LACE stay, GSA

9. 10. – 3. 11. 2023.

Flow dependent SPP perturbations for C-LAEF

Author:

Endi Keresturi

endi.keresturi@cirus.dhz.hr

Supervisor:

Clemens Wastl

Preface

This work is a continuation of last year's LACE stay on the same topic. Some parts of last year's report will be repeated here for clarity and easier following.

Motivation

Can a stochastic perturbation pattern in SPP be adjusted to reflect the state of the flow, i.e., to generate the noise in the areas of high uncertainty? The idea was first proposed in Wastl *et al*. (2019) who writes: "At the moment the perturbation field is created randomly without any consideration of the weather or flow situation. Accordingly, it occurs that strong perturbations are applied in very stable areas (e.g., high pressure areas) without any effect, while on the other hand perturbations are not added in sensible areas with strong convection or near frontal zones. Hence it would be much more reasonable to restrict the perturbations to areas with large uncertainties." Following this line of thought, this work will formulate a new type of SPP – flow dependent SPP.

Introduction

The work here was done by using a configuration of the C-LAEF model called C-LAEF1k. The horizontal resolution was increased to 1 km and the model cycle was upgraded to cy46t1. Changes to the code made during the previous stay were phased to cy46t1.

In total, there are 6 physics schemes that are affected by SPP in C-LAEF. We have implemented flow-dependent perturbations to all physics schemes and all parameters. There are 11 namelist switches for 12 parameters which are perturbed. Those are:

1. ZRDEPSRED – snow reduction factor (LPERT_ZRDEPRED)
2. ZRDEPGRED – graupel reduction factor (LPERT_ZRDEPRED)
3. RCRIAUTI – snow autoconversion threshold (LERT_RCRIAUTI)
4. RCRIAUTC – rain autoconversion threshold (LERT_RCRIAUTC)
5. PSIGQSAT – saturation limit sensitivity (LPERT_PSIGQSAT)
6. RSWINHF – Shortwave cloud thickness inhomogeneity factor (LPERT_RSWINHF)
7. RLWINHF – Longwave cloud thickness inhomogeneity factor (LPERT_RLWINHF)
8. XCTP – Constant for T-P correlations (LPERT_XCTP)
9. XCEP – Constant for V-P correlations (LPERT_XCEP)
10. XCED – Constant for dissipation of TKE (LPERT_XCED)
11. XCMF – Closure coefficient at bottom level for convective mass flux (LERT_XCMF)
12. XFRACZ0 – Coefficient of orographic drag (LPERT_XFRACZ0)

Where the first 5 are from microphysics, 6th and 7th are from radiation, 8th – 10th from turbulence, 11th from shallow convection and the last one is from surfex.

Following Ollinaho *et al.* (2017), in C-LAEF, SPP-perturbed parameters $\hat{P}$ are obtained:

$$\hat{P} = P\, e^{c+w\varphi} \tag{1}$$

Where P is the original constant parameter, ϕ is normally distributed stochastic perturbation pattern and *c* and *w* are some constants. This results in a log-normal distribution for $\hat{P}$ which has a nice consequence that perturbed parameters will not change the sign. The stochastic perturbation pattern ϕ varies in space and time independently for each parameter and ensemble member. After the perturbations are applied according to (1), each parameter is clipped to remain inside its physical limits. Additionally, impact of ϕ can be tuned for each parameter separately by adjusting *c*, *w* or its clipping values. Increasing (decreasing) *w* will increase (decrease) the magnitude of the applied perturbations and increasing (decreasing) *c* will move the log-normal distribution to the right (left) resulting in shifting all the perturbations in positive (negative) direction.

Methodology

In this work, the pattern generator will not be changed, and all its settings will remain the same. Instead, the approach taken here is based on modifying the existing pattern by some weights. The idea is to diagnose which areas in the model are the most unstable for each parameter and then to modify the pattern so that it perturbs more in those areas, i.e., to amplify perturbations. The weights are then added to the perturbation field as *w* in (1), i.e., they multiply the pattern. From (1), we see that in the areas where w=1, nothing will happen and where w>1 (w<1), the perturbations will be amplified (attenuated). This means that we need to have our weights satisfying this condition (because we want to amplify perturbations): $w \in [1, W_{max}]$ where W_{max} is some arbitrary number. One can wonder why the weights are not added so that they multiply *P* directly. The reason is that if weights are greater (lesser) than 1, then all the perturbations are moved to the positive (negative) direction. This means that the magnitude of perturbations can be attenuated and that is not what we want. The goal here is to always amplify the perturbations in the targeted regions and that can be achieved by applying weights as in (1).

The question now remains how to find which areas of the domain to target, i.e., how to determine the magnitude and the spatial distribution of weights? The approach taken here is rather pragmatic – for each of 12 parameters a particular model variable will be used to diagnose sensible areas for that parameter. For microphysics, we have chosen a cloud fraction field since microphysics governs the cloud particle formation, growth, and dissipation. The two radiation parameters are also closely related to clouds, so the same cloud fraction field is used.

For three turbulence and one shallow convection parameters, TKE field was chosen since all of them are related to turbulence and for the final orographic drag parameter a 10 m wind speed is used because wind is a measure of the atmospheric flow near the surface. To remark, in the model routine where the 5th parameter is used, cloud fraction field is not yet available, so the cloud fraction from the last model time step is used instead (by writing it to ezdiag array).

For those fields to be used as weights, we need to transform them to [1, W_{max}]. Cloud fraction is already between 0 and 1 so it is almost ready to be used as weights. First, as this field is given for all model levels separately, it needs to be summed up over all model levels. Then, this new field, let's call it *w'*, is slightly modified:

$$w = MIN\left[\left(\frac{w'}{N_l} \times N\right) + 1,\ W_{max}\right] \tag{2}$$

Where N_l is the number of vertical levels in the model, N is some arbitrary real number and w are the final weights. The division by N_l is done to normalize the values (go to 0-1) because N_l is the maximum value in w'. We multiply by a factor of N=1.5 to increase the impact. 1 is added to ensure that minimum value is 1. Lastly, MIN ensures that no value is bigger than W_{max}.

TKE field is given for all levels and has unbounded values greater than 0. This means that we need to calculate the maximum TKE value (TKE_{max}) over all levels and the whole domain to normalize it. This is not a trivial task since the physics part of the model is heavily parallelized and information is not shared between the processors and the calculation of maximum value requires that we have information from all processors. Nevertheless, the problem was eventually solved (calculating TKE_{max} every timestep by using gpnorm_gfl.F90 routine and then using mpl_send/mpl_recv to share the value between the different MPI tasks) and the maximum value of TKE was obtained. Similarly, as before, TKE weights are obtained from:

$$w = MIN\left[\frac{MAX(TKE_{1..90})}{TKE_{max}} \times N + 1,\ W_{max}\right] \tag{3}$$

Where $MAX(TKE_{1..90})$ is the maximum value of TKE over all vertical levels and N=1.5 to increase the impact (see Fig A1 and text related to it in Appendix). Here, maximum value is used to increase the impact of TKE-weights. On the other hand, using maximum value in cloud fraction-weights results in too high impact, i.e., too often the entire domain will be covered in high weights.

Finally, for wind field, the same procedure as for TKE is implemented. Example of weights obtained by the aforementioned procedure is shown in Figure 1.

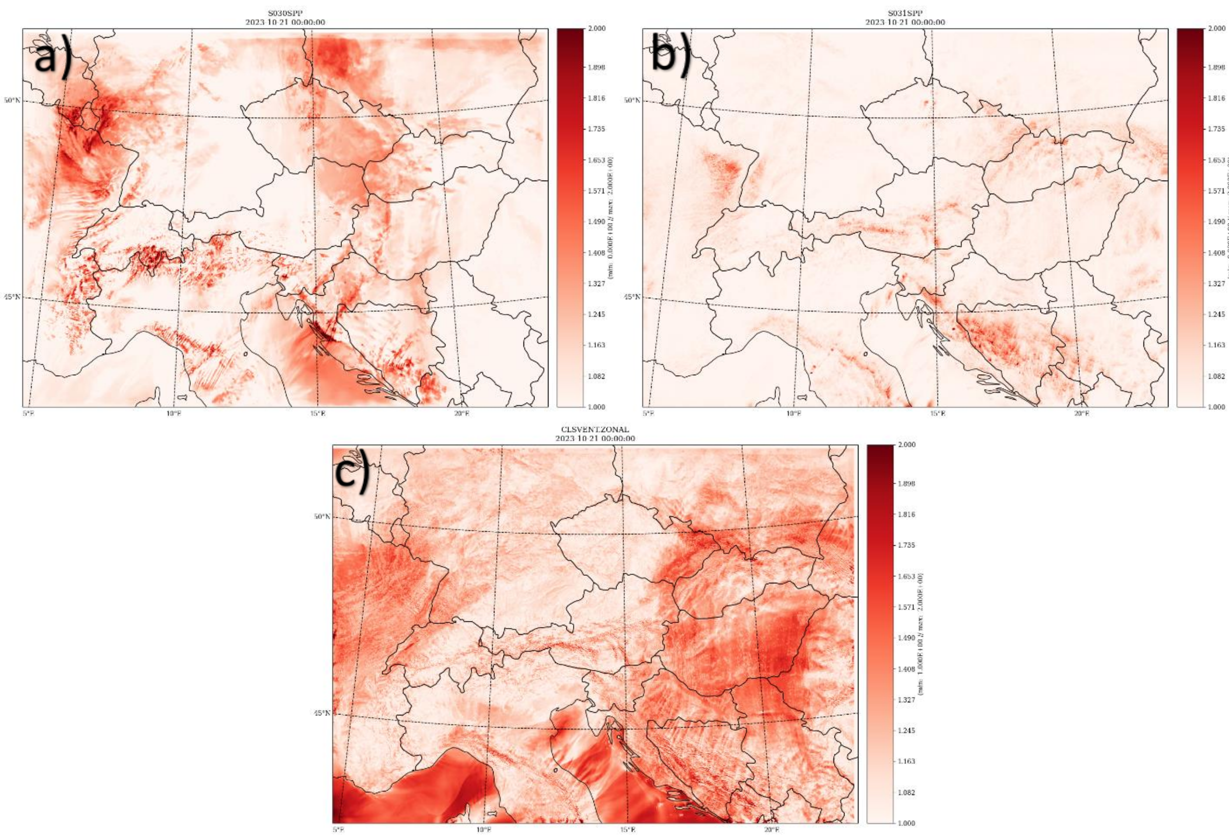

Figure 1. Example of flow-dependent weights obtained from a) cloud fraction, b) TKE and c) 10 m wind speed for ensemble member 2 valid at 21. 10. 2023 at 00 UTC. Keep in mind that the weights shown here are for a case study with high wind speeds.

Results

The impact of adding weights to perturbations was assessed in last year's report (see their Figures 3-5 and 7) and it will not be performed again. In short, the method behaves as expected and adding weights results in positive perturbations becoming more positive, and negative perturbations becoming more negative, i.e., the impact of SPP perturbations is amplified where the weights are higher. This kind of behavior is exactly what we aimed for.

We define 4 experiments which will be used in the rest of this work. **NO_SPP** – reference experiment where no SPP perturbations are used. **SPP** – standard SPP configuration is used. **SPP_FD** – flow dependent perturbations are used and W_{max}=2. **SPP_FD2** - flow dependent perturbations are used and W_{max}=3. The last one was performed to assess the impact of additionally increasing the flow dependent weights' magnitude. Experiments will be run for two case studies with all 17 ensemble members (16 + control), and integration duration of 60 hours.

The first case study is a strong-wind event from 19. October 2023; synoptic situation is shown in Figure 2. A strong low-pressure system moves across Western Europe with extreme southerly

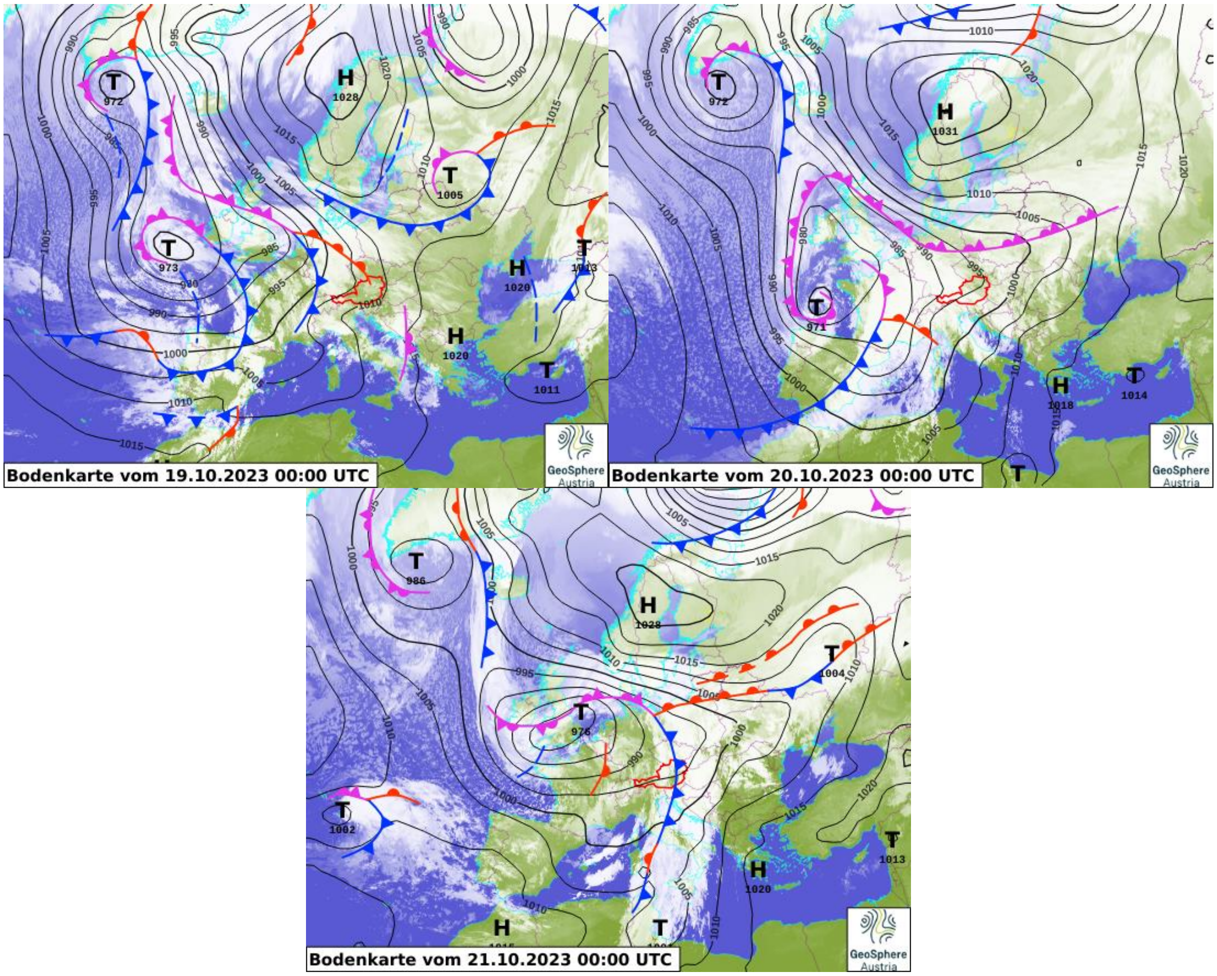

Figure 2. Synoptic situation for the dates written in the bottom-left on each figure.

winds on its front side (wind gusts of about 200 km/h). Also, heavy rain was observed in the southern Alps (100 mm/24 h). The integration was performed for two days (19. and 20.) starting at 00 UTC. Data assimilation was also performed for both the atmosphere and the surface.

Figure 3 shows domain averaged spread for three SPP experiments with respect to NO_SPP and for different surface variables averaged over two model runs. It is clearly visible that SPP_FD has more spread for each variable during the whole integration. The same is true for SPP_FD2 which has more spread than SPP_FD. Figure 4 shows the same but for different vertical levels. The results are similar, but the impact is smaller. This is expected since model physics is more active near the ground. The reduction of spread in all SPP experiments and in the differences between them is visible at the end of model integration. This is explained by the fact that the weather situation is calmer by that time. This also lowers the weights making the impact of flow-dependency smaller (not shown). We have also investigated the impact of flow

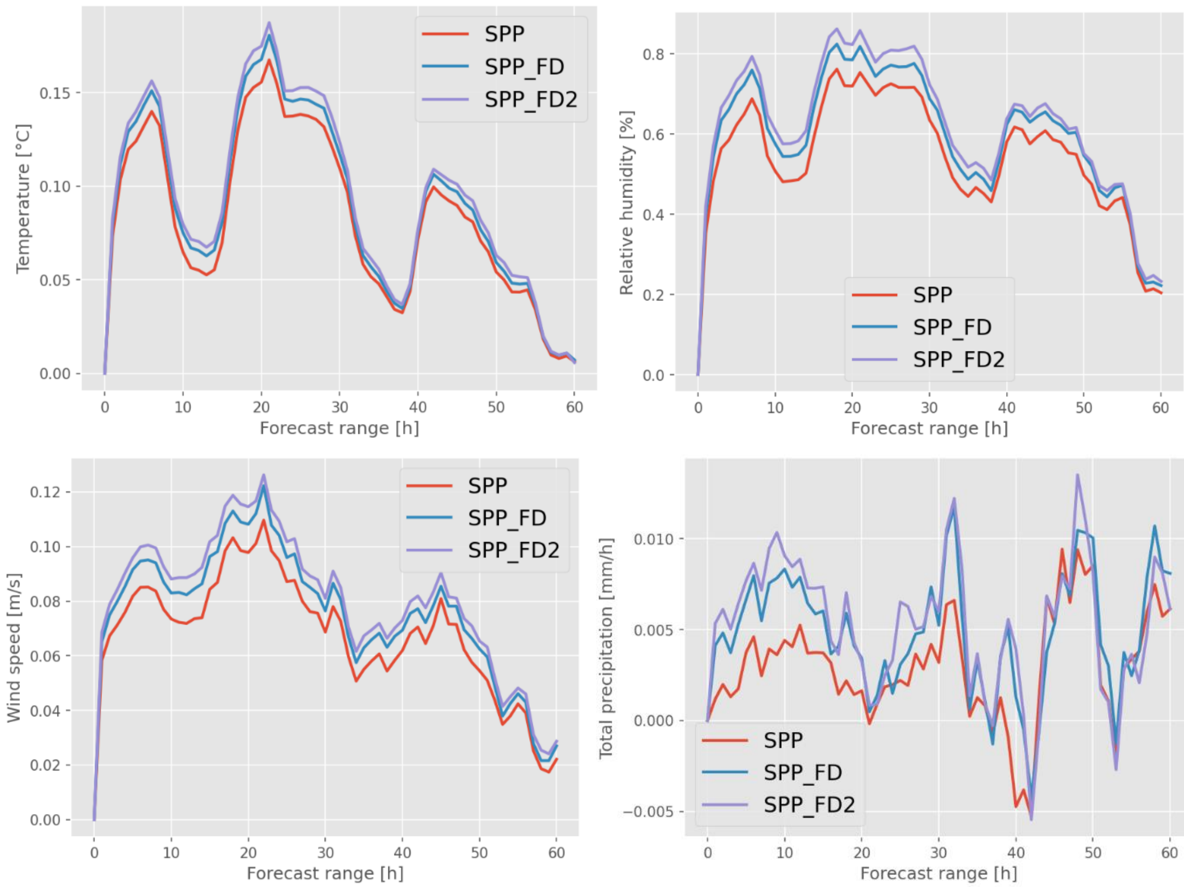

Figure 3. Domain averaged spread (October case) for three SPP experiments with respect to NO_SPP for 4 different variables as written on y-axis.

dependency on mean values on the whole domain (Figure A1) to see if the new method changes the mean state of the model. For temperature and wind speed, no big differences are observed with respect to SPP. For humidity and precipitation there are some minor differences and this, maybe, needs to be investigated further as we don't want to change the mean model state too much.

To illustrate the flow dependency of flow dependent SPP perturbations, Figure 5 shows the difference between SPP_FD2 and SPP wind speed spread over the model domain at different model lead times for C-LAEF run started at 19. 10. 2023. at 00 UTC. The difference between them "moves" with the incoming front and spread is higher in SPP_FD2 in a large part of the domain.

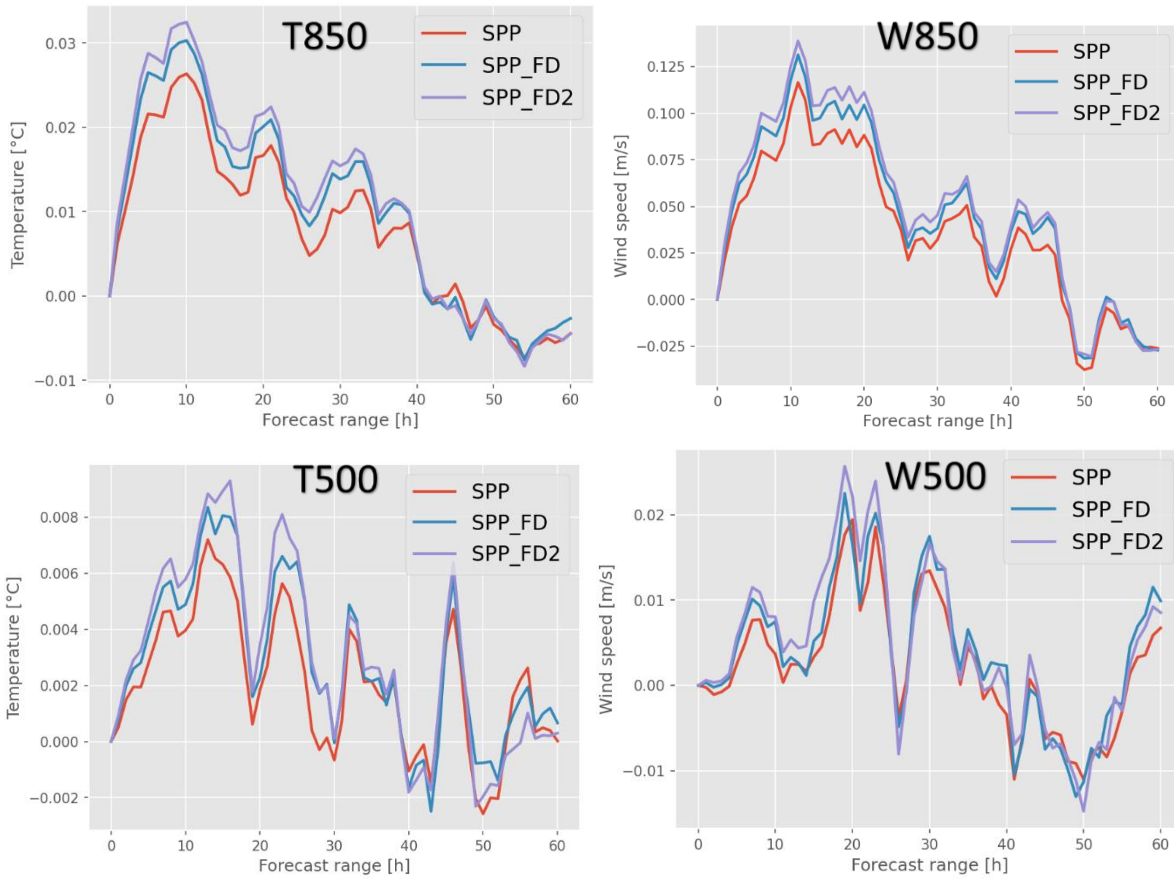

Figure 4. Domain averaged spread (October case) for three SPP experiments with respect to NO_SPP for two different pressure levels - 850 hPa (upper row) and 500 hPa (lower row). Names of variables are indicated on y-axis.

To check the impact on the forecast accuracy and bias, Figures 6-8 show ensemble mean bias, RMSE/spread and spread/RMSE ratio for 3 different variables and for all experiments averaged for 253 stations inside Austria. We see that SPP increases already present biases in the model, and impact of flow dependency is in the same direction, but minimal. RMSE and spread scores reveal that impact of standard SPP on spread is positive and impact on RMSE is slightly negative. However, impact on spread is more positive than negative impact on RMSE, meaning that spread/RMSE ratio is better as can be seen on the plots. Flow dependent perturbations increase spread slightly and increase RMSE slightly less when compared to SPP. The overall impact is similar as in SPP vs NO_SPP - spread/RMSE ratio is slightly better than in the standard SPP. This means that flow dependent perturbations act in the desired direction.

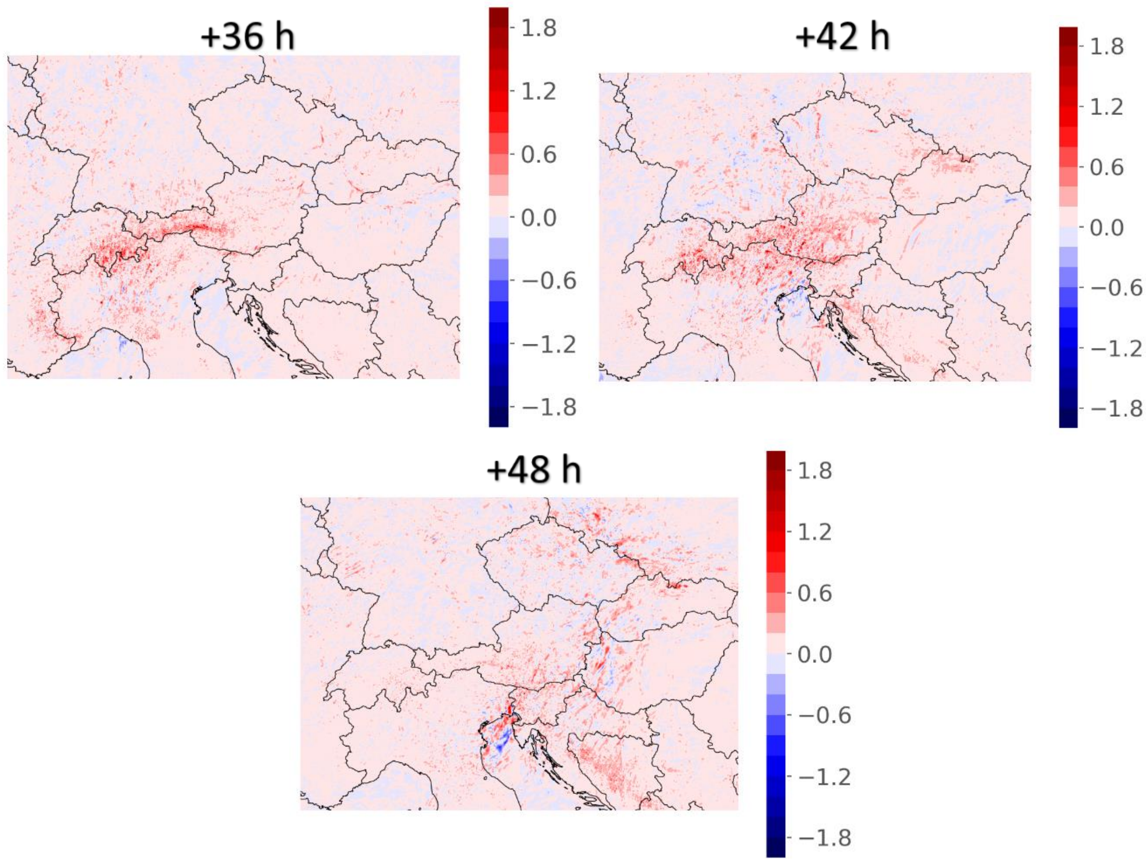

Figure 5. Wind speed ensemble spread difference (SPP_FD2 – SPP) for different lead times and C-LAEF run started at 19. 10. 2023., 00 UTC. Red color indicates higher spread in SPP_FD2.

Looking at precipitation over INCA domain, no systematic differences have been observed during this whole case study. For example, Figure 9 shows 6-h total precipitation from ensemble median for SPP, SPP_FD2, their difference (SPP_FD2 – SPP) and INCA analysis. Both experiments match observations very well and are very similar in total amounts and structure. No systematic difference can be observed, but precipitation in this case is influenced by orography. It would be more interesting to see the differences in a summer convection case on which model physics has more influence.

Figure 6. Bias (up), RMSE/spread (bottom) and RMSE/spread ratio (bottom) for 10 m wind speed. Averaged over 253 stations in Austria and two days in October case.

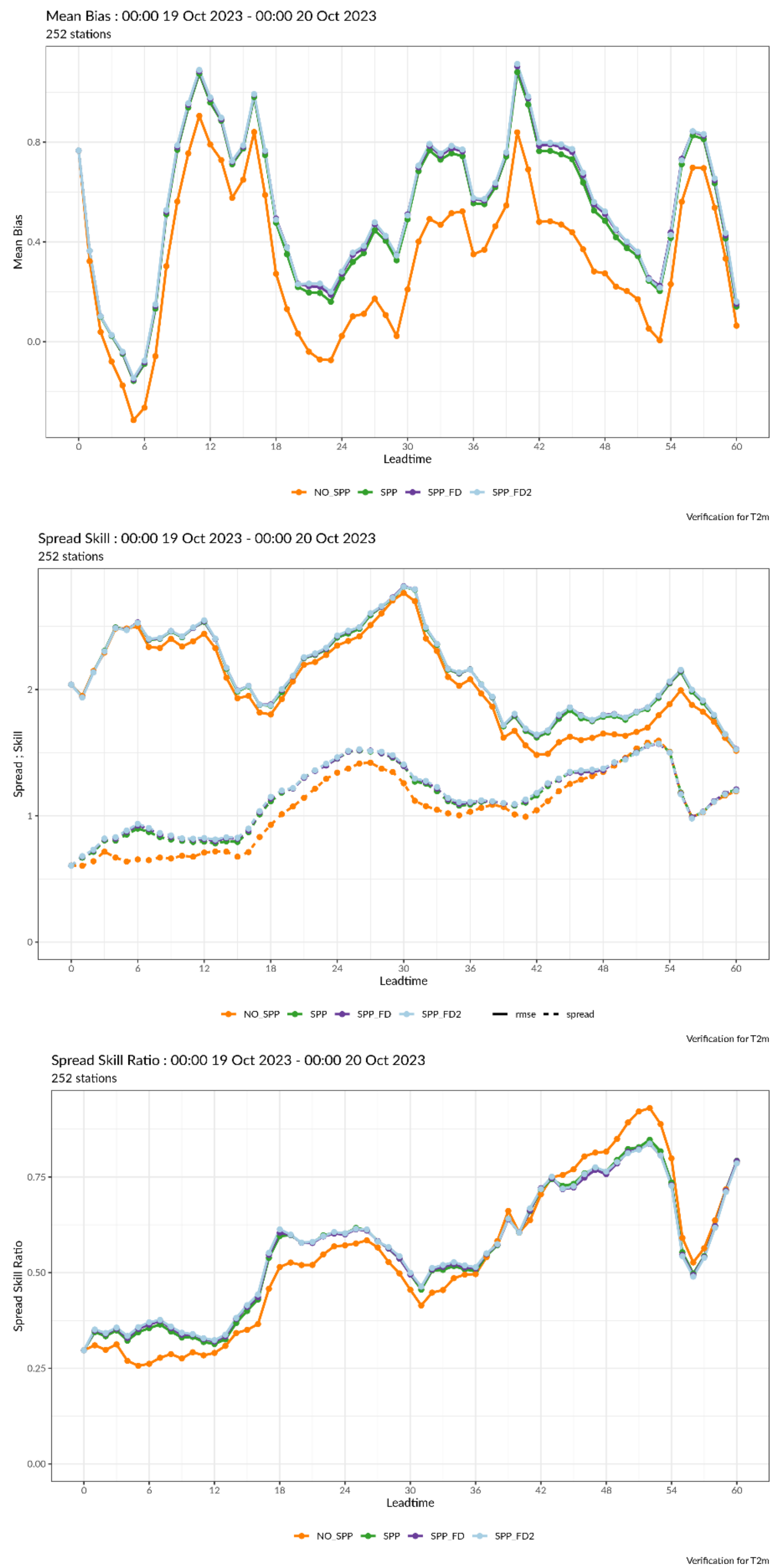

Figure 7. As Figure 6, but for 2 m temperature.

Figure 8. As Figure 6, but for 2 m relative humidity.

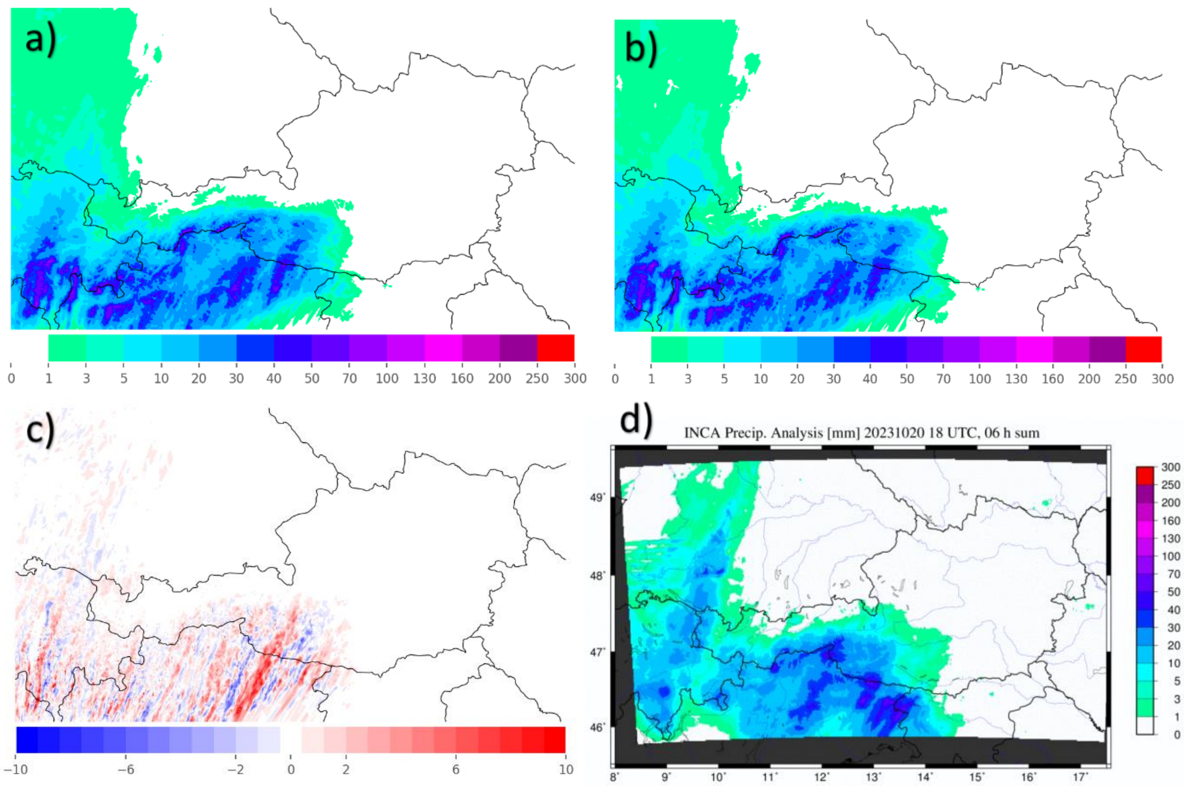

Figure 9. Total 6-h precipitation (lead times: 37-42 h) from ensemble median for a) SPP, b) SPP_FD2, c) their difference (SPP_FD2 - SPP) and d) INCA analysis. Integration has been started on 19. 10. 2023. at 00 UTC.

The second case study is a high-precipitation event from 3. - 5. August 2023. (Figure 10). A cold front from the north Europe caused a cyclogenesis event in Genoa Bay. A shallow cyclone formed and started to move to the east towards Croatia and Slovenia. In front of the moving storm with the inflow of the warm and moist air, heavy convection was triggered. Parts of Slovenia, Croatia and Austria received heavy precipitation (more than 200 mm in 48 h) causing major floodings mostly in Slovenia.

The integration was performed for three days (3. - 5.) starting at 12 (3.) and 00 (4. - 5.) UTC. This time we will drop SPP_FD experiment as SPP_FD2 better illustrates the differences. Data assimilation wasn't available this time and the model was started from interpolated ECMWF's forecasts. This will, of course, heavily affect and degrade the model's performance, so we will not assess the absolute model's performance. Like before, domain averaged spread is higher for SPP_FD2 experiment for all variables and almost all lead times (Figure 11). For vertical levels, SPP_FD2 has more spread, but results are more neutral than for the October case (Figure 12).

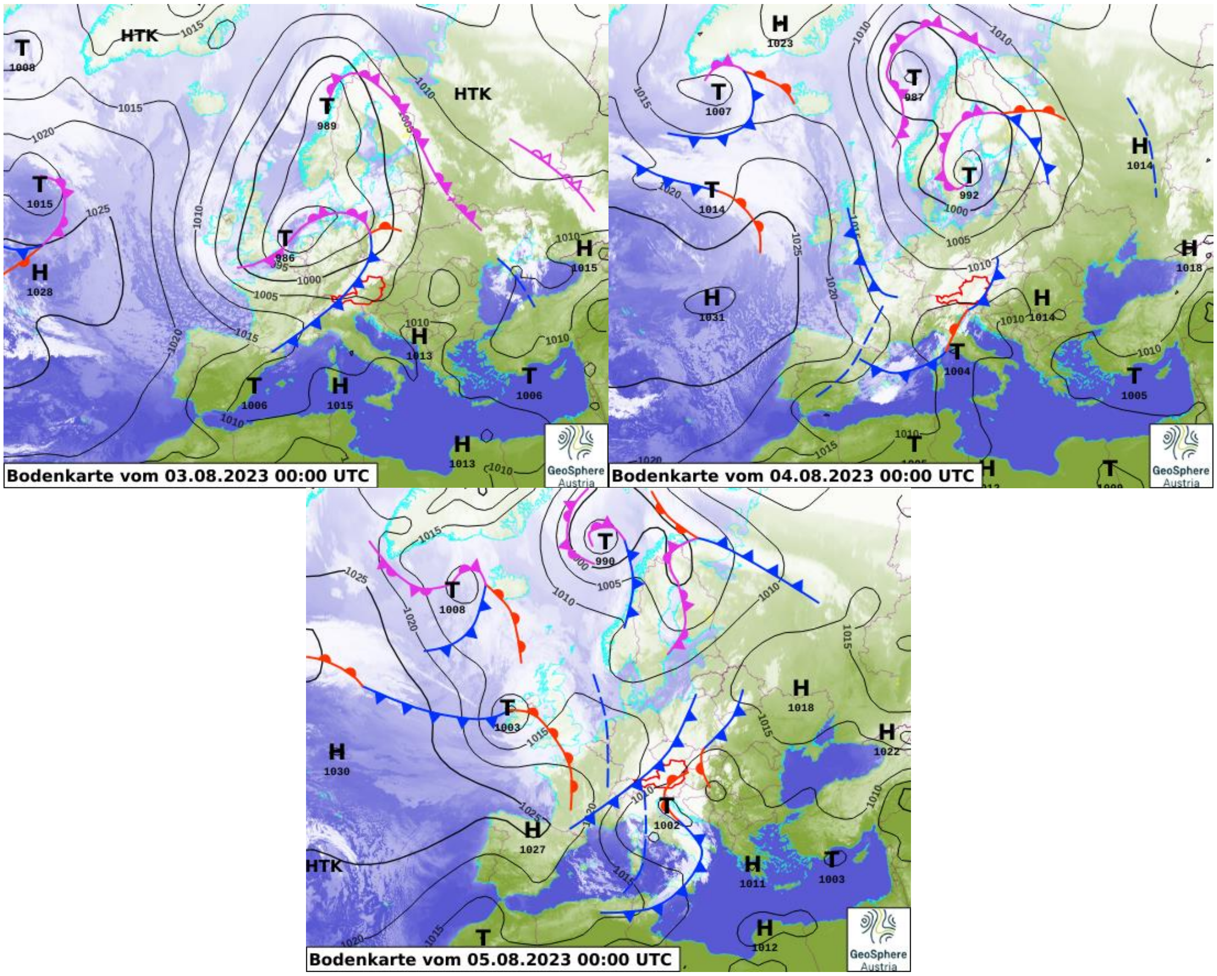

Figure 10. Synoptic situation for the dates written in the bottom-left on each figure.

As for the October case, we have investigated how flow dependency affects mean values on the whole domain to see if the new method changes the mean state of the model. Similar conclusions apply as in the October case, except that now there is a very small effect (~0.01 °C) on temperature as well (not shown).

Looking at the precipitation, which is now convective, more systematic differences can be observed. Figure 13 shows 6-h total precipitation from ensemble median for SPP, SPP_FD2 and their difference (SPP_FD2 – SPP). We see that red is the dominant color on the plot meaning that FD_SPP2 gives more precipitation which agrees with the fact that SPP_FD2 is a little bit more humid and gives a little bit more precipitation (Figure 11). However, the structure of precipitation field remains very similar. Closer inspection of hourly precipitation fields confirms that difference is almost always in precipitation quantity and that the spatial distribution as well as existence of isolated convective cell is almost unaffected.

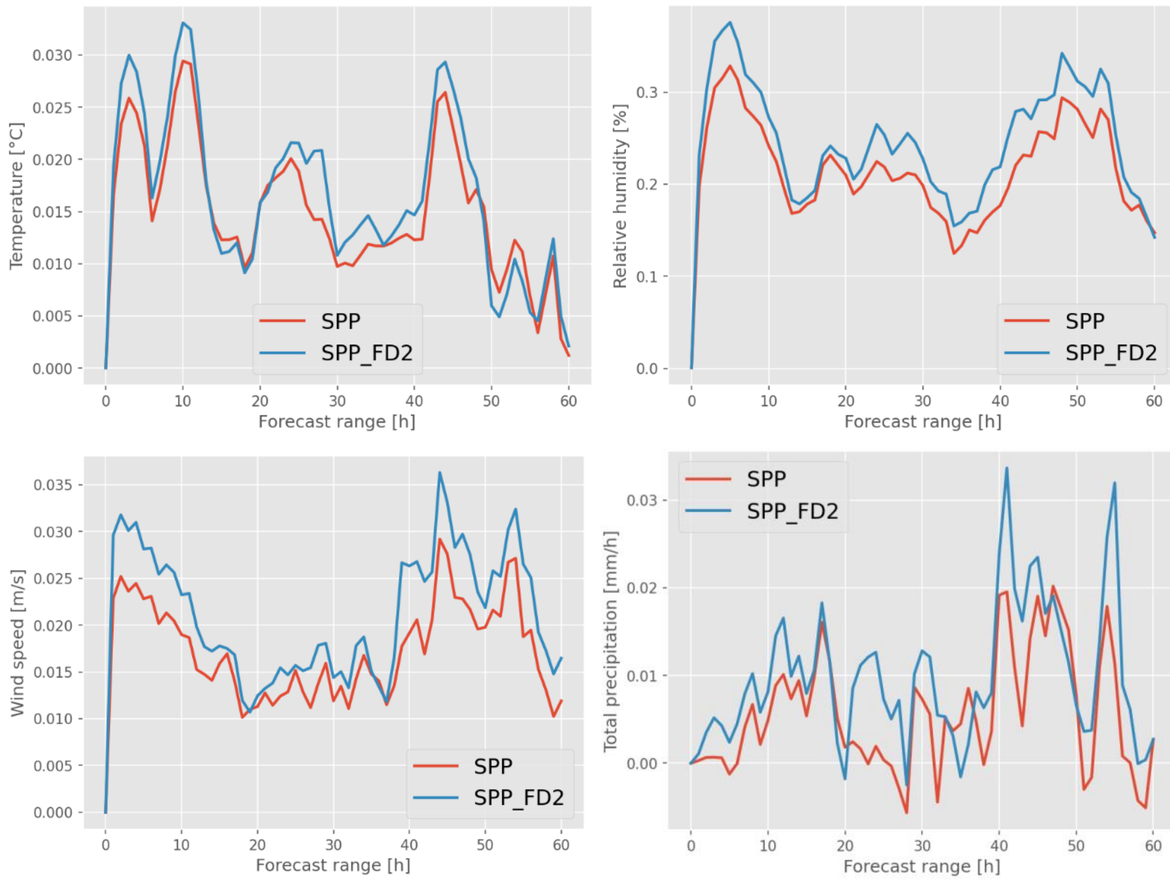

Figure 11. The same as Figure 3, but for the August case.

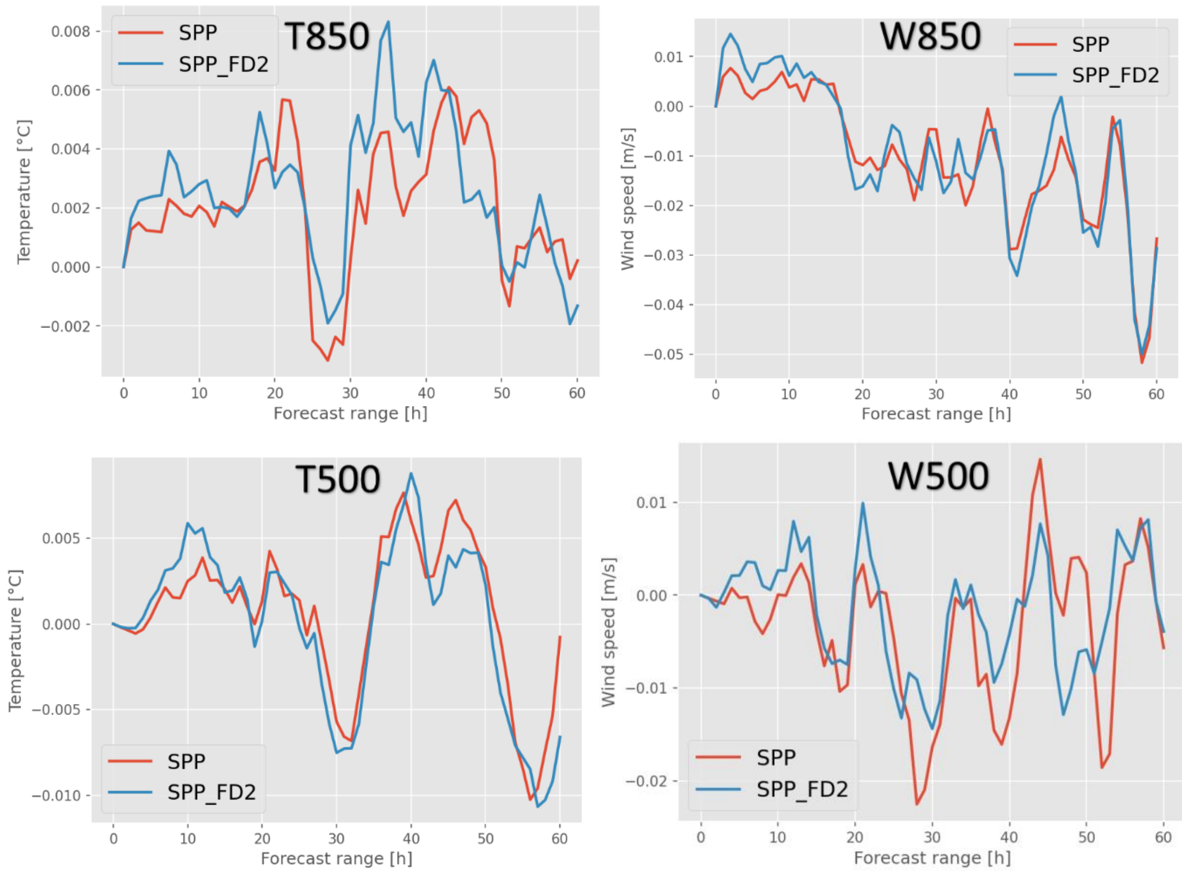

Figure 12. As Figure 4 but for August case.

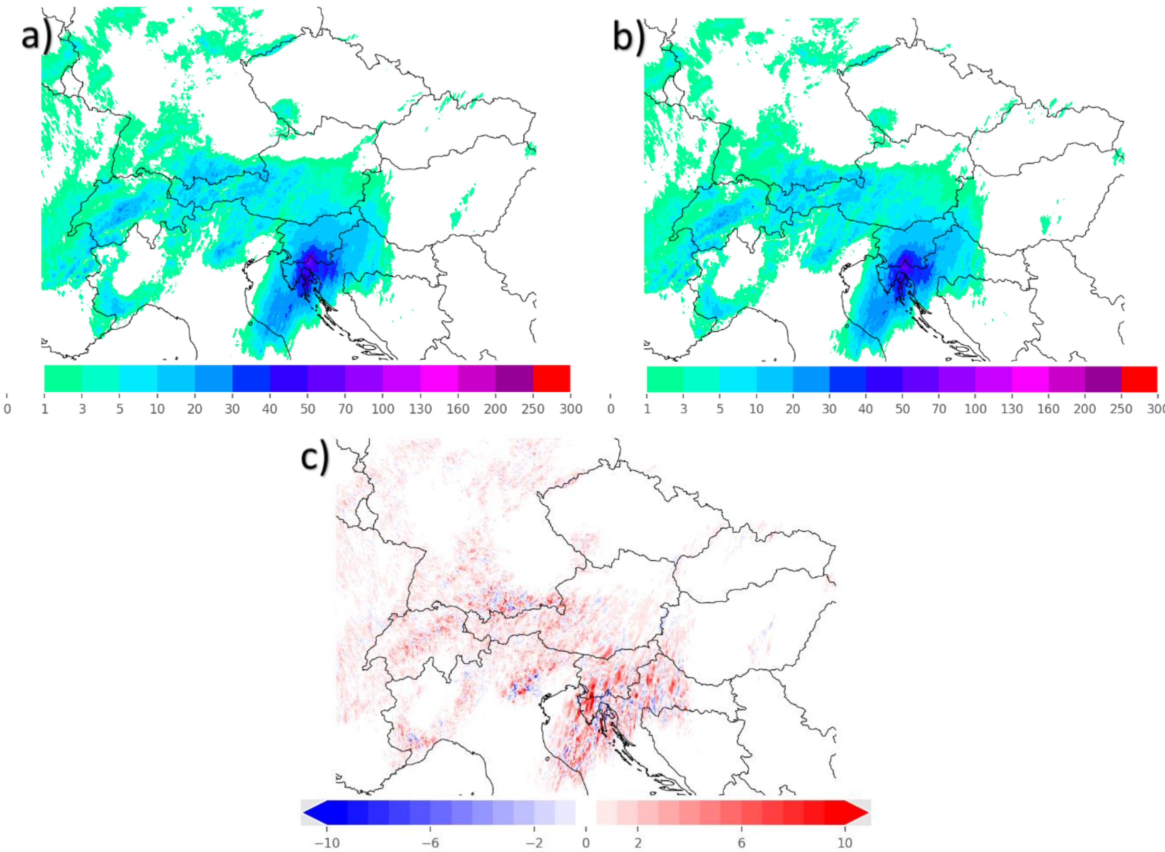

Figure 13. Total 6-h precipitation (lead times: 13-18 h) from ensemble median for a) SPP, b) SPP_FD2 and c) their difference (SPP_FD2 - SPP). Integration has been started on 4. 8. 2023. at 00 UTC.

Conclusion and future work

A new model perturbation method was introduced – flow dependent SPP. The method was implemented to C-LAEF for all perturbed parameters (switch LSPP_FLOWD in the namelist). The method behaves expectedly, meaning that in the targeted regions perturbations are amplified and the differences "move" with the weather. Spread is increased for all variables and for almost all lead times. Impact is smaller for higher vertical levels as we move away from the ground, but this is expected. Effect on precipitation forecast over Austria in large-scale flow forcing case (October) is almost negligible but is more noticeable in a summer convection case. Still, differences remain visible mostly in precipitation quantity and not in spatial distribution. Verification performed on October case reveal that impact of flow dependency is slightly positive on ensemble spread and less negative on RMSE meaning that RMSE and spread ratio is improved. Overall, the new method acts in the desired direction.

Next, a more comprehensive verification is needed. This includes a long-term evaluation and a few additional case studies (both winter and summer) with atmosphere and surface assimilation. We propose to do this evaluation after the final C-LAEF configuration (domain size, resolution, etc.) has been chosen. In addition, some problems in C-LAEF were diagnosed (see Appendix for details) and they need addressing.

Appendix

During this work two bugs were found and fixed in the SPP code. The first one was related to the perturbation of VSIGQSAT parameter from microphysics. The parameter was perturbed and then never used again. Instead, the old constant value was used. The second one was related to a memory leak. The problem was first spotted in *ifs.stat* file where heap memory usage kept increasing during the whole model integration. This only happened when SPP was turned on and the resulting memory usage was about 3 times higher than non-SPP run at the end of 60 h integration. The problem was eventually traced down to a allocation and deallocation inconsistency in *stepo.F90* routine.

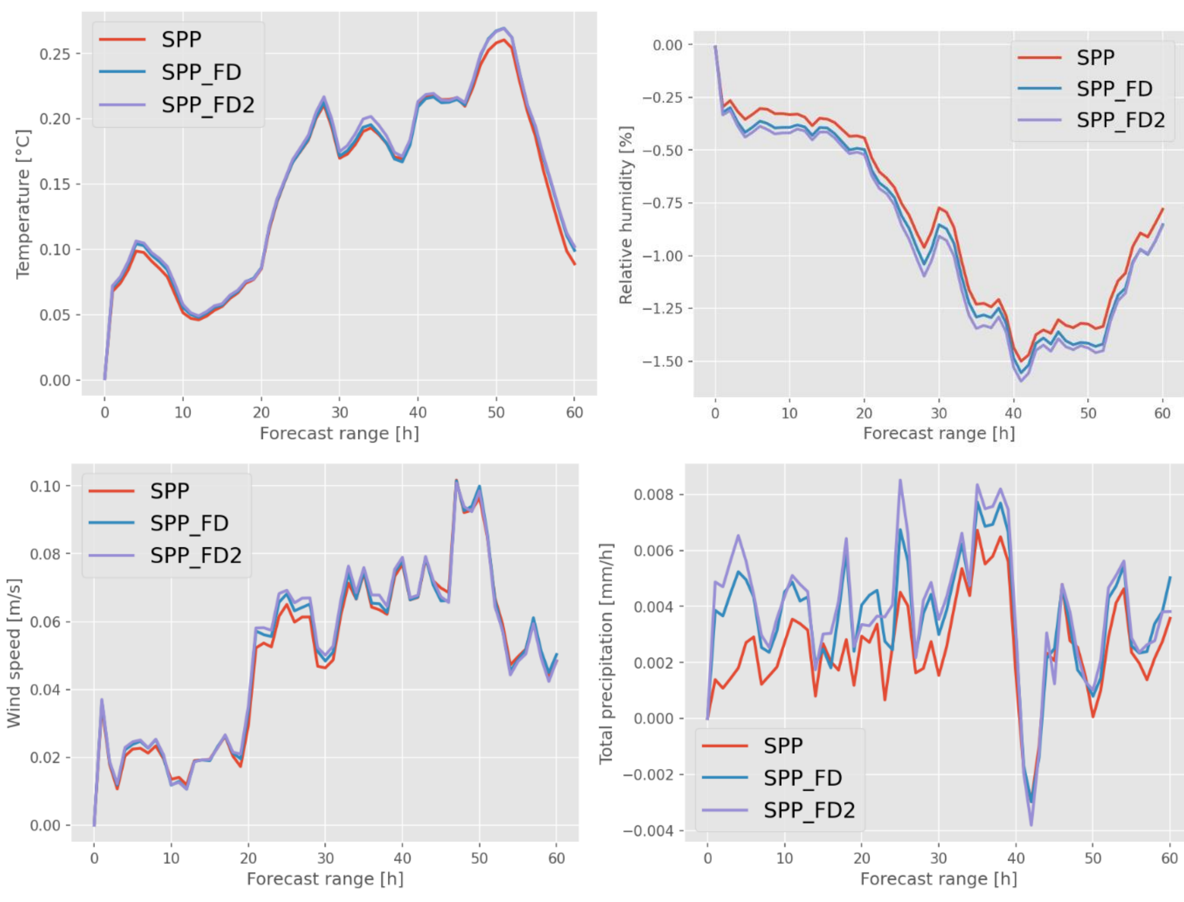

Figure A1. Domain averaged values (October case) for three SPP experiments with respect to NO_SPP for 4 different variables as written on y-axis.

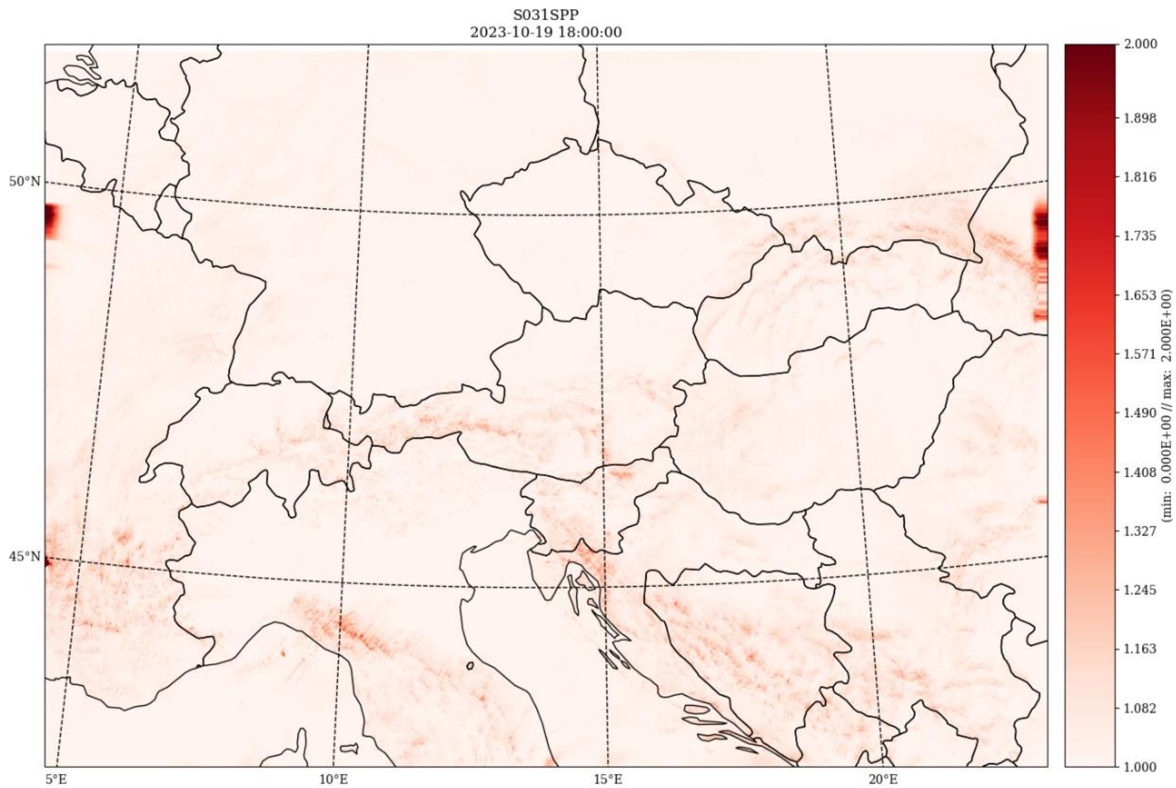

Figure A2. An example of anomalous TKE-weights value in the I-zone. This then results in weights being too small in the rest of the domain.

Additionally, we would like to mention a few noted problems in C-LAEF 1k configuration. "Trajectory underground" problem is experienced too often which indicates some instabilities in the model dynamics. Additional work should be invested into optimizing namelist parameters for SPP because that may help lower the negative impact on RMSE. TKE field has some unrealistically high values in intermediate zone (Figure A2) which then negatively impacts its usage as weights, because then TKE_{max} doesn't adequately represent values inside the domain. Although, this problem can be alleviated by increasing factor N in (3).

References

Ollinaho, P., Lock, S.-J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozzo, A., Forbes, R.M., Haiden, T., Hogan, R.J. and Sandu, I. (2017). Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble. *Q.J.R. Meteorol. Soc.*, **143**: 408-422. https://doi.org/10.1002/qj.2931

Wastl, C., Wang, Y., Atencia, A., and Wittmann, C. (2019). A Hybrid Stochastically Perturbed Parametrization Scheme in a Convection-Permitting Ensemble. *Monthly Weather Review* **147**, 6, 2217-2230. https://doi.org/10.1175/MWR-D-18-0415.1