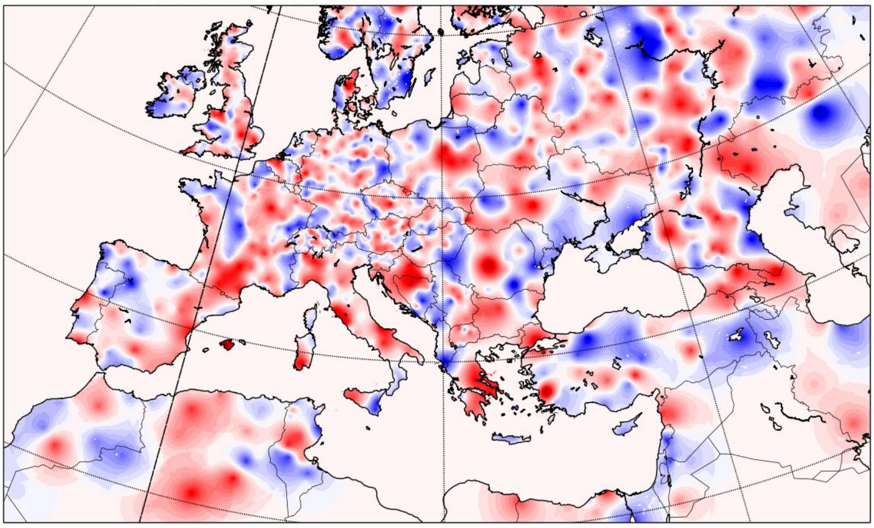
Report on stay at ZAMG

24/04~19/05, 2017, Vienna, Austria

IC and model perturbations for new 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. Coupling (boundary perturbation)
II. Surface assimilation cycle
III. OBS perturbation
    A. External program
    B. Internal (screening)
IV. Model perturbation
    A. Stochastic Physics
    B. Multiphysics

Conclusions
Appendix

::Acknowledgement

I am very grateful to Florian Meier and Alena Trojáková for their valuable help with the technical issues concerning ODB changes between the model cycles. I also very much appreciated the kind invitation for common hiking day event organized by VHMOD group. It was indeed very refreshing and my "batteries" had been completely recharged for the next working days. Many thanks to the whole ZAMG team.

::Foreword

The aim of this 4-weeks RC LACE stay at ZAMG was to implement and test the different perturbation methods within the new ALADIN-LAEF system. It was necessary step towards the operational upgrade of spatial resolution to 5 km with the 60 vertical levels and model upgrade to cycle 40t1 with ALARO-1 physics and all known bugfixes (further referred as LAEF5).

::I Coupling (boundary perturbation)

The ALADIN-LAEF system is driven by the global ECMWF ensemble. The perturbed boundary conditions are technically downscaled from the global grib files. When ECMWF changed their computational grid to the cubic octahedral in early 2016, the interpolation of data became rather complicated task. The formerly used configuration 901 can not be applied now to convert the IFS gribs to ALADIN FA files.

For the operations, the ENS data are still produced at ECMWF also on the original grid (they are upscaled). But one can not fully benefit from the increased grid-point resolution this way. Furthermore, these data are not archived, hence for the research experiments some other tools have to be utilized, like GL. Regardless of GL tool limitations considering the vertical interpolations, we have found another issue during this stay. From currently unknown reason it extracts the surface temperature (sea + land) from provided climatological files instead of getting them from the skin (skt) and sea-surface (stt) temperature fields available in ECMWF grib files. As a result the SST is not changing within the initial conditions of different model runs for given month! The land-surface temperature is not necessarily a big issue as far as the local assimilation cycle is involved. However, for the dynamical downscaling such boundary conditions are not suitable.

::II Surface assimilation cycle

We have implemented surface assimilation cycle for new LAEF5 domain using the local cy40t1 homepack under the ecmwf user kmxy (M. Bellus) on cca cluster, which contains several bugfixes for quadratic coupling interpolation, T2m interpolation, surface SPPT development, etc. (see the list of touched routines below). This was rather big code update considering the operational version, which is still running on cy36t1. The first test for 10 days assimilation loop on dataset from May 2016 seems to be fine. We have also compared the assimilation increments to those from SK operational assimilation cycle at SHMU (see Fig.1).

The code is based on the export version of cy40t1_bf5 with the following modifications and it is used for all mentioned experiments within this report (i.e., this is valid also for the OBS perturbation, stochastic physics and multiphysics experiments tackled in chapters II, III and IV):

surface SPPT mod

- arpifs/adiab/cpg.F90
- arpifs/phys_dmn/mf_phys.F90
- arpifs/phys_dmn/sppten_isba.F90

QCPL bugfix

- arpifs/module/elbc0b_mod.F90

T2M bugfix

- arpifs/phys_dmn/actkecls.F90

other bugfixes (above bf5)

- aladin/coupling/ecoupl1.F90
- aladin/fullpos/fpfillb.F90
- aladin/fullpos/suefpg3.F90
- arpifs/dia/suppdate.F90
- arpifs/fullpos/sufpc.F90
- arpifs/fullpos/sufpd.F90
- arpifs/fullpos/suvpos.F90
- arpifs/namelist/namfpc.nam.h
- arpifs/namelist/namxfu.nam.h
- arpifs/phys_dmn/apl_arome.F90
- arpifs/phys_radi/rrtm_rtrn1a_140gp.F90
- arpifs/pp_obs/apache.F90
- arpifs/pp_obs/ppobsac.F90
- arpifs/setup/suafn2.F90
- arpifs/setup/su0yomb.F90
- odb/pandor/module/bator_init_mod.F90
- odb/pandor/module/bator_module.F90
- odb/pandor/module/bator_util_mod.F90
- utilities/pinuts/module/fa_cadre_mod.F90

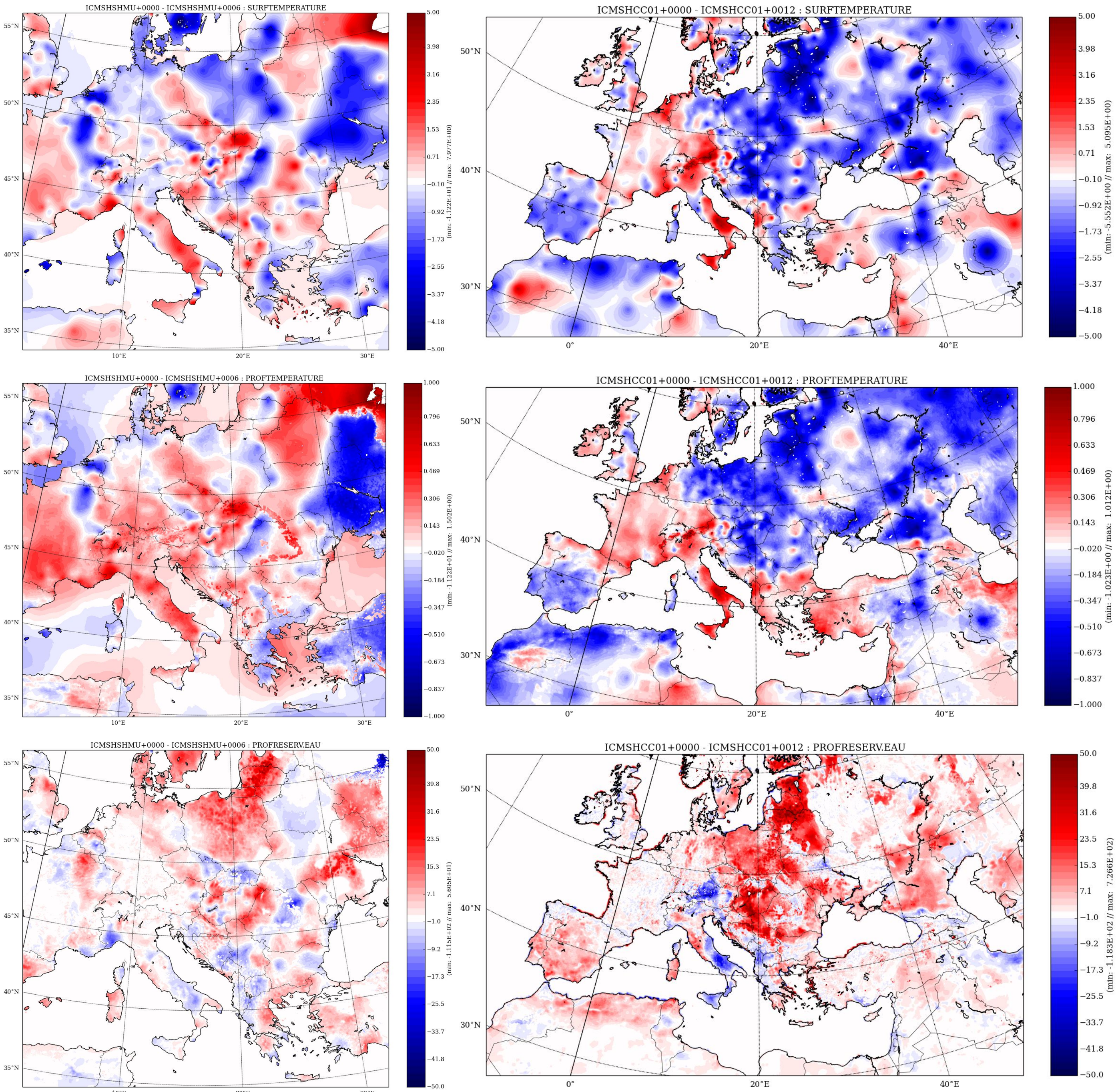


Fig.1: *The assimilation increments (analysis - guess) for SHMU oper suite 4.5 km (left) and LAEF 4.8 km (right) - both on the cycle 40t1, from top to bottom: surface temperature, soil temperature, soil water reservoir. The corresponding color scales are equal for direct intercomparison. Note that the analyzed day is different for SHMU and LAEF, and also LAEF assimilation cycle was tested (warmed up) only for one week period.*

Below is an example of the global statistical scores for LAEF5 (before and after the surface assimilation) averaged over the involved observation sites for one case. Assimilated parameters are being T2m and RH2m only, which are available from the OPLACE archive (number of used measurements can be seen in the brackets):

BEFORE:

```
GEOPOTENTIEL                    OBS-MOD =   43.870 SIGMA =   65.646 (01775)
HUMIDITE RELATIVE A 2M          OBS-MOD =   -0.035 SIGMA =    0.116 (02281)
TEMPERATURE A 2M                OBS-MOD =    0.308 SIGMA =    1.738 (02293)
VENT U A 10M                    OBS-MOD =   -0.024 SIGMA =    1.874 (01997)
VENT V A 10M                    OBS-MOD =    0.134 SIGMA =    1.754 (01997)
NON REPERTORIE                  OBS-MOD =    0.000 SIGMA =    0.001 (01919)
```

AFTER:

```
GEOPOTENTIEL                    OBS-MOD =   43.870 SIGMA =   65.646 (01775)
HUMIDITE RELATIVE A 2M          OBS-MOD =   -0.033 SIGMA =    0.109 (02281)
TEMPERATURE A 2M                OBS-MOD =    0.212 SIGMA =    1.313 (02293)
VENT U A 10M                    OBS-MOD =   -0.024 SIGMA =    1.874 (01997)
VENT V A 10M                    OBS-MOD =    0.134 SIGMA =    1.754 (01997)
NON REPERTORIE                  OBS-MOD =    0.000 SIGMA =    0.001 (01919)
```

For the future, there are several possibilities how to create an assimilation loop in ALADIN-LAEF system. The easiest option, from the technical as well as from the cost point of view, is to stick with the current operational configuration, i.e. no introduction of separate assimilation cycle. That means, there would be no additional integration while the assimilation guess is taken directly from the previous production run (12h forecast) - option a). More sophisticated approach, but also with the additional cost, would be the option b) or c) from the table below (see Tab.1). It involves a separate assimilation cycle, but with the advantage to wait for the real time boundary conditions. Therefore, more accurate assimilation guess would be created. The additional two integrations up to 12h - option b) or 4 integrations up to 6h - option c) are required in such case with more complicated data flow in the operational production chain. Because we do not want to change too many things at once (resolution, code version, new ALARO-1 multiphysics, stochastic physics for the surface, etc.), we may decide to start with the easiest option a) and the implementation of separate assimilation cycle postpone until the future LAEF upgrades.

Tab.1: *Assimilation cycle possibilities within ALADIN-LAEF operational suite.*

	prod	separate assim	assim cycle	assim guess
a)	12h lagged	no	12h lagged	none (prod)
b)	12h lagged	yes	12h real-time	2x12h
c)	12h lagged	yes	6h real-time	4x6h

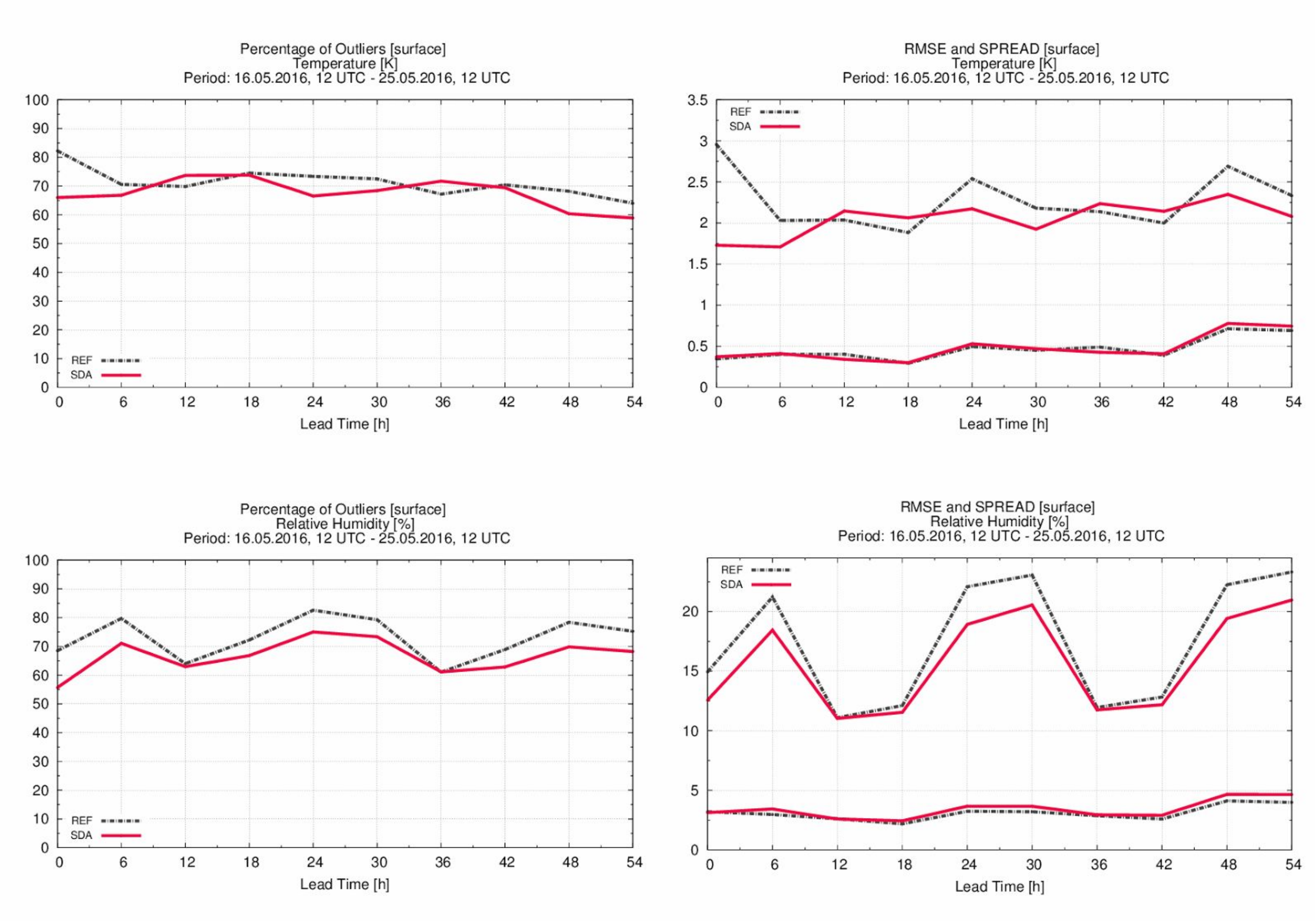


Fig.2: *Temperature (up) and relative humidity (bottom) verification scores for the period of 10 days from 16 to 25 May 2016, 12 UTC run of the full LAEF ensemble (16 members). The percentage of outliers (left) and RMSE of the ensemble mean with ensemble spread (right) are shown for the reference (downscaled ECMWF LBCs, black dashed) and surface data assimilation experiment (SDA, red).*

In the figure above (Fig.2) one can see the statistical scores for the surface data assimilation experiment (SDA) without perturbed observations in comparison with the dynamical adaptation. However, to have a fair intercomparison of the individual LAEF5 perturbation components (like ESDA with the perturbed observations, stochastic physics and multiphysics), all the experiments will be from now on verified against the new reference - surface data assimilation (SDA). Therefore, only the small improvements against this new reference are expected. This is also important because of previously mentioned issue with GL tool. Reference based on SDA is not affected, while the pure downscaling (i.e. dynamical adaptation of global EPS) can not correctly represent the quality of the forecast due to the inappropriate initial land-surface and sea-surface temperatures.

::III OBS perturbation

Following the previous chapter, we have now 2 possibilities how to construct the ensemble of surface data assimilations (ESDA), i.e. to perturb the OBS: a) with the external program ECMAPERT (by Andrea Storto) which is currently used in the ALADIN-LAEF operations but on the old cycle 36t1, or b) by model configuration screening (LPERTURB=.T., NAENSEMBLE=1, NAEMEMBER=$mem in &NAMSCC namelist).

The disadvantage of choice a) might be in the problematic maintenance and tricky compilation of the external code under the newer model cycles. The option b) has also a potential drawback, we can't perturb the observations just after the quality control as it is the case right now (in order to keep all perturbed values involved). Instead, this will be the opposite situation with perturbation at first, then quality check. However, this shouldn't be necessarily a problem at all. One can also look at it as on a higher security level. Although, this was never tested for the surface assimilation procedure, in the situation when ECMA ODB is going to be manipulated by 3DVar-like screening with the subsequent CANARI analysis.

A. External program

The external tool for OBS perturbation within the ECMA database was compiled via gmkpack. However, some manual "hacking" was required in the new project definition under the directory ~/.gmkpack/link/ecmapert for files "entry" and "name" (while the other files were just copied from project BATOR).

```
/entry                  odb/*/*/[pP]ertecma.o
/name                   ECMAPERT
```

The source code has been updated from cy36 to cy40 and can be found on cca here: /home/ms/at/kmxy/src/ecmapert/ecmapert_cy40.tar.

In principle, the SQL queries were changed in pertobs.sql as follows:

```
version cy36:
     SELECT
obsvalue,fg_depar,final_obs_error,hires@update_1,hires@update_3
       FROM hdr,body,errstat,update_1,update_3
```

```
version cy40:
     SELECT
obsvalue,fg_depar,final_obs_error,hires@update[1],hires@update[3]
       FROM hdr,body,errstat,update[1],update[3]
```

B. Internal (screening)

The new OBS perturbation method has been scripted into the existing CANARI (canari.pl) module of LAEF5, where function &ecmapert was modified accordingly. It utilizes the screening configuration (NCONF=2, LPERTURB=.T., NAENSEMBLE=1, NAEMEMBER=$mem). The variable $mem is used to generate the SEED and it is equal to member number. The logical switch LSCREEN=.T. is required (otherwise model aborts).

In the following pictures (Fig.3) one can see an example of surface temperature perturbation by: a) external program ECMAPERT and b) internally - using screening. The results are comparable, both plots have the same scale (+/-3K). Each of the methods have their benefits and drawbacks. For instance, the major disadvantage of internal perturbation lies in a bit slower processing and it also burns more CPU (obviously, the model has to be loaded), but on the other side the maintenance in the future would be much easier in comparison with the external solution.

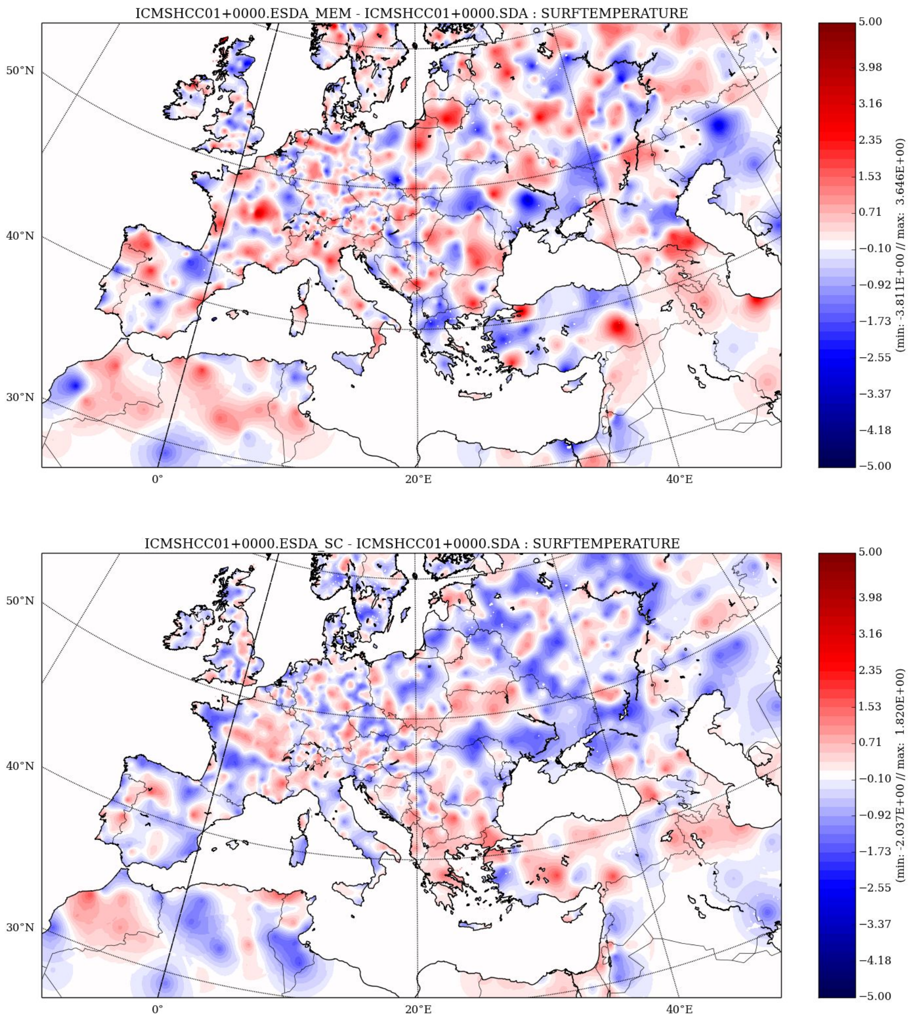


Fig.3: *The temperature perturbation at the surface (pertOBS analysis - reference analysis) for LAEF 4.8 km domain on cycle 40t1. There is ensemble member 01 perturbation by external method (up) and by internal method (bottom) - both initialized with the same SEED number. Deep*

soil temperature perturbation corresponds very well to this too, with exactly the same structure but by one order of magnitude smaller values (not shown).

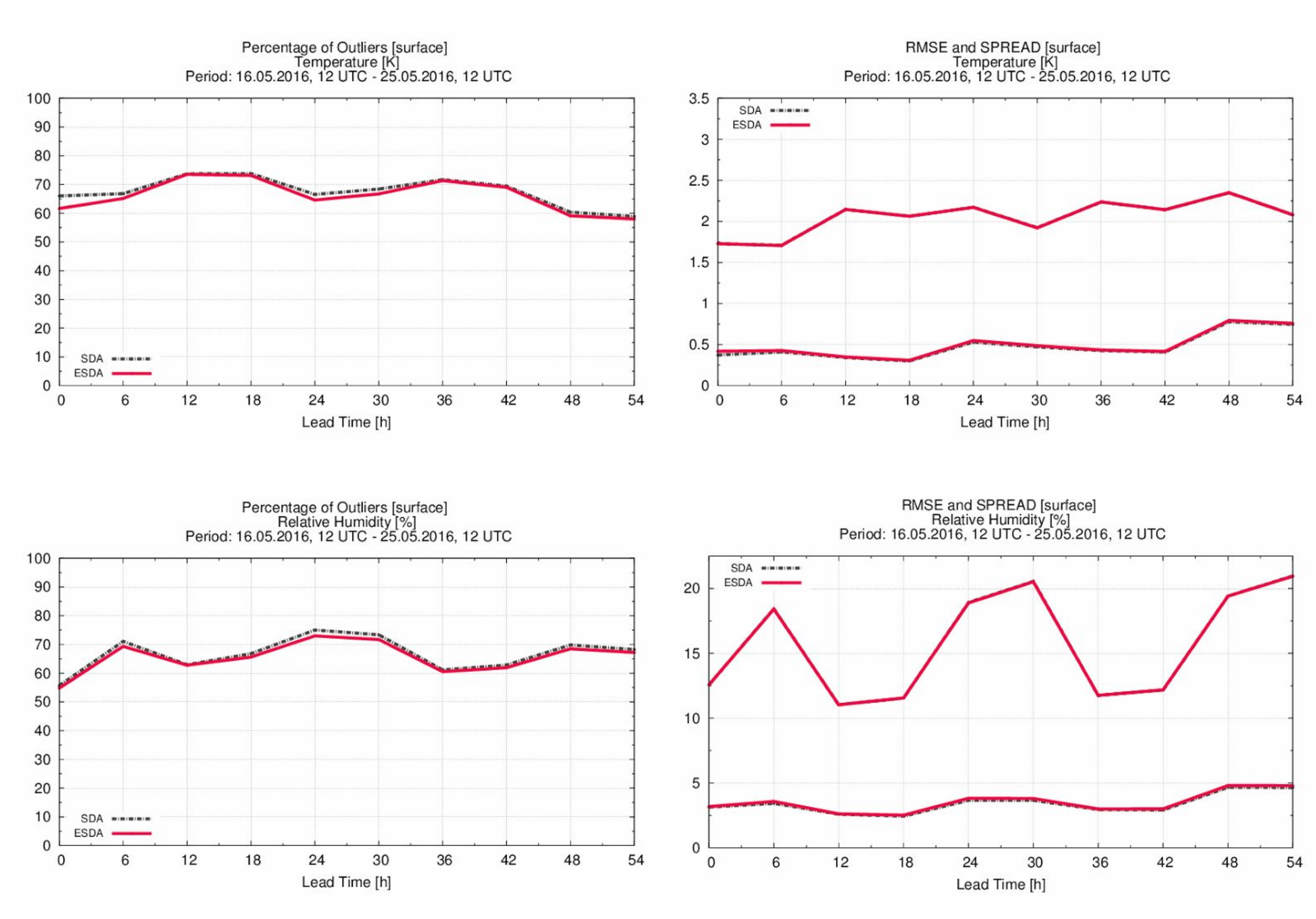


Fig.4: *Temperature (up) and relative humidity (bottom) verification for the period of 10 days from 16 to 25 May 2016, 12 UTC run of the full LAEF ensemble (16 members). The percentage of outliers (left) and RMSE of the ensemble mean with ensemble spread (right) are shown for the reference (SDA, black dashed) and ensemble of surface data assimilations experiment with the perturbed observations (ESDA, red).*

The slight enhancement is most obvious for the shortest lead times. That is according the expectations, since only the ICs were perturbed. However, some positive impact still persists even for longer forecast ranges when we look at the percentage of outliers. Although, there is no significant change in BIAS or RMSE scores, the positive impact of perturbed ICs can be seen for ensemble parameters like spread (increased) and outliers (decreased). It is also mostly pronounced for temperature and humidity fields, whose perturbed measurements are being assimilated into the model.

::IV Model perturbation

The stochastic perturbation of physics tendencies (SPPT) applied on the surface prognostic fields was first time implemented into the ALARO code within ALADIN-LAEF system in 2014 (M. Bellus, “Stochastically perturbed physics tendencies of surface fields in ALADIN-LAEF system”, RC LACE report, 2014). Next year the SPPT code was ported from cy38t1 to the fresh export version of cy40t1 and the initial problems/bugs of new cycle related to the SPPT were successfully fixed (M. Bellus, “Stochastic perturbation of physics tendencies in new cy40t1 with ALARO-1 physics”, RC LACE report, 2015). This code was now re-compiled on the base of cy40t1_bf5 with several other bugfixes (see the complete list of routines in chapter II).

Unlike the stochastic physics, which is not currently used in the operations, the multiphysics parameterizations have been used in the operational version of ALADIN-LAEF for a while to simulate the model uncertainty. Currently, there are 16 different integration namelists applied, which is indeed not very practical from maintenance point of view. It was decided to reduce this huge amount of different configurations by only four with the concentration on the latest ALARO-1 development.

A. Stochastic physics

The stochastic perturbation is called each time step in grid-point space where surface prognostic fields are perturbed. These are the surface temperature, surface liquid water content, surface frozen water content, snow albedo, snow reservoir water content, snow density and water intercepted by vegetation. Seven fields altogether. The direct perturbation of deep soil prognostic fields (such as deep soil temperature or deep soil moisture) was intentionally avoided because they naturally change slower in time, with some delay with respect to the surface.

An example of surface temperature perturbation by SPPT and the soil temperature perturbation in response to it is shown in the following figure (Fig.5). Upon close investigation of the given stochastic patterns it can be noticed, that the perturbation of the surface and soil is spatially well consistent, while the surface temperature is perturbed a bit stronger in comparison with the soil. That is indeed how it should work, since the soil temperature is not directly perturbed by SPPT but rather through the response from the perturbed surface.

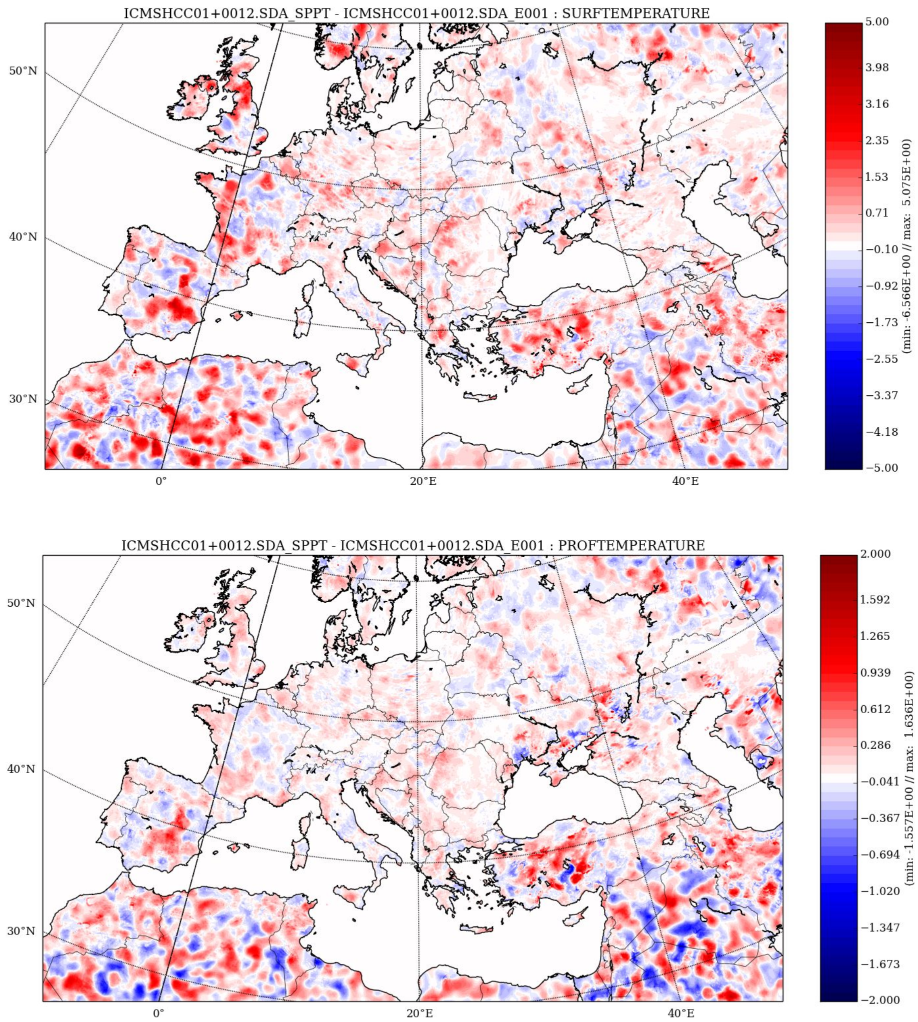


Fig.5: *The model perturbations of surface temperature (up) and its impact on soil temperature (bottom) after 12 hours of integration, note that the scales are not equal here. Stochastic perturbation of physics tendencies (SPPT) is used to perturb surface prognostic fields, the difference is computed against an unperturbed run (both experiments involve surface assimilation without the OBS perturbation).*

The significant reduction of the percentage of outliers can be observed for both perturbed fields - temperature and relative humidity. In contrast with the previous experiment from chapter III, where only the ICs were perturbed through the observations, here the impact is visible rather for the forecast ranges. The stochastic physics has a positive impact on the ensemble spread as well. It has been enlarged mainly for the temperature while RMSE stays untouched (see Fig.6).

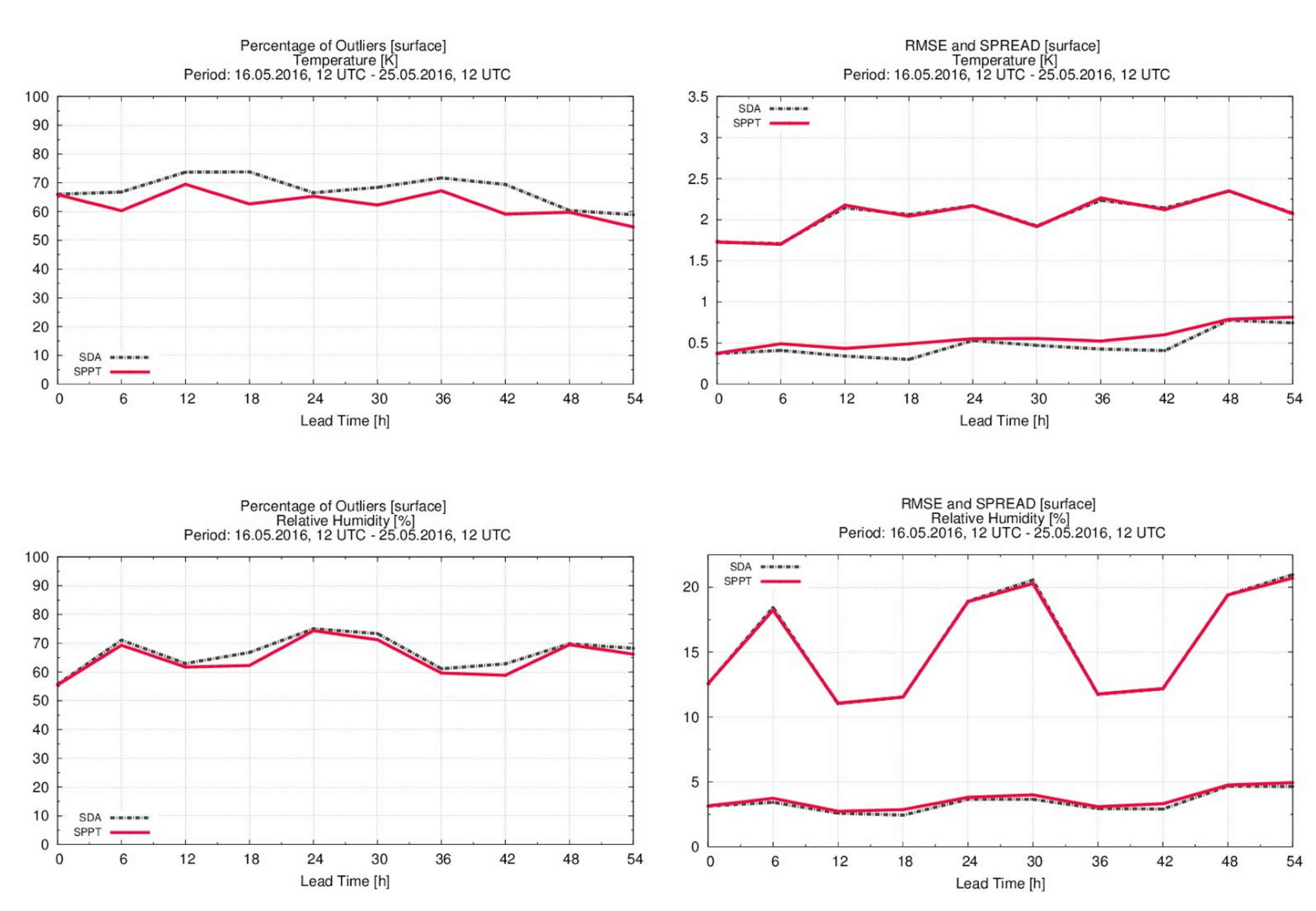


Fig.6: *Temperature (up) and relative humidity (bottom) verification for the period of 10 days from 16 to 25 May 2016, 12 UTC run of the full LAEF ensemble (16 members). The percentage of outliers (left) and RMSE of the ensemble mean with ensemble spread (right) are shown for the reference (SDA, black dashed) and stochastically perturbed physics tendencies of surface prognostic fields experiment (SPPT, red).*

B. Multiphysics

Here the reduced set of different physical parameterizations was tested and compared against the unperturbed run. As for all the other experiments, surface data assimilation was applied in the experiment as well as in the reference.

The ensemble members “04, 08, 12, 16” have ALARO-1 recommended settings (i.e. according the export version). While the members in three other groups “01, 05, 09, 13” (Tab.2), “02, 06, 10, 14” (Tab.3) and “03, 07, 11, 15” (Tab.4) differ in some settings related to the microphysics, turbulence and deep convection. These are the ALARO-1 settings recommended by Christoph Wittmann and Simona Tascu, after their many experiments containing much bigger range of different configurations.

Tab.2: Changes against the reference ALARO-1 settings for ensemble members 01, 05, 09, 13.

	set	ref	description of parameter
LAB12	F	T	ALARO-1 microphysics: Abel-Boutle 2012 (independent of ACRANEB2 and TOUCANS)
LCVGQM	F	T	ALARO-0: modulation of humidity convergence closure
LCVGQD	F		
LENTCH	F	T	ALARO-0: memory in adaptive detrainment
LSCMF	F	T	ALARO-0: mesh fraction's influence on the entrainment rate
LSMGCDEV	T		
LXRCDEV	F	T	ALARO-0: Xu-Randall used in adjustment

Tab.3: Changes against the reference ALARO-1 settings for ensemble members 02, 06, 10, 14.

	set	ref	description of parameter
CGMIXLEN	EL3	EL0	ALARO-1 TOUCANS: equivalent of 'AY' in ALARO-0
CGTURS	QNSE	MD2	ALARO-1 TOUCANS: turbulence model II
LPRGML	F	T	ALARO-0: situation-dependent mixing length
C3TKEFREE	1.39	1.183	ALARO-1 TOUCANS
C_EPSILON	0.798	0.871	ALARO-1 TOUCANS
ETKE_OLAM	0.324	0.29	ALARO-1 TOUCANS
NUPTKE	0.504	0.5265	ALARO-1 TOUCANS

Tab.4: Changes against the reference ALARO-1 settings for ensemble members 03, 07, 11, 15.

	set	ref	description of parameter
CGMIXLEN	EL3	EL0	ALARO-1 TOUCANS: equivalent of 'AY' in ALARO-0
CGTURS	QNSE	MD2	ALARO-1 TOUCANS: turbulence model II
LCVGQM	F	T	ALARO-0: modulation of humidity convergence closure
LCVGQD	F		

LENTCH	F	T	ALARO-0: memory in adaptive detrainment
LPRGML	F	T	ALARO-0: situation-dependent mixing length
LSCMF	F	T	ALARO-0: mesh fraction's influence on the entrainment rate
LSMGCDEV	T		
LXRCDEV	F	T	ALARO-0: Xu-Randall used in adjustment
C3TKEFREE	1.39	1.183	ALARO-1 TOUCANS
C_EPSILON	0.798	0.871	ALARO-1 TOUCANS
ETKE_OLAM	0.324	0.29	ALARO-1 TOUCANS
NUPTKE	0.504	0.5265	ALARO-1 TOUCANS

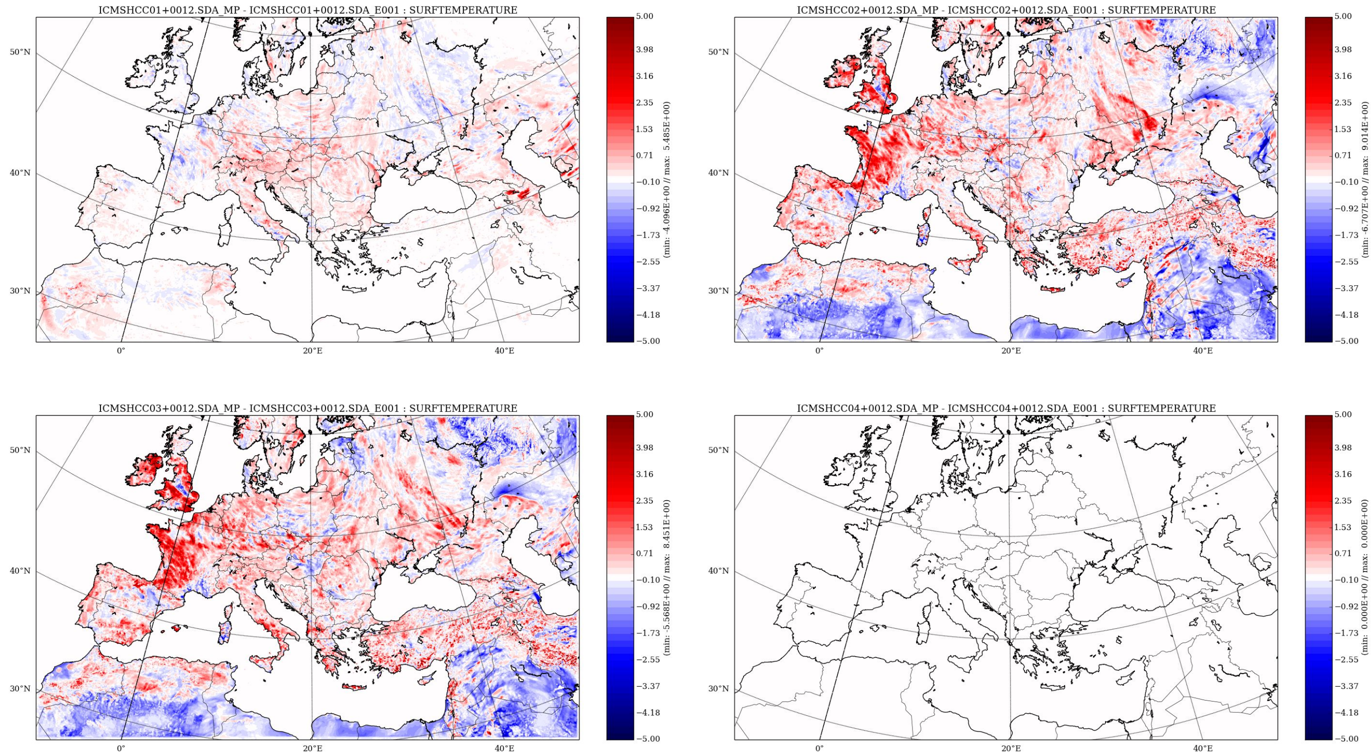


Fig.7: *The surface temperature perturbation due to multiphysics (MP) after the 12 hours of integration. Each panel represents one member of given MP group - altered microphysics and deep convection (top left), altered turbulence scheme (top right), altered turbulence, microphysics and deep convection (bottom left) and default ALARO-1 settings (bottom right). The bottom right panel shows constant zero field, because default ALARO-1 physics was per se used also in the reference (this also proves that MP implementation is correct).*

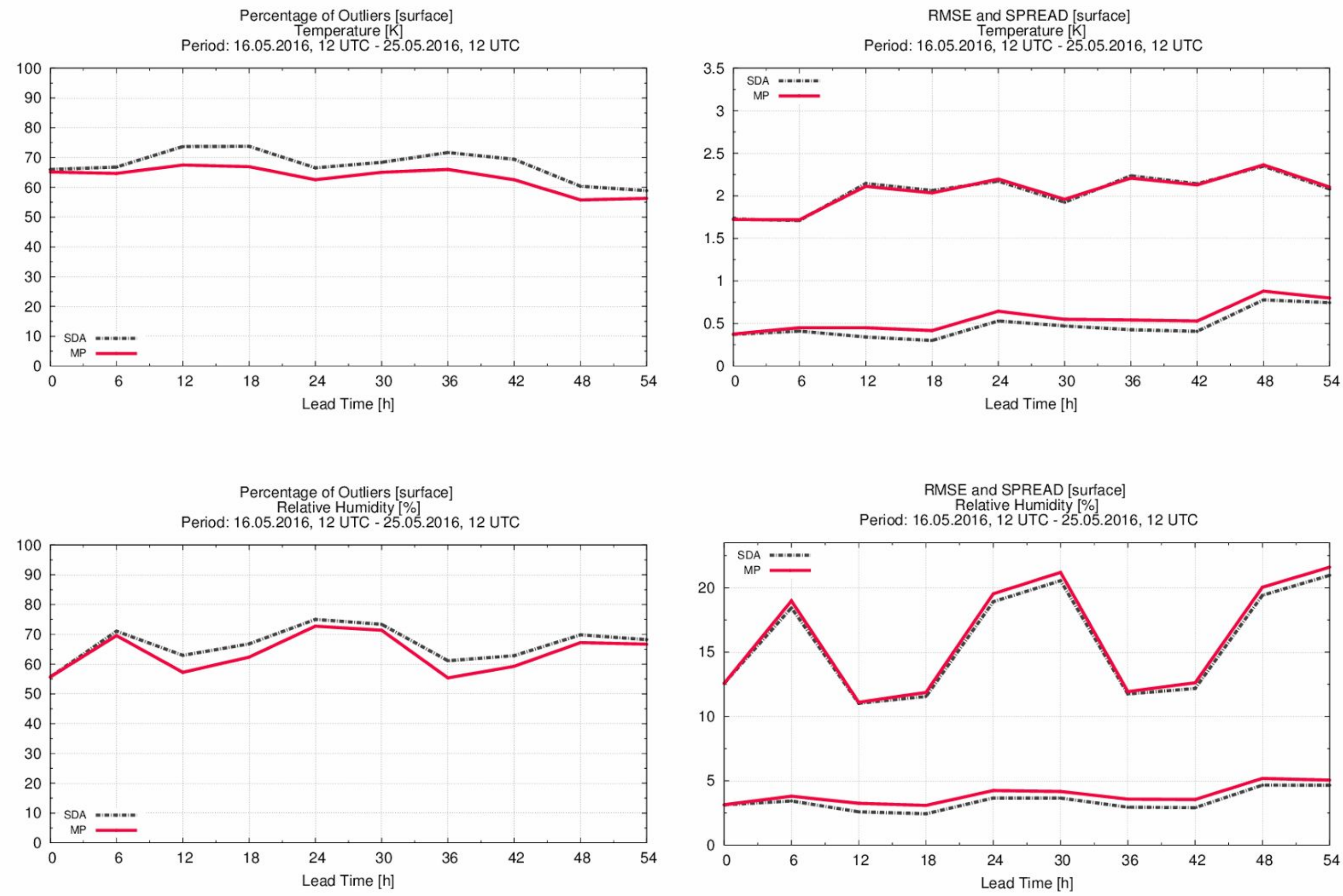


Fig.8: *Temperature (up) and relative humidity (bottom) verification for the period of 10 days from 16 to 25 May 2016, 12 UTC run of the full LAEF ensemble (16 members). The percentage of outliers (left) and RMSE of the ensemble mean with ensemble spread (right) are shown for the reference (SDA, black dashed) and multiphysics experiment (MP, red).*

An example of surface temperature perturbation by multiphysics for one selected ensemble member represented each of the four defined groups is shown on Fig.7. Furthermore, the statistical scores from the verification of multiphysics experiment are compared against the unperturbed SDA on Fig.8. One can see quite significant impact on the reduction of percentage of outliers for both temperature and relative humidity as well as the increase of the ensemble spread. However, at least for the relative humidity such spread enlargement goes along with a slight deterioration of RMSE too.

::Conclusions

Several methods for simulating uncertainties in the ensemble system ALADIN-LAEF were implemented and tested for the new higher resolution domain, new model cycle with upgraded

ALARO-1 physics and even trying new approach for OBS perturbation. The impact of each individual method is clearly positive and in accordance with the expectations and theory. Also the technical correctness of the implementation was confirmed.

Finally we carried out very briefly a comparison against the operational ALADIN-LAEF 11 km as well (not shown). It must be said, that the perturbation methods applied per partes in new LAEF5 system have not defeated the full operational configuration of 11 km ensemble. While the RMSE of ensemble mean for MSLP has been reduced, the ensemble spread was decreased at the same time for all the individual experiments. However, the CRPS has been significantly improved for MSLP in LAEF5, mostly for the first 18-20 hours in comparison with ALADIN-LAEF 11 km. The situation for temperature at 2m has been less beneficial, with smaller ensemble spread and worsening of CRPS. There could be several reasons for such results:

a) the joint impact of all perturbation methods have not been studied yet;
b) the upper-air IC perturbation by blending was not applied in any of the experiments;
c) only very short assimilation cycle was carried out (in order to save computer resources and because of time constraints).

As a next step the LAEF5 should be verified against the operational ALADIN-LAEF 11 km for reasonably long (assimilation) period and with all available uncertainty sources. It is necessary to show, that the new LAEF5 is better and more skilled than the current operational version. As long as it goes to the higher spatial resolution, it might be possible that the global statistical scores won't be enough to show this. Hence, it is strongly recommended to find some case studies as well, where the LAEF5 skills could be clearly demonstrated.

::Appendix

```
-> testing member 01 only (in DEBUG mode)
LAEF5F_SDA     - surface CANARI - OK
LAEF5F_ESDA    - surface CANARI with perturbed OBS via ECMAPERT (external) - OK
LAEF5F_ESDA_SC - surface CANARI with perturbed OBS via screening (internal) - OK

-> full EXPs with 16 members (results saved to ECFS)
PERIOD: 2016051600 ~ 2016052600 (=> 10 days verification for 12 UTC run)
START:  LAEF5Bx (2016051512 +12h forecast as first guess for assim cycle, it has correct
SURFTEMPERATURE unlike downscaled LBCs by GL tool)
OBS:    OPLACE data only

1) LAEF5F_ESDA_SC         - started Thu May 11 16:23:21 GMT 2017  ~2.11 mil BU
   Ensemble of surface data assimilations with perturbed OBS in ECMA ODB using internal
   feature of 3DVar - screening. First guess is also different for each member.

   data (ECFS)
   ===========
```

```
GRIBS: ec:/kmxy/GRIBS/LAEF5F_ESDA_SC.gribs.20160516-20160525.tar.gz
ICMSH: ec:/kmxy/LAEF5F_ESDA_SC
```

2) **LAEF5F_SDA** - started Mon May 15 07:34:29 GMT 2017 ~2.00 mil BU
Ensemble of surface data assimilations without perturbed OBS, just different first guess is used for each member.

```
data (ECFS)
===========
GRIBS: ec:/kmxy/GRIBS/LAEF5F_SDA.gribs.20160516-20160525.tar.gz
ICMSH: ec:/kmxy/LAEF5F_SDA
```

3) **LAEF5F_SDA_SPPT** - started Tue May 16 12:27:25 GMT 2017 ~1.82 mil BU
Ensemble of surface data assimilations without perturbed OBS, just different first guess is used for each member. Model is perturbed by SPPT for the surface prognostic variables (Sigma=0.25, Tau=7200, L=500000).

```
data (ECFS)
===========
GRIBS: ec:/kmxy/GRIBS/LAEF5F_SDA_SPPT.gribs.20160516-20160525.tar.gz
ICMSH: ec:/kmxy/LAEF5F_SDA_SPPT
```

4) **LAEF5F_SDA_MP** - started Thu May 18 11:56:42 GMT 2017 ~1.70 mil BU
Ensemble of surface data assimilations without perturbed OBS, just different first guess is used for each member. Model is perturbed by multiphysics (4 ALARO-1 namelist groups).

```
data (ECFS)
===========
GRIBS: ec:/kmxy/GRIBS/LAEF5F_SDA_MP.gribs.20160516-20160525.tar.gz
ICMSH: ec:/kmxy/LAEF5F_SDA_MP
```

/\/\bell@2017