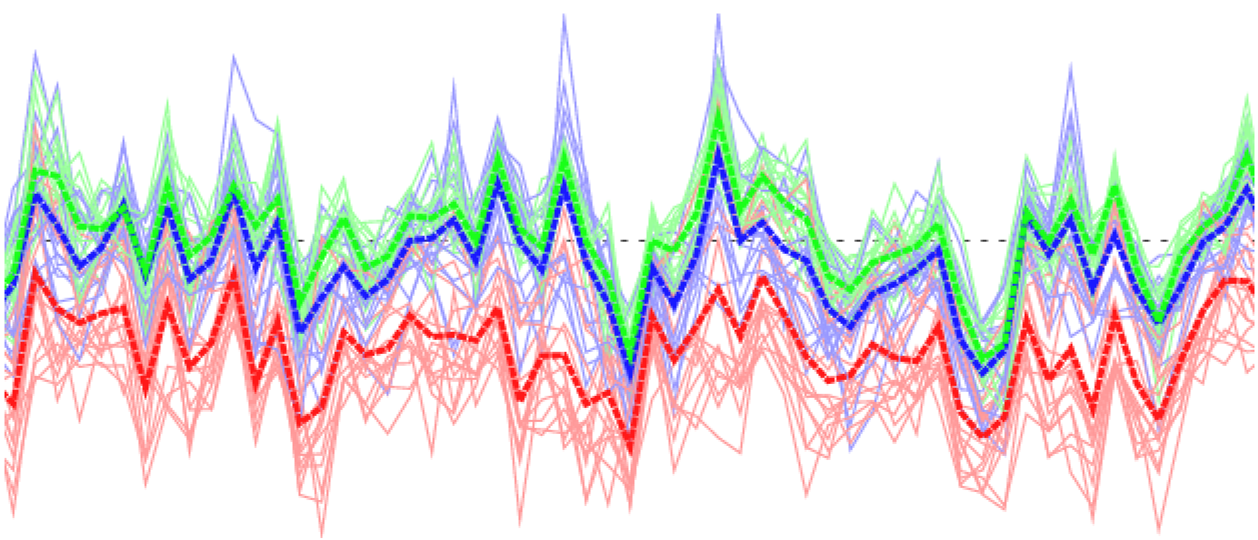
Report on stay at ZAMG

13/05 – 21/06/2013, Vienna, Austria

Time Consistent versus Space Consistent coupling and the revision of the Ensemble of surface Data Assimilations by CANARI in ALADIN-LAEF



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

::Author
Martin Belluš (SHMU)
martin.bellus@shmu.sk

::Table of Contents

Acknowledgement
Foreword

I. Externalization of Perl functions and the other code upgrades
II. Time Consistent versus Space Consistent coupling
III. Surface assimilation of RH2M
IV. Verification of new ALADIN-LAEF

Conclusions
References

::Acknowledgement

I would like to express my acknowledgement to the local ZAMG team for their hospitality and friendly atmosphere. I am grateful to Francois Bouyssel for his valuable hints regarding relative humidity assimilation problem by CANARI. My compliment belongs to Simona Tascu, for her courage to rewrite and improve the LAEF Verification package. I'd like to thank to Maria Derkova for inspiring talks in the past, present and future. And many thanks to my family for their patience...

::Foreword

Technical maintenance of existing R&D version of ALADIN-LAEF code towards the operational implementation and several upgrades of the code functionality were carried out during this stay. First of all, the externalization of shared Perl functions was done for the main LAEF "bricks", such as for breeding, surface assimilation and blending scripts (see more details in chapter I). For easy experiment handling, the code for so called "time consistent" and "space consistent" coupling was merged (see more in chapter II, together with the experiment outputs and new rising ideas). Revision of Relative Humidity surface assimilation cured the problem with null assimilation increments for surface liquid water content (see more in chapter III). Full ALADIN-LAEF system for 16 members including new development - mainly the ensemble of surface assimilations by CANARI based on the perturbed observations (with the corrected RH2M assimilation) and multiphysics (new tuning done by Christoph Wittmann) was integrated on new domain (dx=10.9km) for 3 months in 2011. The verification scores for this experiment are confronted with the "old" LAEF system and pure ECMWF downscaling in chapter IV.

Furthermore, in cooperation with Simona Tascu (who was working on LAEF Verification package optimization at ZAMG), the driving script for the surface LAEF Verification was rewritten from Shell to Perl. More security checks are now included and evaluated before the verification starts. At the same time less activity is needed from the user, since big part of the initial setup is now performed automatically by Perl script. Necessity of filling the duplicated informations by user was avoided and the verification settings are now separated from the execution code. New plotting scripts written in Perl (using gnuplot) were also prepared for the basic visualization of the new verification package results. New functionality was added - the option to plot statistical scores by days and especially to plot the scores for the individual ensemble members as well. We believe, that this kind of visualisation can help us better understand the behaviour and quality of the ensemble system.

::I. Externalization of Perl functions and the other code upgrades

Because of already big scope of the existing applications for breeding, surface assimilation and blending, it was inevitable to externalize and unify the different functions largely used within them. New common Perl module (*Support.pm*) with all the functions was created. It is much better for the maintenance, since the potential changes need to be applied only on one place in the future. At the same time, the application source code for the LAEF "bricks" has been simplified by approx 200 lines each. That was also a good opportunity to check and revise the complexity of the code. New *Support.pm* module contains the following subroutines:

&modif_namel($file, %hash)
inputs: namelist, hash with tags and their substitutes
output: modified namelist

Supported tags were extended to both “__X__” and “{X}” format, i.e. one can write the template for namelist like this:

```
&NAMBLEND
  CL_FNAME1='__FNAME1__',
  CL_FNAME2='__FNAME2__',
  CL_FNAME3='__FNAME3__',
  L_SPEC_HYDRO=.T.,
  Z_NSIGN=__NSIGN__,
/
```

or

```
&NAMBLEND
  CL_FNAME1='{FNAME1}',
  CL_FNAME2='{FNAME2}',
  CL_FNAME3='{FNAME3}',
  L_SPEC_HYDRO=.T.,
  Z_NSIGN={NSIGN},
/
```

To keep and maintain just the templates of namelists is very handy, especially for ensemble system and the system where several procedures take place and differ only according to given setup via namelist.

&check_date($dd, $mm, $yyyy)
inputs: day, month, year
output: return code 0 or exit (procedure is stopped if date is not valid)

&leap_year($yyyy)
inputs: year
output: 0 (not a leap year) or 1 (leap year)

This function is used by *&check_date* to determine the number of days in February.

&shift_date($dd, $mm, $yyyy, $HH, $N)
inputs: day, month, year, hour, shift (in hours)
output: dd, mm, yyyy, HH shifted by N hours (it works in both directions)

&spent($start)
inputs: starting time in seconds (for actual time it can be e.g. “$start = time();”)
output: HH:MM:SS (passed time from the start is printed to STDOUT)

&gettrunc($file)
inputs: any FA file
output: NSMAX, NMSMAX (spectral truncation of the data in FA file)

&wait_ff($file, $expiration, $action)
inputs: any accessible file, expiration time in seconds, desired action
output: return code 0, 1 or exit (die)

It checks the file existence. If file doesn't exist after expiration has run out, it either exits (if defined action is "exit") or continues with return code 0. If file exists, it continues immediately with return code 1. This function can be used in operational implementation to handle the observation files or any other input files.

Here is the list of all ALADIN-LAEF "bricks" and the subroutines they are using from *Support.pm* module:

blend (upper air blending)	*&gettrunc, &modif_namel, &spent, &check_date*
breed (upper air and p_s breeding)	*&modif_namel, &spent, &check_date, &shift_date*
canari (surface assimilation of perturbed observations)	*&gettrunc, &modif_namel, &spent, &check_date, &shift_date, &wait_ff*
laeff (LAEF forecast integration)	*&modif_namel, &spent, &shift_date, &check_date*

Furthermore, the setting of boundary conditions (LBC files) was moved to the main configuration file *Conf_app.pm*. The reason was again the higher portability and easier experiment handling. Three functions were introduced to substitute the date, ensemble member and forecast range in the LBC filenames and paths:

&LBC_PATH($yyyy, $mm, $dd, $HH)
&LBC_ARCH($member)
&LBC_FILE($member, $range)

The above changes helped much for an easy implementation of lagged coupling scheme as well. It can be activated simply by exporting the ENV variable CNF_LAGGED. Since the used LBC files are now fully controlled by calling &LBC_PATH, &LBC_ARCH and &LBC_FILE subroutines with the given date arguments, it is easy to shift the time internally in corresponding functions.

The external tool for getting the boundary conditions *get_lbc.pl* was also modified. It is possible to fetch the files from ECFS as well as from local FS (which is determined automatically by the defined file path). It is capable to work with file archives and also with the individual files (which is determined automatically by the file extension).

Next to the other existing ENV switches like CNF_REPRO (to reproduce the exact surface perturbations between several experiments) new switch to omit the OBS perturbation at all was also implemented. By exporting CNF_NOPERT only the surface assimilation will be done (without previously perturbed observations in ODB). This is meant for the pure assimilation procedure in the control run, obviously.

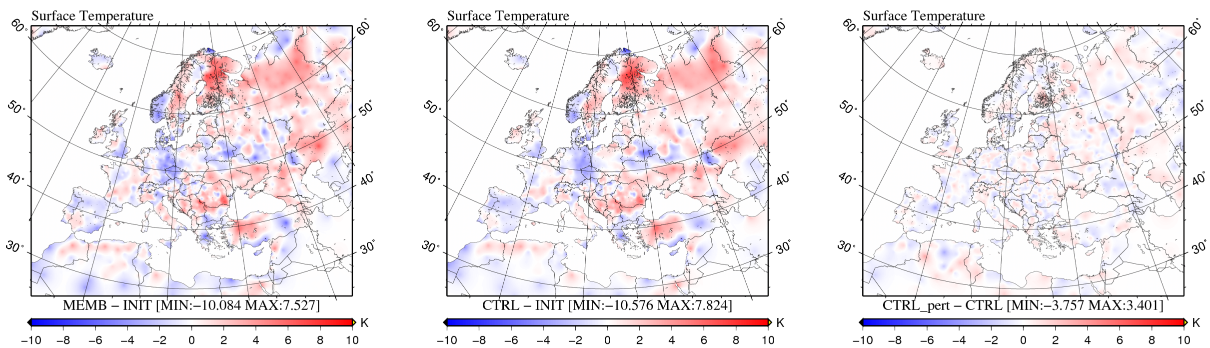


Fig 1: *Surface Temperature assimilation increments based on perturbed observations for selected member (left), the same increments for control run but on unperturbed observations (middle) and the difference between perturbed and unperturbed assimilation of control run, were one can see exactly the contribution coming from OBS perturbation (right).*

::II. Time Consistent versus Space Consistent coupling

For an easy experiment handling regarding the coupling method approach, the code for so called “time consistent” and “space consistent” coupling was merged under the switch to be found in the main configuration file *Conf_app.pm*. According the *$coupl* variable (TCC or SCC) the appropriate coupling method is used and also output sub-directory with such name is created to store the results (R&D version of ALADIN-LAEF).

From the case study performed for 15th of May 2011 the following summary can be concluded:

- There are no differences between two investigated coupling methods at range 00 (i.e. for "model analysis"). It only means that the 1st coupling file is taken into account just from the 1st time-step.

- Then the "signal" (the difference between TCC and SCC forecasts) is clearly present since the beginning of the integration and it comes solely from the coupling zone as expected (see Fig 2, 3, 4 - left).

- The "signal" is further advected towards the domain center, but not homogeneously (see next point).

- The differences in Temperature and other fields at 2m and also on the vertical model levels are later materialized along the frontal zones (see Fig 2, 3, 4, 6 - right and Fig 5).

- According to the trace of cumulative Stratiform Precipitation differences, there is obviously a slight spatial shift of the frontal zones positions (between the two different coupling methods), which is visible also after several hours of integration (see Fig 5 - right).

- The differences in Convective Precipitation fields between TCC and SCC runs are also organized in obvious relation to the frontal zones. However, their spatial distribution is locally more random, probably in accordance with the behaviour of convective phenomena (see Fig 5 - left).

- The "signal" is present on the vertical model levels as well (see Fig 3).

- The impact on statistical scores computed for the ensemble mean (all 16 LAEF members taken into account) over the whole domain and for longer time period (one month) is negligible (see Fig 8). However, for individual cases the selected coupling strategy can have significant impact mainly within the synoptically active areas, where the differences between the two approaches can locally reach 2-4 degrees for Temperature, 10-30% for Relative Humidity and several millimeters for Precipitation fields depending on the strength of the frontal zones.

These experiences revealed a new idea for utilization of the coupling strategy in LAM EPS:

It seems, that such phenomena could be possibly used to create a targeted perturbation for LAM EPS. It would act locally within the areas where it matters the most, i.e. along the frontal zones or synoptically active locations. Its impact on the global verification scores is neutral and unlike the other perturbation methods (ensemble of data assimilations for surface, upper-air breeding) it acts rather locally, not homogeneously over the whole domain and can apparently affect even the long forecast ranges.

On the following maps we show the divergence of the ALADIN-LAEF forecasts due to two coupling methods. One has to realize, that the only difference between the experiments is in the first coupling file. Even the OBS perturbations are kept the same, using the seed numbers reproducibility (via environmental variable CNF_REPRO). For better understanding of the “signal” behaviour, we did the 54-hours animation with hourly step for each investigated field. From technical reasons, only the maps for first range and 48 hour forecast are shown. At the end, the synoptical analysis valid for that 2nd forecast day should demonstrate the relation between the “signal” and actual frontal zones.

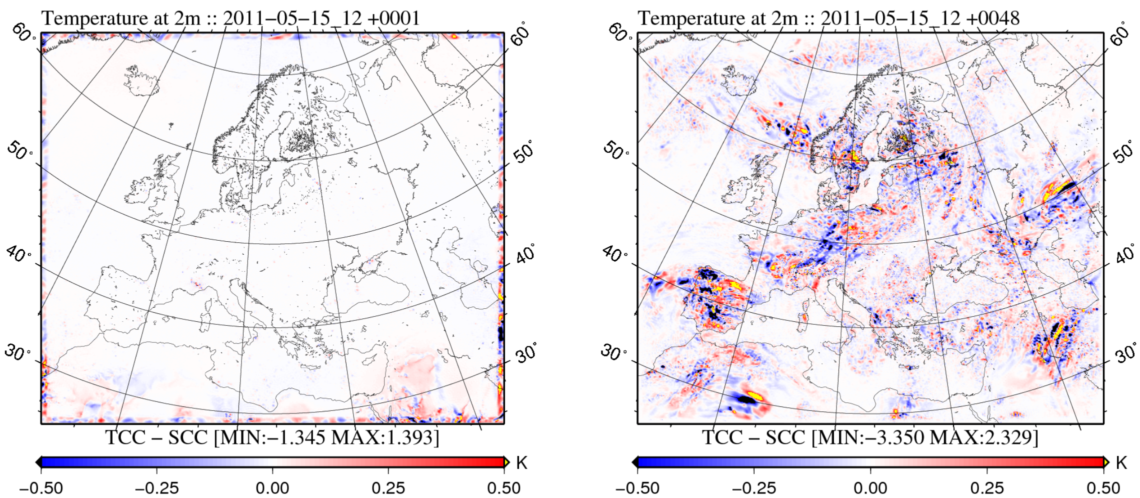


Fig 2: *The difference between TCC and SCC experiments for Temperature at 2m after 1st hour of integration (left) and after 48 hours, i.e. valid for 17th of May 2011, 12 UTC (right).*

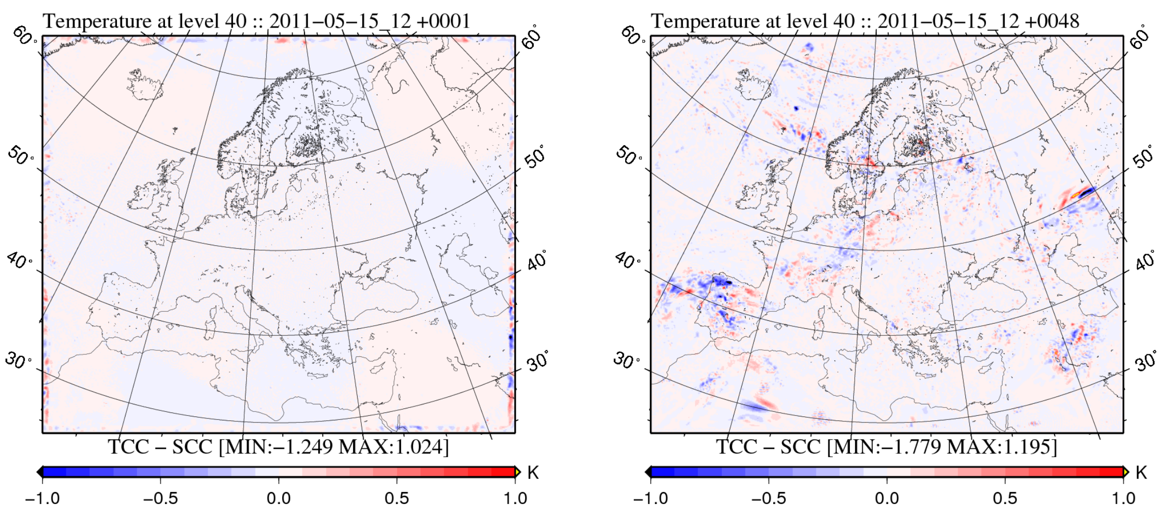


Fig 3: *The difference between TCC and SCC experiments for Temperature at level 40 after 1st hour of integration (left) and after 48 hours, i.e. valid for 17th of May 2011, 12 UTC (right).*

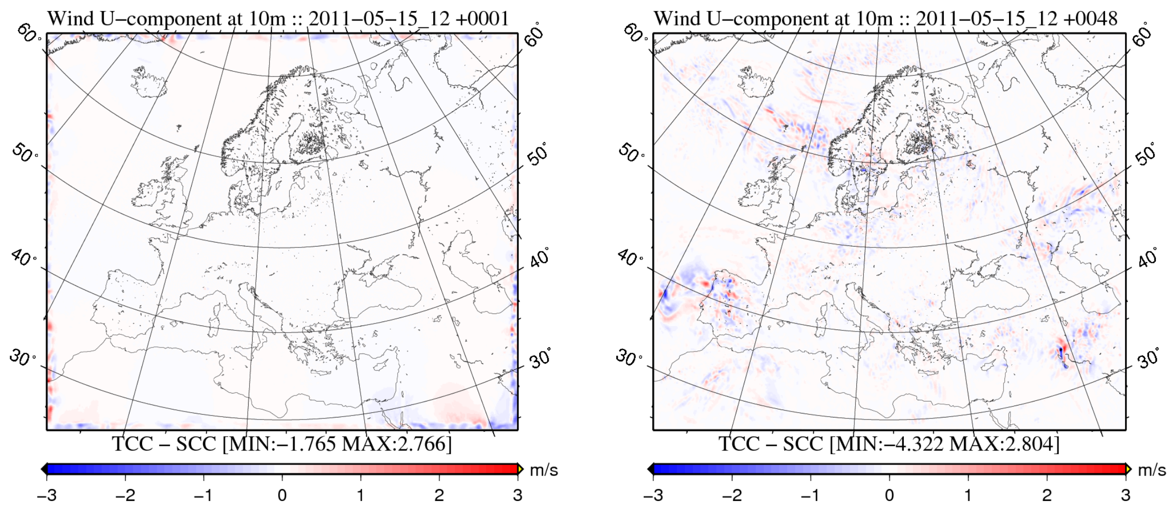


Fig 4: *The difference between TCC and SCC experiments for U-component of the Wind at 10m after 1st hour of integration (left) and after 48 hours, i.e. valid for 17th of May 2011, 12 UTC (right).*

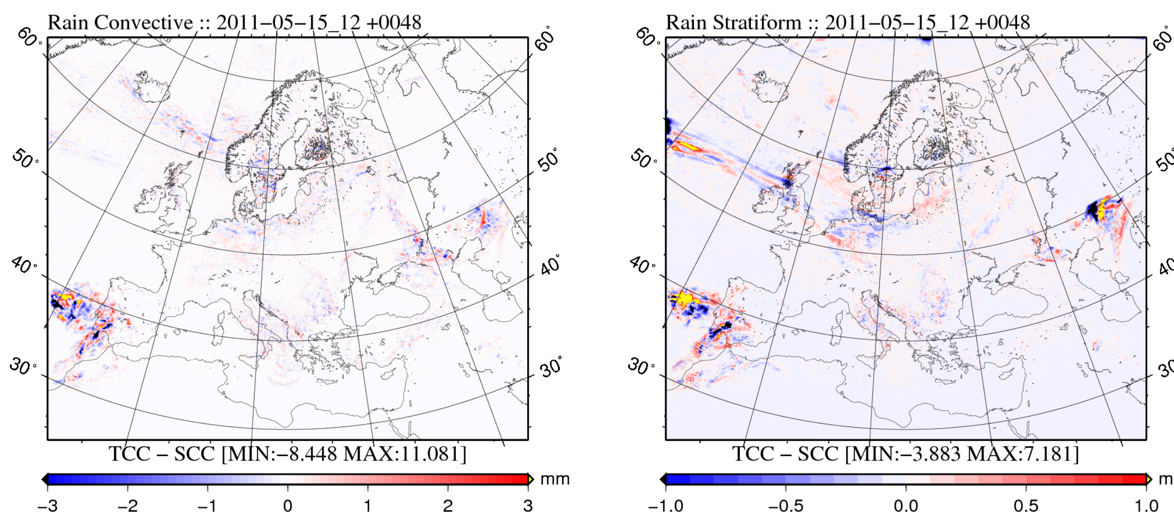


Fig 5: *The difference between TCC and SCC experiments for accumulated Convective Precipitation (left) and Stratiform Precipitation (right) valid for 17th of May 2011, 12 UTC.*

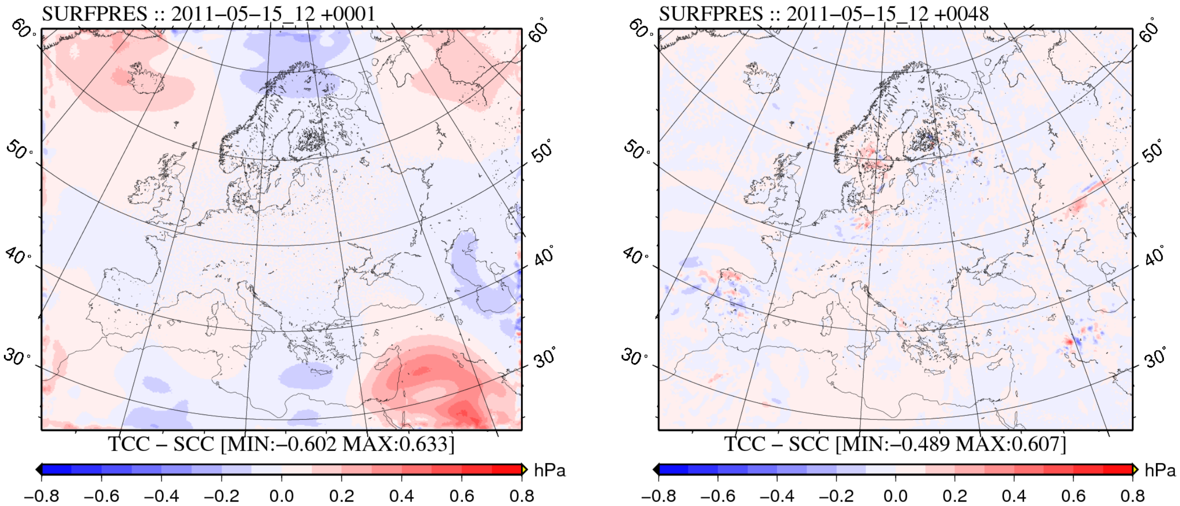


Fig 6: *The difference between TCC and SCC experiments for Surface Pressure after 1st hour of integration (left) and after 48 hours, i.e. valid for 17th of May 2011, 12 UTC (right).*

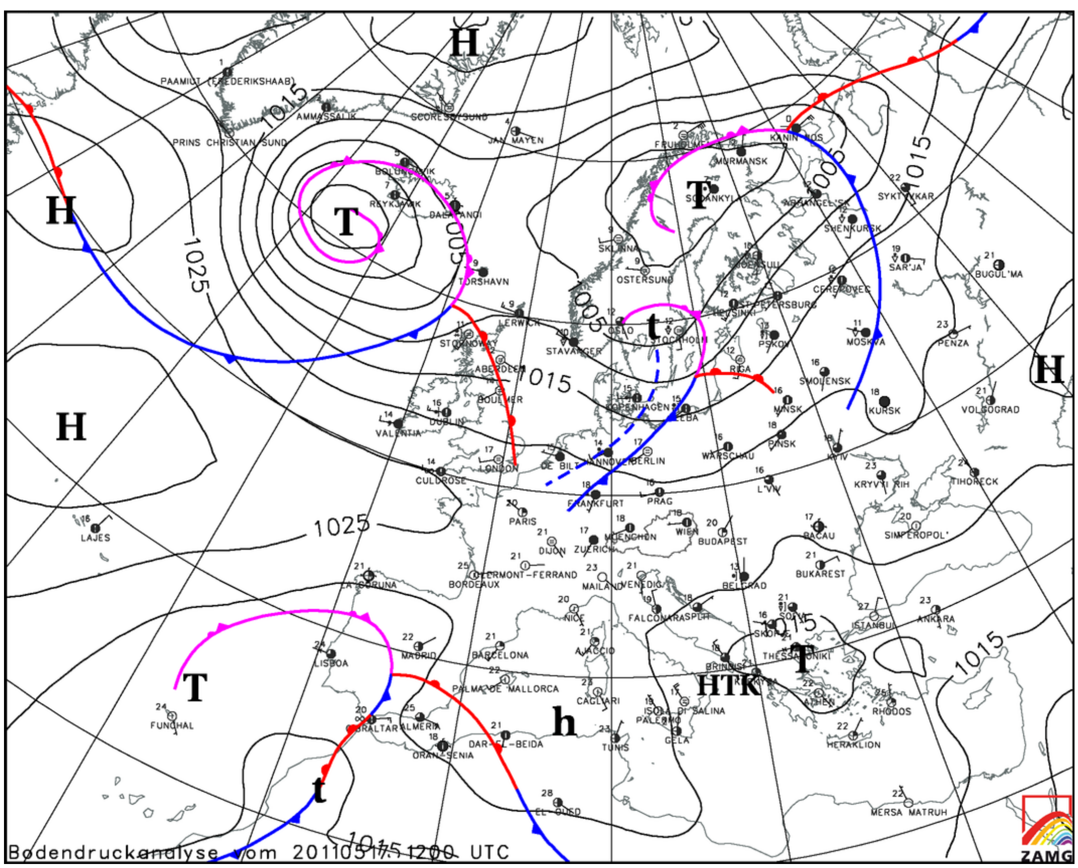


Fig 7: *Weather map analysis valid for 17th of May 2011, 12 UTC.*

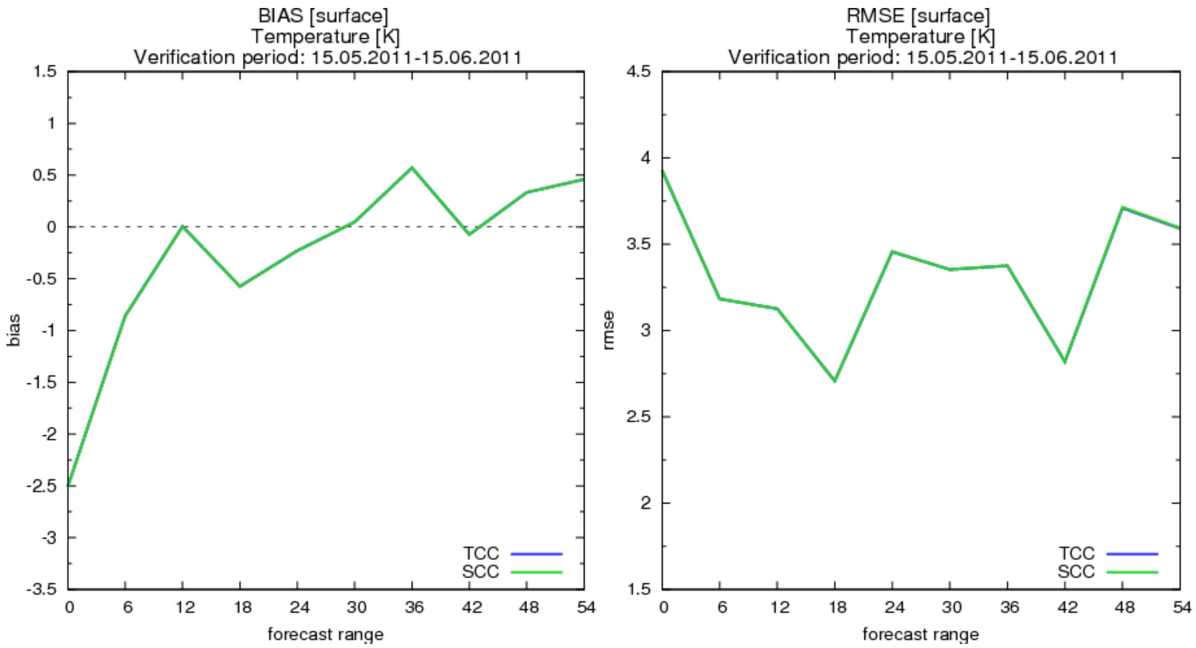


Fig 8: *Temperature BIAS (left) and RMSE (right) of ensemble mean for one month of time consistent coupling (TCC) vs space consistent coupling (SCC) experiments. The scores are not bit-identical, but the differences are beyond the resolution of given charts.*

For the Surface Pressure field (Fig 6) we can observe also a different kind of the initial disturbance. Here, for the first several time-steps, the oscillations related to model spinup can be observed. These initial state imbalances can be caused by the assimilation increments due to the perturbed observations and/or due to the lateral boundary condition inconsistency. They move fast over the whole domain and have large scale structure. After the first 3-4 hours of integration they completely disappeared, i.e. model fields came into the balance state as a result of dynamical adjustment. Then, for the next forecast ranges, our well known “signal” can be observed in Surface Pressure field as well. To the contrary, it is rather small scale and stationary, fixed to the active areas (i.e. cyclogenesis).

::III. Surface assimilation of RH2M

Shortly after the primary implementation of surface assimilation by CANARI in ALADIN-LAEF system (back in the end of 2012) it was observed, that the assimilation increments for the surface liquid water content are null all together. Although the impact seemed to be quite small, the assimilation of Relative Humidity was indeed not correct. It appeared, that the problem is due to some missing surface fields in the input file - the guess. It happened because of the prior breeding perturbation, where all surface fields were ignored by the upper-air perturbation tools and hence removed from the guess (better say - not saved into the output file). The question was, which surface fields are really necessary for a proper Relative Humidity assimilation.

Certainly, all the surface prognostic fields are required for the surface assimilation and they were already present. They were copied separately after the breeding procedure from the original 12h LAEF guess. Nevertheless, the superficial assimilation moisture increments were null. There are the following surface prognostic fields in ARPEGE/ALADIN files:

PROFTEMPERATURE	- deep soil temperature
SURFTEMPERATURE	- surface temperature
PROFRESERV.EAU	- deep soil liquid water content
SURFRESERV.EAU	- surface liquid water content
PROFRESERV.GLACE	- deep soil ice content
SURFRESERV.GLACE	- surface ice content
SURFRESERV.INTER	- interception water content (water on the leaves)
SURFRESERV.NEIGE	- surface snow depth water equivalent

Surface moisture assimilation increments are given by the differences between the analysed and predicted T_{2m} and RH_{2m} values following the equation:

$$\Delta w_s = \alpha_s^T \Delta T_{2m} + \alpha_s^{RH} \Delta RH_{2m}$$

However, the dependency on other meteorological fields is there through the coefficients. According to e-mail communication with Meteo-France expert Francois Bouyssel, the optimum coefficients for soil moisture analysis are modulated or switched off depending on several meteorological fields like precipitation, cloudiness, surface evaporation, and so on. If one of these is missing, the analysis increments can be set to zero. Hence, after the code inspection, we realized that our problem is related to the missing fields from XTRP2 group (PSP_X2 in *cacsts.F90*). It means, another seven fields were required (which were otherwise initialized by zero values if not present):

SURFPREC.EAU.CON	- convective precipitation
SURFPREC.EAU.GEC	- stratiform precipitation
SURFPREC.NEI.CON	- convective snow fall
SURFPEC.NEI.GEC	- stratiform snow fall
ATMONEBUL.BASSE	- cumulated low cloud cover
SURFXFLU.MEVAP.E	- instantaneous evaporation flux
SURFXEVAPOTRANSP	- instantaneous evapotranspiration

From the technical point of view, all these model fields were copied into the input file for the CANARI surface assimilation from LAEF 12h guess using *blendsur* aladin tool. Afterwards, the assimilation increments for superficial liquid moisture content were finally computed as one can see on the following plots (see Fig 9).

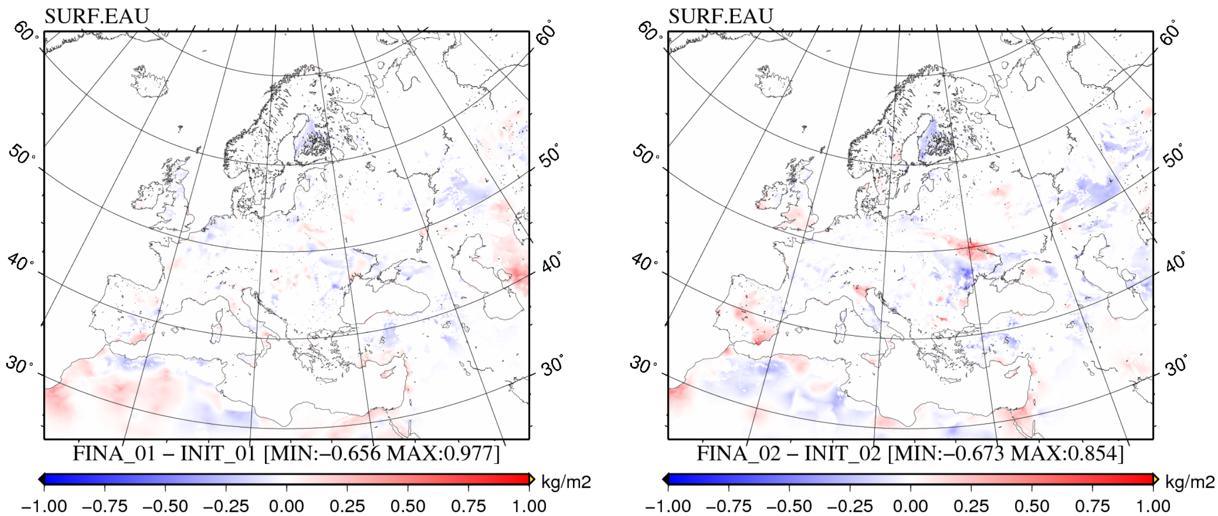


Fig 9: *Surface liquid water content assimilation increments after revised Relative Humidity assimilation for ensemble member 01 (left) and 02 (right). The differences between the members are due to perturbed observations.*

::IV. Verification of new ALADIN-LAEF

New ALADIN-LAEF system contains the upper-air initial perturbation of model variables (*t, u, v, q*) and p_s perturbation by breeding method. Furthermore, it contains the ensemble of surface data assimilations by CANARI based on randomly perturbed T_{2m} and RH_{2m} observations. Upper-air spectral blending by digital filter is used to combine the small-scale (LAEF) and large-scale (ECMWF) initial perturbations, where a non-physical signal out of the interpolation of global ensemble into the finer grid is replaced by more reliable and physically meaningful signal from LAM guess. At the same time it ensures the consistency between the initial and boundary conditions. Finally, the different physical parameterizations and related settings for the integration of individual LAEF members are used to simulate the model uncertainty.

To test our new ALADIN-LAEF configuration, the experiment was carried out for the historical period of 3 months (15th of May till 15th of August 2011). New ALADIN-LAEF assimilation cycle was initialized at 10th of May. However, the verification was done only for last 2 months of the experiment, because the first 20-30 days are "spoiled" by warming up of the assimilation cycle. That is the standard behaviour, especially for humidity assimilation, since deep soil moisture processes take long time to get into the balance. The comparison of new system with the current one and with the pure downscaling of ECMWF EPS forecast was done and the results are going to be presented on the following pages (see Fig 11-18).

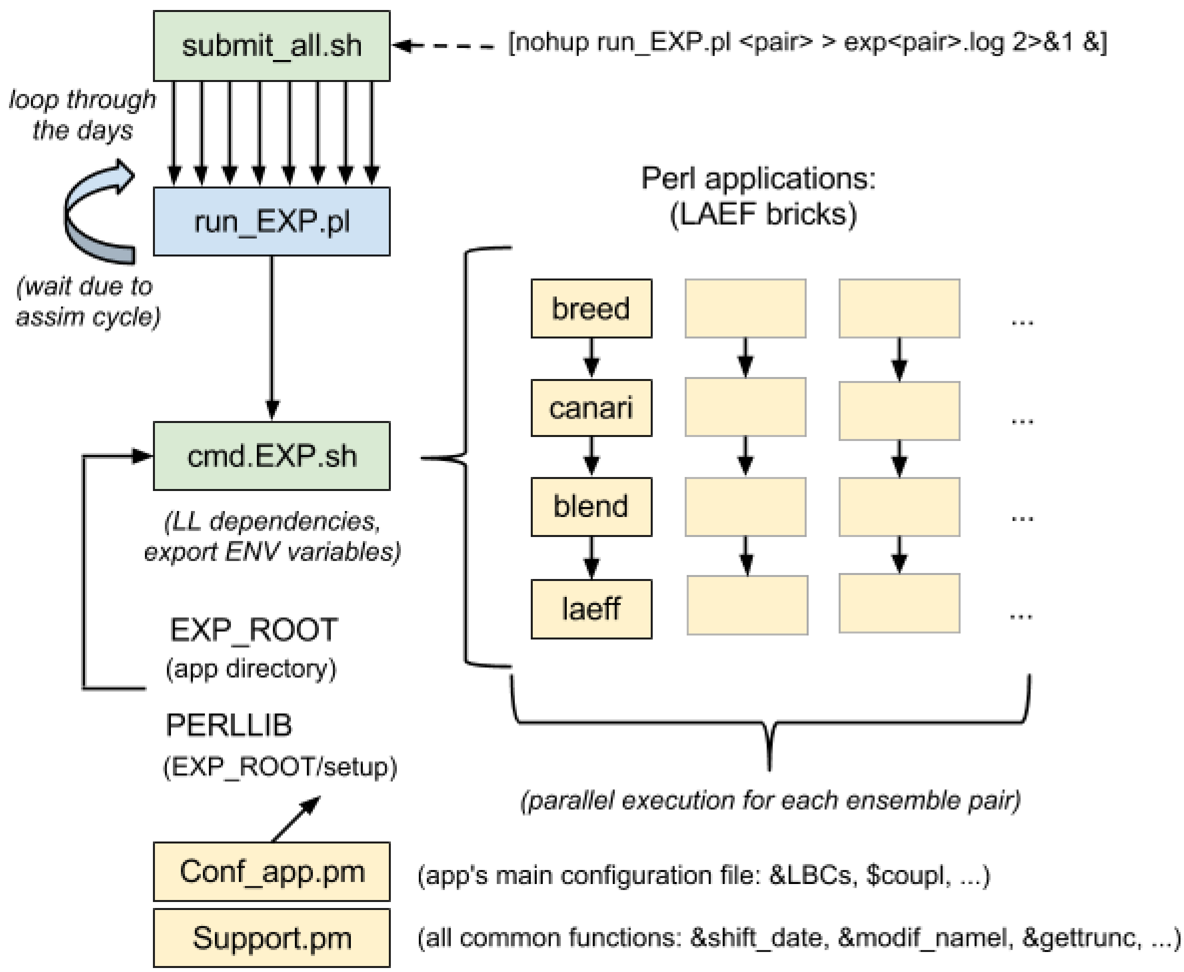


Fig 10: *New ALADIN-LAEF development scheme used for the experiments (such canari and blend "bricks" are used in the operational ALADIN-LAEF suite as well).*

The verification results for 2011 data-set can be summarized in the following points:

- New tuning of multiphysics is rather aggressive for some members (namely for the members 4, 13, 16 ...and maybe also some others).

- There is a bigger initial spinup for new ALADIN-LAEF system due to the perturbed observations (that was expected).

- Generally (and even in spite of the first 2 points) the scores are clearly better for new ALADIN-LAEF system in comparison with the old one and with the pure downscaling of ECMWF EPS forecast.

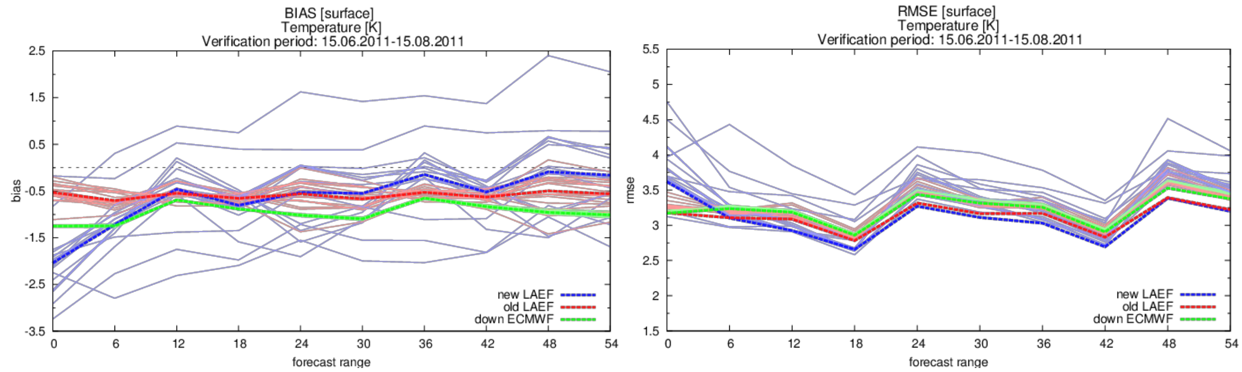


Fig 11: *Ensemble mean BIAS for the Temperature at 2m (bold dashed lines, left) and RMSE (bold dashed lines, right) and the BIAS/RMSE respectively for 16 individual ensemble members (thin lines). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

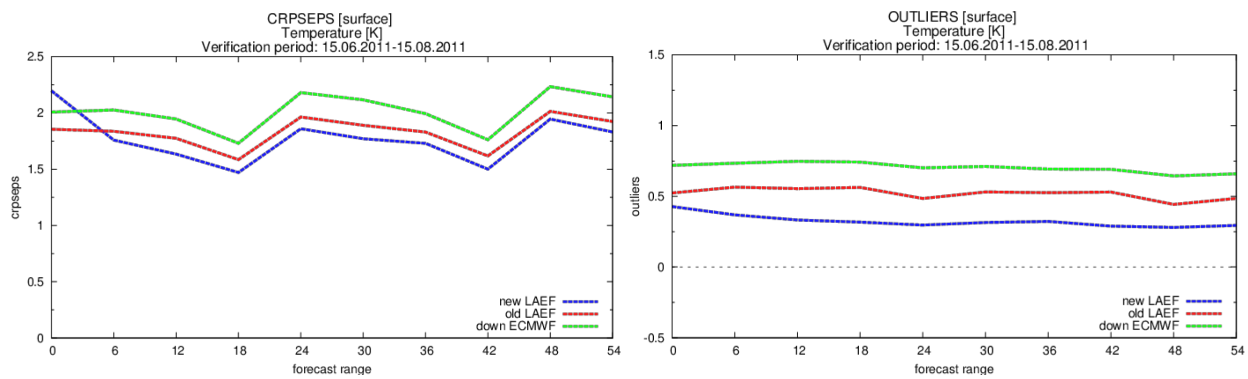


Fig 12: *Continuous Ranked Probability Skill Score for the Temperature at 2m (left) and Outliers (right). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

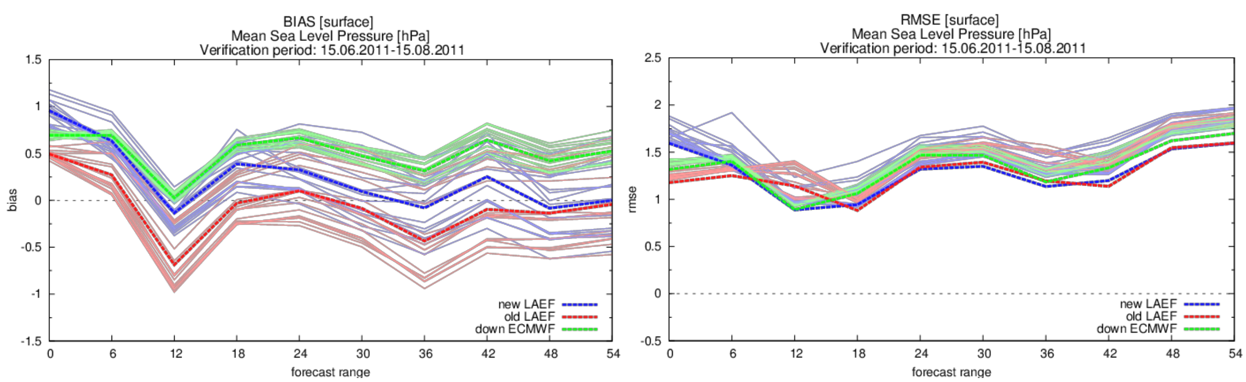


Fig 13: *Ensemble mean BIAS for MSLP (bold dashed lines, left) and RMSE (bold dashed lines, right) and the BIAS/RMSE respectively for 16 individual ensemble members (thin lines). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

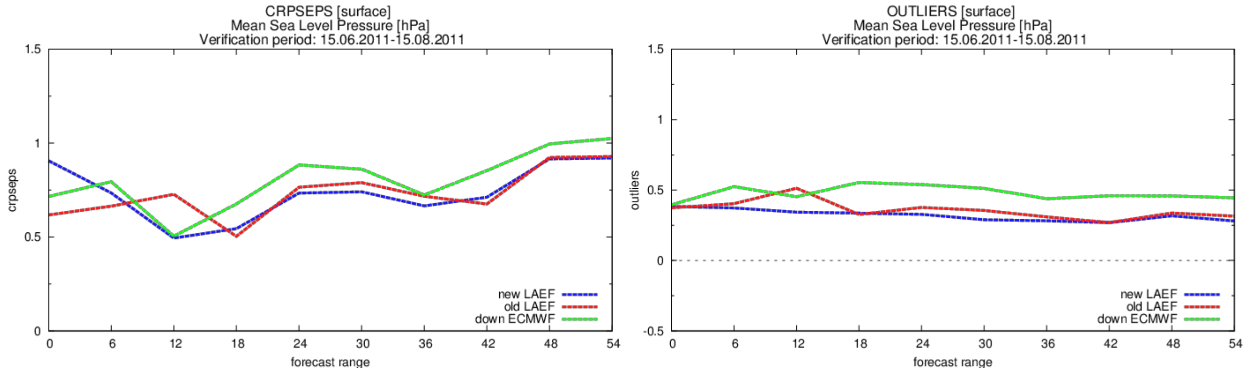


Fig 14: *Continuous Ranked Probability Skill Score for MSLP (left) and Outliers (right). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

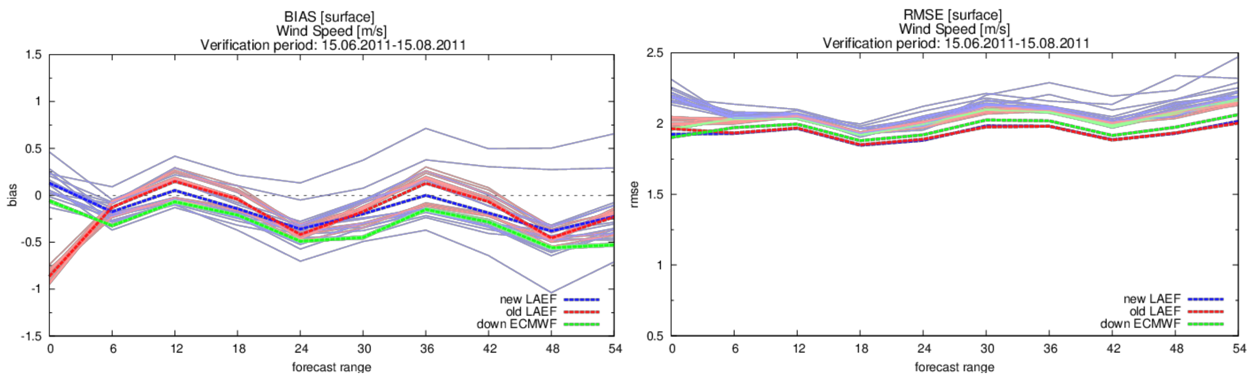


Fig 15: *Ensemble mean BIAS for the Wind Speed (bold dashed lines, left) and RMSE (bold dashed lines, right) and the BIAS/RMSE respectively for 16 individual ensemble members (thin lines). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

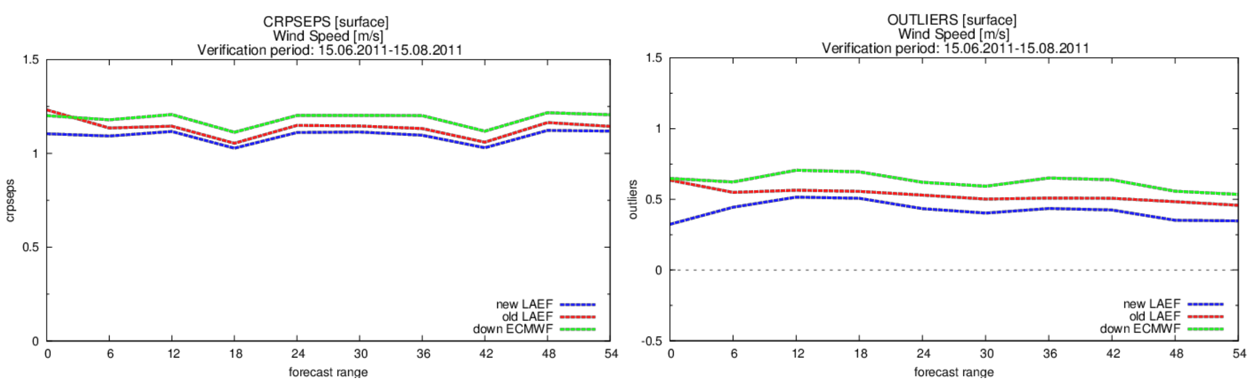


Fig 16: *Continuous Ranked Probability Skill Score for the Wind Speed (left) and Outliers (right). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

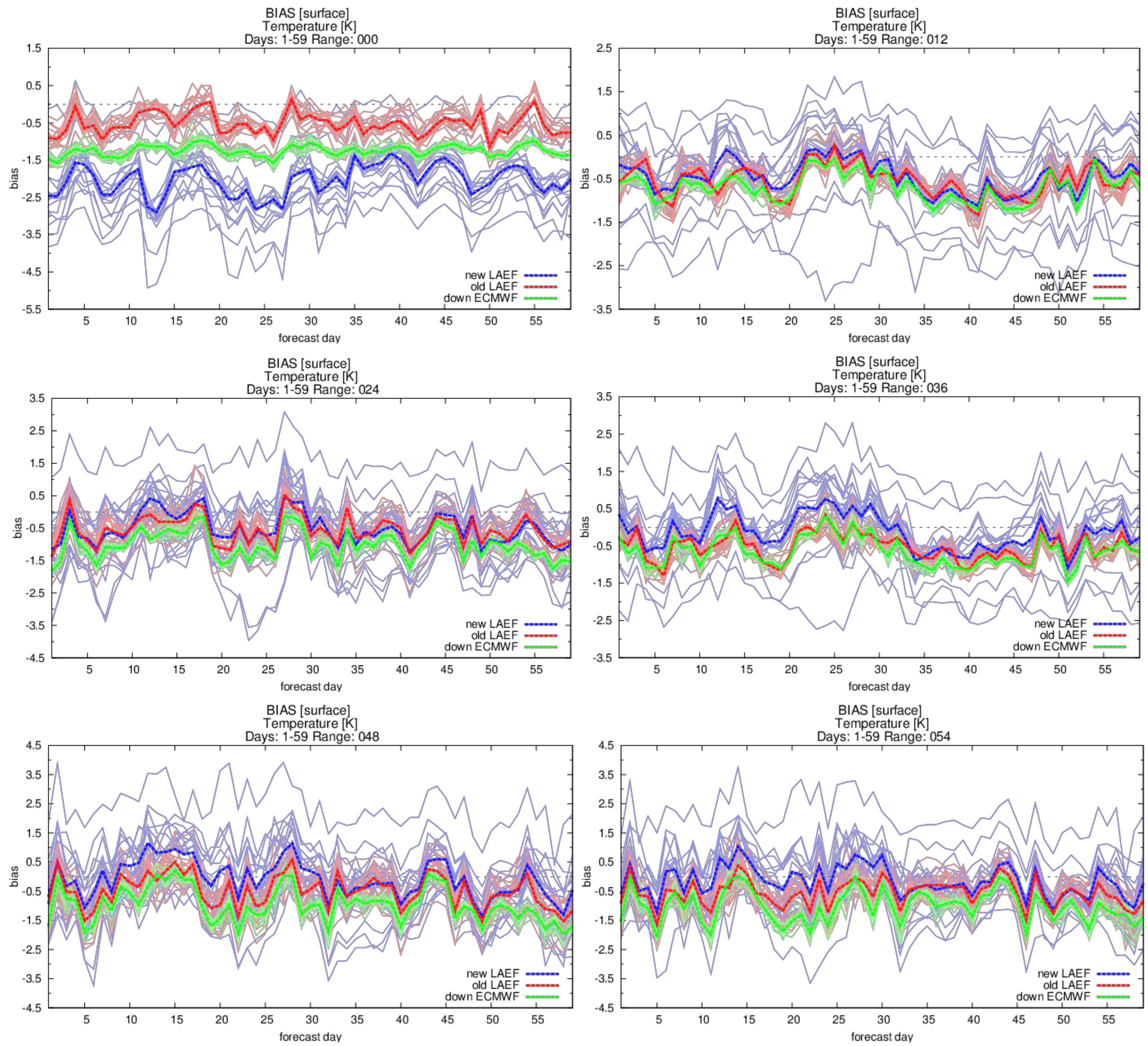


Fig 17: *Ensemble mean BIAS (bold dashed lines) for the Temperature at 2m by forecast days and the same for 16 individual ensemble members respectively (thin lines) for range 00, 12, 24, 36, 48 and 54 (ordered from the upper left corner to the bottom right). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

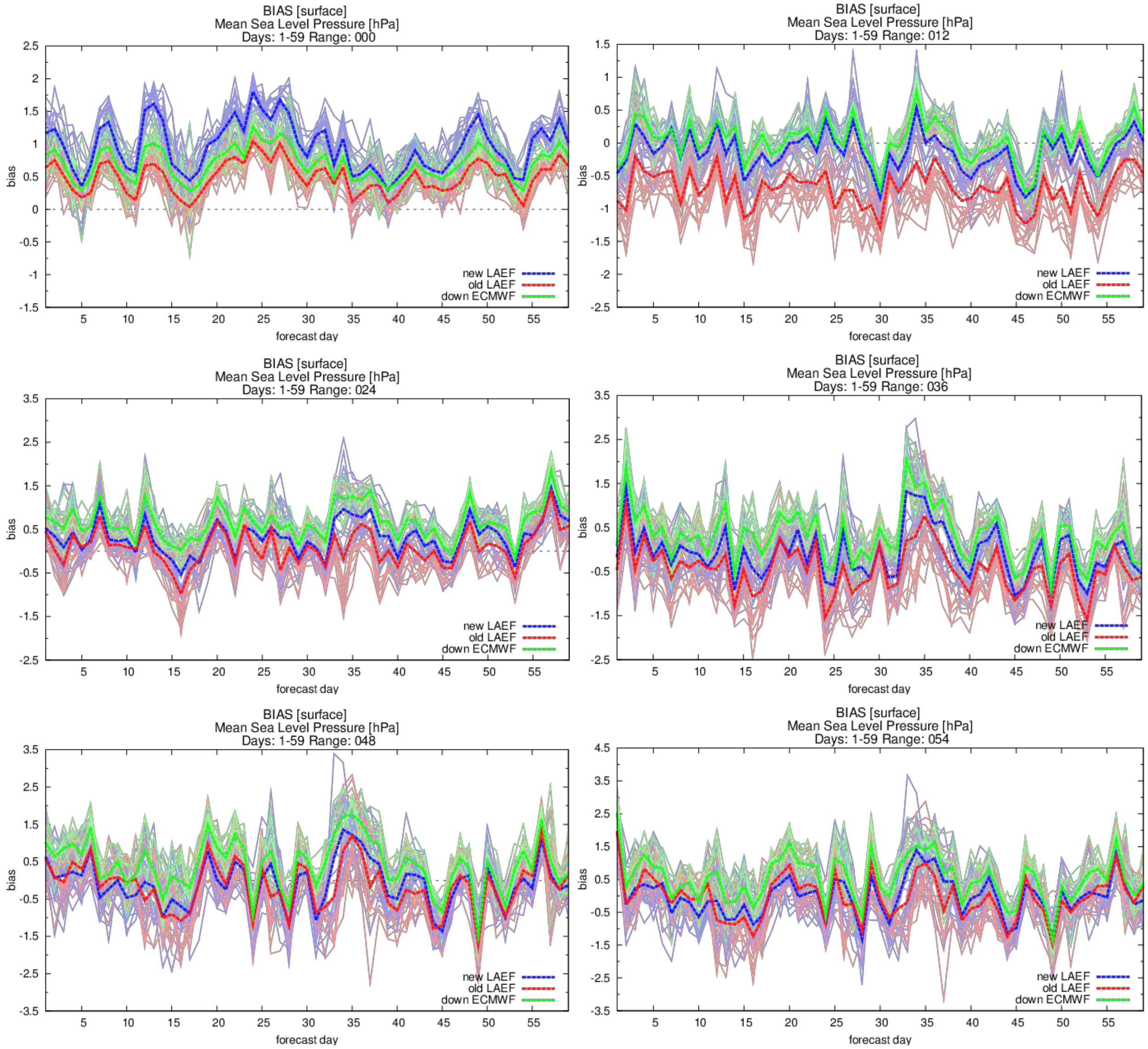


Fig 18: *Ensemble mean BIAS (bold dashed lines) for MSLP by forecast days and the same for 16 individual ensemble members respectively (thin lines) for range 00, 12, 24, 36, 48 and 54 (ordered from the upper left corner to the bottom right). New LAEF is blue, old LAEF is red and ECMWF downscaling is green.*

All verification charts are available at ZAMG server vvhmod:
/daten2/mgruppe/bellus/verif_plot/img/
BPBC-LAEF-DOWN_15.05.2011-15.08.2011/ (3 months)
BPBC-LAEF-DOWN_15.06.2011-15.08.2011/ (2 months)

::Conclusions

Two different coupling methods were tested within the new ALADIN-LAEF system. The first one called “space consistent” coupling uses the LAM perturbed initial conditions (i.e. INIT file) in the position of the first coupling, while the other coupling files came from the driving model (ECMWF EPS in our case). This configuration can be used to reduce the initial spinup of the model, since one of the spinup sources is the imperfect match between interior and boundary conditions at the initial time (Fischer and Auger, 2010).

The second method called “time consistent” coupling uses all LBCs from the driving model, including the first coupling at range 00. Hence, all the lateral boundary tendencies are consistently coming from the same driving model. In such case it is normally recommended to perform some digital filter initialization, because the initial conditions and the 1st coupling are in principle not consistent. This can be very likely the source of some numerical noise. However, if blending technique is used as a pseudo-assimilation procedure before the integration (or as a tool to combine large-scale and small-scale initial perturbations in our case), the additional DFI is not necessary.

Quite surprising results of the above experiments lead us to the fresh idea. In the context of LAM EPS application this could be probably used for an easy and cheap generation of a targeted perturbation. It seems, that such perturbation would act only locally within the areas where it is really needed the most - along the frontal zones. We believe, that in combination with the other used perturbation techniques, it has the potential to bring additional improvement to the ALADIN-LAEF system.

Regarding the verification scores of new ALADIN-LAEF system and observed exaggerated behaviour of marked ensemble members, the same feature was confirmed by Florian Weidle as well (while looking at the pre-operational EPS-grams from the parallel ALADIN-LAEF E-suite). Hence the quick recommendation before switching new system into the operations would be retuning of the integration namelists for those selected members. Otherwise, new ALADIN-LAEF system with all its upgrades proved to be the most advanced version of our LAM EPS and it is suggested for the operational utilization.

There was a great job done by Simona Tascu concerning the LAEF verification package optimization and improved user accessibility. I was really glad to had a chance to participate at least partially on this topic by my minor contribution. However, there are still things to be done and also many ideas how to improve the verification package as a whole. There is no doubt, that the reliable and fast verification tool is very important for the research. Hence, it is highly recommended to continue with this kind of work.

References:

Fischer, C., Auger, L., 2010: Some Experimental Lessons on Digital Filtering in the ALADIN-France 3DVAR Based on Near-Ground Examination. Monthly Weather Review, vol. 139, p. 774-785.