Prague, 28.03.2025

TOWARDS SMOOTH INTEGRATION OF ALARO WITH SURFEX

Author: Gabriel Stachura

Supervisor: Ján Mašek

Introduction

The report sums up work which has been done during a 3-week-stay in Prague in March 2025. The stay was a part of the SU3.9 task in the RWP 2025. A few currently unresolved issues were recommended to tackle during the stay. At first, work from the previous stay should be completed which involved:

- figuring out how SURFEX snow fields are (or should be) scaled during transformation to the grid-box averaged snow fields in atmospheric file
- adding snow from urban vegetation (town gardens) to the total snow reservoir field in an atmospheric file (SURFRESERV.NEIGE)
- examining currently available way of initialization of snow fields in ISBA-ES (explicit snow) scheme from ARPEGE LBCs containing single-layer D95 snow fields

Besides, we intended to tune some parameters in D95 snow scheme as the scheme was thought to cause a warm bias in T2M in winter (with respect to ALARO with the old ISBA and EBA snow scheme). Although after deeper investigations we revealed that the bias comes mainly from wrong initialization of soil moisture, we ran some basic tests on sensitivity of three tunable parameters in the D95 snow scheme. Therefore, the main part of the report is split into three sections:

1. Inclusion of snow from urban vegetation in the total snow reservoir field in the atmospheric file

2. Addressing warm bias of T2M

3. Tuning of selected D95 parameters

We do not mention in this report about initialization of ES scheme from the global ARPEGE since we still did not succeed to run it. However, with support of Patrick Samuelsson we identified and analyzed the piece of PREP code which initializes ES scheme from ECMWF single-layer snow scheme. Based on that it was concluded that running FULLPROS-PREP from APREGE file to initialize multi-layer ISBA-ES scheme should also be feasible, however, some adaptation is still required.

Common experiment setup

Experiments involved running a 72h forecast starting from either 12th January 2021 00 UTC or 5th February 2024 00 UTC. The dates were deliberately selected to capture significant snowfall that occurred in most of the CHMI domain (the former date) or winter condition with little snow (the latter date). Unless otherwise specified, surface state was initialized from a FULLPOS-PREP file (e927 configuration) which was produced based on old ISBA analysis in CHMI. Considering SURFEX configuration, 8.0+ version was used (this is 8.0 with NWP modifications) with 3 ISBA patches (bare ground, high vegetation, low vegetation), force-restore approach in the soil, D95 snow scheme and TEB with LGARDEN option on (unless otherwise specified).

Adding snow from town gardens to total snow water reservoir

During the previous stay, we modified SURFEX code and introduced a new diagnostic miscellaneous field $TWSNOW_{TOWN}$, which is a sum of SWE on roofs and roads weighted by their relative fractions:

$$TWSNOW_{town} = TWSNOW_{roof} f_{roof} + TWSNOW_{road} f_{road}$$

The field is a tile-averaged quantity and is saved within SURFEX structure DGMT in DIAG_MISC_TEB_n subroutine. The summation is valid if no urban vegetation is considered (LGARDEN=F in the NAM_PGD_SCHEMES namelist). Once the key is activated, urban vegetation becomes a part of urban fraction and has separate mass and energy budget. As far as snow evolution is considered, it is important to notice that while for roofs and roads a simple 1-D snow model is run (under SNOW_COVER_1LAYER subroutine), it is ISBA model that is run for urban vegetation and the choice of the snow scheme can be done in PREP step via namelist variable CSNOW_GD. This can have different value than CSNOW, selecting the snow scheme for nature. Therefore, we modified the data flow so that also SWE from garden is passed to the DIAG_MISC_TEB_n subroutine. The updated formula looks as follows:

$$TWSNOW_{town} = TWSNOW_{roof} f_{roof} + TWSNOW_{road} f_{road} + TWSNOW_{garden} f_{garden}$$

After model physics is calculated, the routine GET_SURF_VARN is called which main purpose is to deliver surface quantities to atmospheric fields. The routine fetches SWE from nature and town tile. Our modification here concerns scaling the fields with tile fractions. The scaling is done while remapping the variables from a tile to the whole domain:

```
PTWSNOW(:) = 0.
DO JI = 1, KI_NATURE
        ! IMASK is a nature mask
        ! ZFIELD7 is a variable where SWE from nature was imported
        PTWSNOW (IMASK(JI)) = ZFIELD7(JI)*PNATURE(IMASK(JI))
ENDDO
DO JI=1, KI_TOWN
        ! IMASK is a town mask
        ! ZFIELD4 is a variable where SWE from nature was imported
        PTWSNOW(IMASK(JI)) =PTWSNOW(IMASK(JI))+ZFIELD4(JI)*PTOWN(IMASK(JI))
ENDDO
```

As a result, PTWSNOW is saved to the atmospheric output file as a land-averaged quantity (grid-box averaged quantity if no lake or sea is present). To make sure the snow balance is correct, we calculated the difference between SURFRESERV.NEIGE and fraction-weighted sum of WSNOW_ISBA, WSNOW_RF1, WSNOW_RD1 and WSNOW_GD1. Due to missing fraction of roofs, roads and garden in the output file, a uniform snow amount equal to 1 kg/m^2 was prescribed over urban patches. Then, the balance on snow-covered areas simplifies to: $SURFRESERV.NEIGE - (f_{town} + WSNOW_{nature}f_{nature}) = 0$. As our results conserve the balance, we know that the total snow field is correctly diagnosed.

Examining sensitivity of tunable snow parameters in D95 snow scheme

Three snow parameters used in D95 snow scheme have been identified as tunable:

- XTAU_SMELT [s] – a lower limit of timestep in melting rate calculation to prevent timestep-dependence (default value 300s.)
- XWSNV [1] – a coefficient that scales surface roughness length in calculation of snow fraction over vegetation
- XWCRN [kg/m^2] - critical SWE value that affects snow fraction over bare ground

All parameters can be changed via namelists NAM_SURF_CSTS or NAM_SURF_ATM. The default and tested values are listed in Tab. 1.

experiment name	REF (default)	S030	S031	S032
XTAU_SMELT	300	0	0	0
XWSNV	5	5	2.5	2.5
XWCRN	10	10	10	5

Tab. 1 Combination of values of the three parameters tested in the experiments.

Looking at domain-averaged evolution of SWE in Fig. 1 it is clear to see that diminishing XTAU_SMELT (green curve) to any value lower than the model timestep increases snowmelt. Although during the forecast period the amount of snow was generally increasing (which means that accumulation prevailed over melting), at the end of a forecast the difference between the experiments grows to 2 kg/m^2 on average. As far as reduction of XWSNV is considered (orange curve), the difference in SWE is very negligible, however, the total snow fraction is locally increased by around 10% (Fig. 2a). Its impact on T2M is unfortunately insignificant. A little more impact is caused if we add reducing XWCRN by half (blue curve). Consequently, snow fraction is increased by at least 10% on the whole area of snow cover occurrence (Fig. 2b). Locally the increase is much higher, up to 40%. It leads to more radiation being reflected and thus slightly lower T2M. The difference is not very significant (around 0.1-0.2°C lower than the reference run), but still bigger than for other experiments.

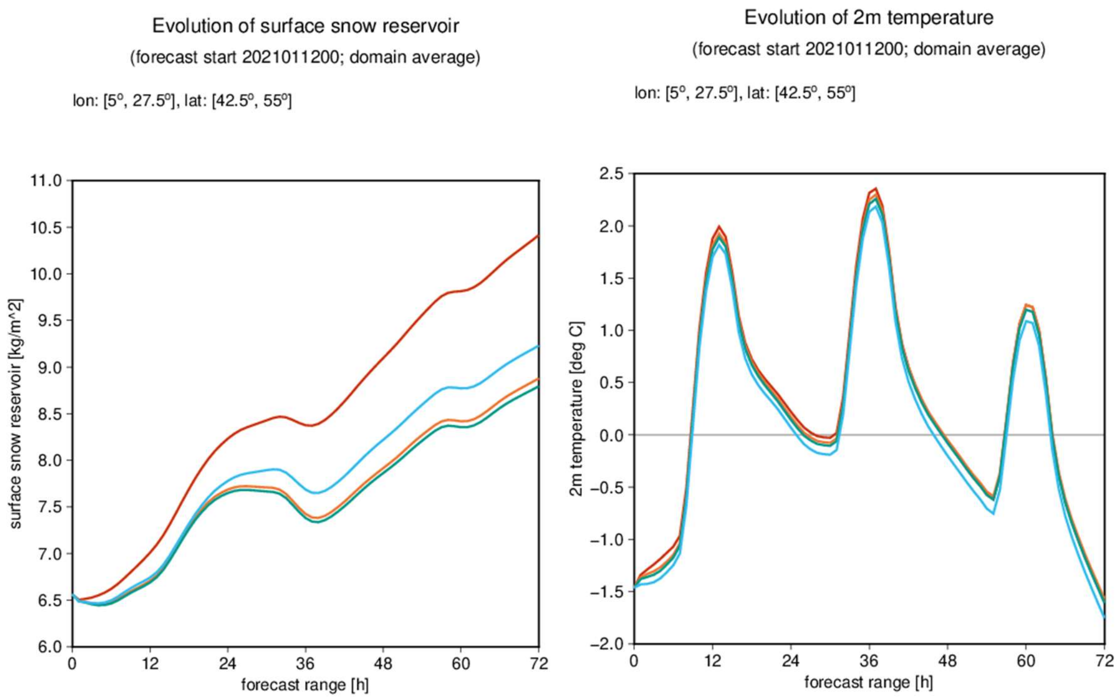

Fig. 1 Domain-averaged evolution of SWE (left) and T2M (right) over forecast length for different experiments. Red color denotes the reference, green – S030, orange – S031, blue – S032.

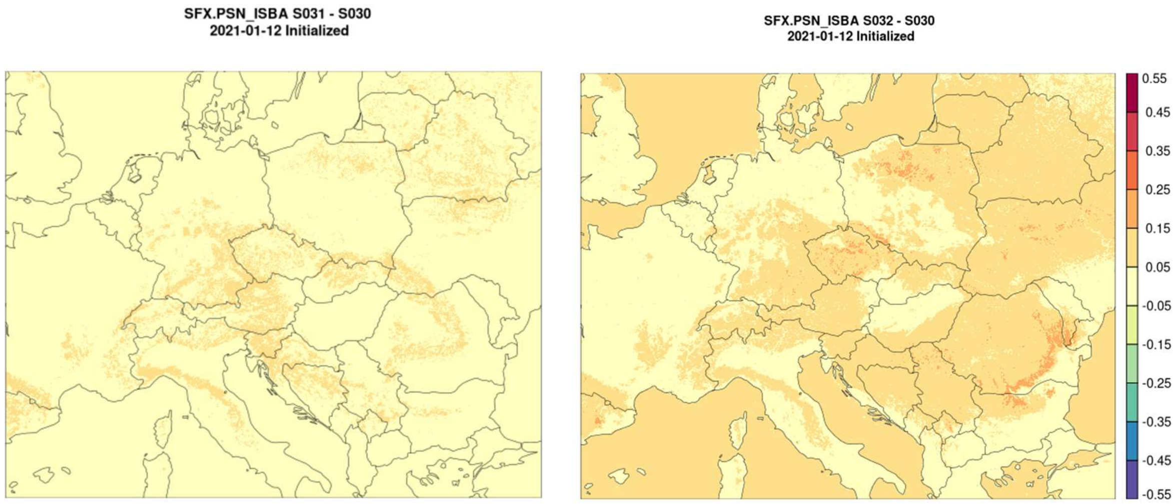

Fig. 2 Spatial distribution of difference in snow fraction between S031 vs S030 experiments (left) and S032 vs S030 experiments (right).

Tracing warm bias of T2M

Warm bias of T2M was initially thought to be induced by the D95 snow scheme. However, since tuning of its parameters affected the bias very little, after careful investigations of spatial distribution of the differences between ALARO+SURFEX and the reference ALARO+ISBA (Fig. 3), we found out that the deviations are the greatest in areas where snow cover is very thin or has just melted (e.g. northern Germany, Hungary, Moldova). It prompted us to search for the origin of bias somewhere else. Eventually we found out that this is because of too warm deep soil temperature (TG2 in SURFEX). The bias gradually increases with time and exceeds 10°C at the end of a forecast (Fig. 3d). As we discovered that even the initial state is too warm, we believe that this bias may be caused by improper initialization of deep soil moisture and/or temperature. After considering each patch separately, we identified that it occurs mostly in the bare ground patch unless the top soil layer is much more humid than the reference (Fig. 4).

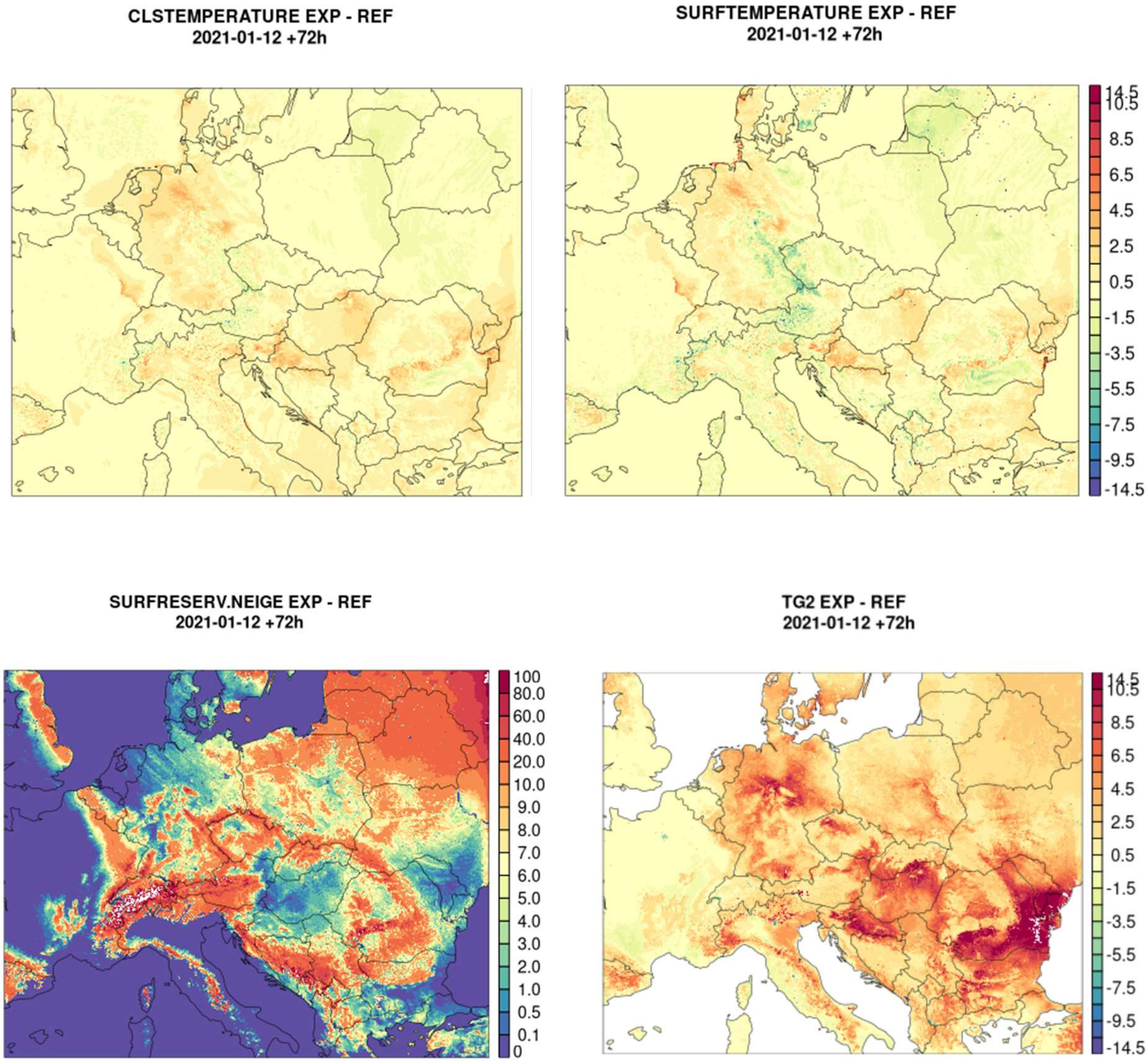

Fig. 3 Spatial distribution of difference between an experiment (ALARO+SURFEX) and the reference (ALARO+ISBA) of: T2M (a), surface temperature (b) and deep soil temperature (calculated for SURFEX as an average across patches) (d); spatial extent of snow cover in the experiment (c).

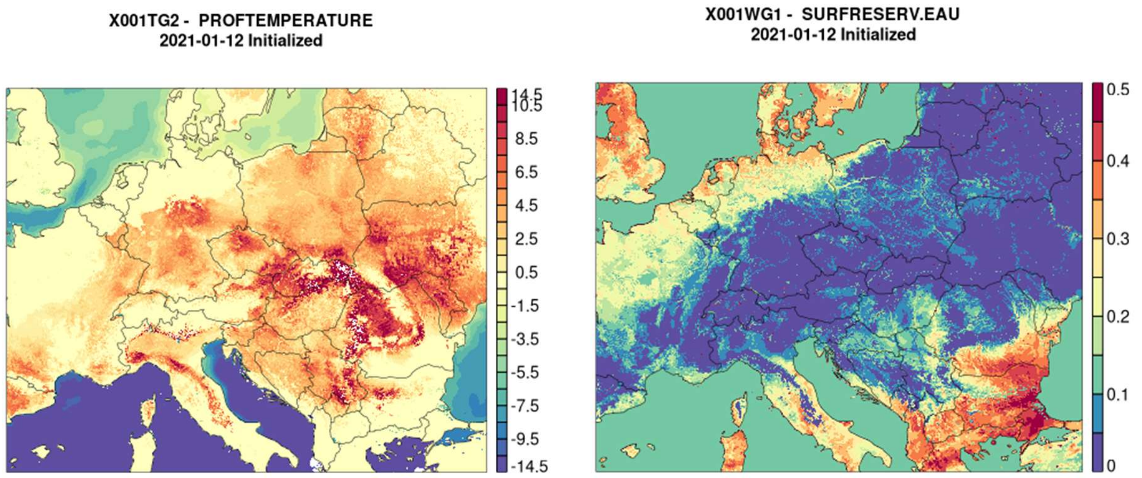

Fig. 4 Spatial distribution of (left): difference between deep soil temperature in patch 1 (ALARO+SURFEX) and PROFTEMPERATURE (ALARO+ISBA); right: difference between liquid water soil content in patch 1 (ALARO+SURFEX) and SURFRESERV.EAU (ALARO+ISBA).

Based on these findings, we decided to experiment with different surface initial conditions. With support of Adrien Napoly from Météo France, we successfully ran PREP from the global ARPEGE analysis but for a different case - 5th February 2024 (since in 2021 they did not use cy46 yet). However, model integration starting from obtained surface initial file crashed due to infinite values of total water reaching the ground (PPG in HYDRO_SOIL routine)[1]. The crash happens only when ISBA is run for urban vegetation, so we tried to switch off LGARDEN option while keep having TEB on. Then the integration went smooth. Scores of three configurations are compared on Fig. 5, evaluated for runs starting on 3-5.2.2024 at 00 UTC:

1. ALARO+SURFEX forecast initialized from ALARO+ISBA analysis by FULLPOS-PREP (red);
2. ALARO+SURFEX forecast initialized from ARPEGE+SURFEX analysis by PREP configuration (green);
3. and reference ALARO forecast with the old ISBA (black).

Although both DA systems diagnose soil moisture in a similar way (from T2M and RH2M), there are major differences e.g. in soil properties (clay and sand fraction data), soil depths

[1] In fact we managed to produce forecast but only if trapping of invalid operations and division by zero operations was switched off (environmental variable VE_FPE_ENABLE). The output looks reasonable except for two surface fields which involve NaNs and infinite values: rain intercepted by vegetation (SFX.RRVEGC_ISBA) and dripping from vegetation (SFX.DRIVEGC_ISBA)

and horizontal resolution of the analyses. The APREGE-initialized forecast is colder from ALARO+SURFEX by around 0.5-1°C, which not only removes the warm bias but actually falls into cold bias. In both forecasts distinct diurnal cycle of the error is visible, which is less pronounced in the reference. Despite cold bias, the APREGE-initialized forecast performs similarly to the reference ISBA as far as RMSE and SD of T2M is considered. The scores are much more promising for 2-m relative humidity - there RMSE and SD are clearly better than for other forecasts. The bias has considerable diurnal variation, but it does not drift away with time. Only in case of wind the scores are slightly deteriorated (not shown here), but this seems to be tunable by e.g. manipulating fake trees option.

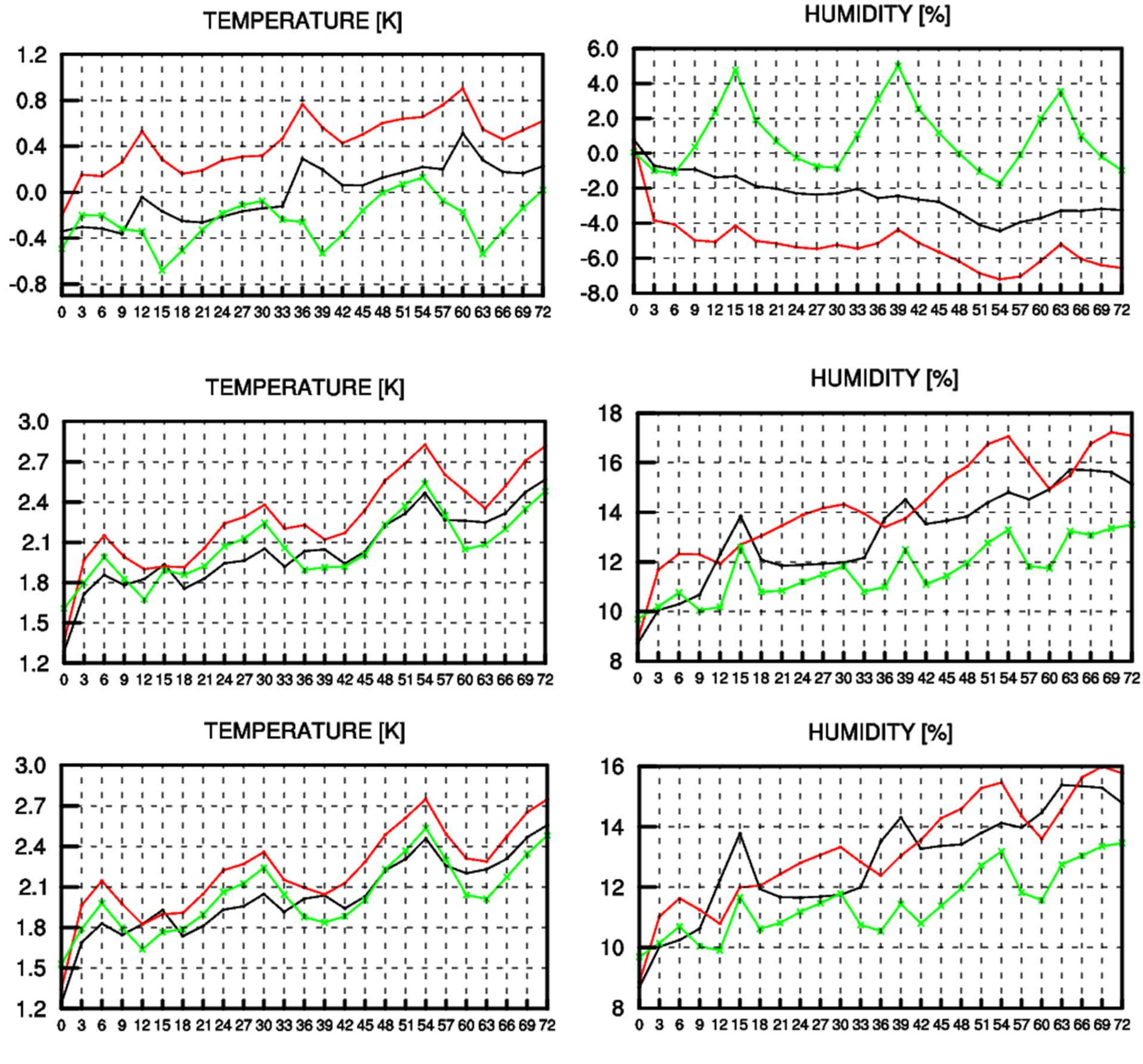

Fig. 5 Evolution of T2M and RH2M bias (top row), RMSE (middle row) and standard deviation (bottom row) with forecast length for the three configurations: ALARO with ISBA (black), ALARO with SURFEX initialized by ISBA analysis (red) and ALARO with SURFEX initialized by the global ARPEGE analysis (green).

Conclusions

Snow in urban vegetation have been added to total snow amount in town and consequently to total snow reservoir field stored in an atmospheric output file. Now the field have complete contribution from two land tiles: nature (ISBA) and town (TEB). Including snow from other tiles (e.g. over frozen likes and sea) is feasible, however, its usefulness for DA is questionable and we were advised by DA community that their algorithm may not be prepared for it. Therefore, we decided to stop at this stage. What should be done in future is checking weather summation works correctly when using multiple ISBA patches and with MEB.

As for tuning of D95 snow scheme, we came to conclusion that in our case decreasing of XTAU_SMELT below the value of forecast timestep has the biggest impact on snow evolution, however it doesn't affect T2M that much. The biggest change in T2M is induced by manipulating the coefficient affecting snow fraction over bare ground (XWCRN).

Tracing of warm bias in T2M in ALARO+SURFEX led us eventually to problems with initialization of deep soil temperature and soil moisture. The comparison of forecasts with surface initialized by ISBA analysis (FULLPOS-PREP) and global ARPEGE analysis (PREP) shows big sensitivity of verification scores on this element. Difference in soil properties between data assimilation system and forecast system seem to be particularly harmful when the former uses old e923 physiography, while the latter ECOCLIMAP II physiography. The problematic warm bias present in ALARO+SURFEX forecast changes sign and reduces in magnitude when ARPEGE analysis obtained with ECOCLIMAP I is used for surface initialization. Taking into account all considered verification scores and impact on other meteorological fields, the outcome is very promising. Since proper initialization is to be treated by local data assimilation with SURFEX, further work on this subject is to be done in close collaboration with DA group.