Report on stay at ZAMG

28/09/2015 – 06/11/2015

Tests of possible SPPT developments

Stochastically Perturbed Parameterized Tendencies (SPPT) is a scheme which has been operationally and successfully used in global IFS model by ECMWF (Buizza et al., 1999, Palmer et al., 2009). In the previous years there was a growing interest around model error representation also in limited area ensemble systems, especially in convection-permitting ensembles. That was a motivation inside ALADIN community to implement the scheme in the limited area version of ARPEGE-IFS code which was done by Francois Bouttier, Météo France, and tested in an AROME-EPS framework (Bouttier et al., 2012).

The author of this paper has also done several tests with AROME-EPS on 2.5km horizontal resolution for the Hungarian domain, without noticing the sufficient impact of the scheme (Szintai et al. 2015). Last year on a similar LACE stay the author implemented some ALADIN code modification which made SPPT available also in case of using ALARO physics package (Szűcs 2014). The recent tests were done mainly with ALARO physics on 8km horizontal resolution because of the lower computational costs in comparison with AROME. However it can be expected that tests were quite "scheme-oriented" so the same conclusions would be valid in case of using AROME. Tests were focusing on three bigger topics and these are detailed in the present document.

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

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

Christoph Wittmann (ZAMG)
christoph.wittmann@zamg.ac.at

In present paper first some general problems are described around the functionality of spectral random pattern generator in LAM versions. Then the impact of the modification of vertical tapering function is examined. Finally the "dimensional extension" of SPPT is described as a new idea of the author. Two possible methods are presented for this purpose and both of them are tested with different settings.
10 cases were collected from the spring and summer of 2015 when there was strong convective activity over Hungary. All the presented verification results are calculated from 36-hours-long model runs started at 18UTC on the day before the above-mentioned convective cases. Near-surface verification was calculated against Hungarian SYNOP observations while high-atmospheric verification was calculated against ECMWF analysis. The list of the examined runs: 20150427 20150505 20150513 20150530 20150614 20150707 20150725 20150801 20150816 20150818.

1. Introduction about the Stochastically Perturbed Parameterized Tendencies (SPPT) scheme

This section introduces SPPT scheme very briefly after Buizza et al. 1999 and Palmer et al. 2009. The purpose is to give a very short overview about the field and make reference equations for the following sections where developments are detailed. First let me present the partial differential equation system of the atmosphere on a very schematic way:

$$\frac{dx}{dt} = F(x;t)$$
$$x(t=0) = x_0 \tag{1}$$

In Eq. 1 *x(t)* vector contains the state variables (e.g. temperature, humidity, wind components) describing the atmospheric state in every grid point. ***F*** denotes the forecast model and x_0 is the corresponding initial condition (IC) of the equation. While SPPT is responsible for the representation of the model uncertainty we can focus on the perturbation of F instead of defining initial condition perturbations added to x_0. Such IC perturbations are downscaled from the global model (PEARP) in the hereafter represented results.
Model state at time *T* can be described as the time integral of Eq. 1:

$$x(T) = \int_{t=0}^{T} F(x;t)\,dt = \int_{t=0}^{T} \left(A(x;t) + P(x;t)\right) dt \tag{2}$$

In Eq. 2 ***F*** is separated to two parts: the explicitly-handled non-parameterized (***A***) and the small-scale parameterized (***P***) processes. ***A*** is described by the model dynamics and ***P*** by the model physics and usually the second one is thought to be the most uncertain part of the numerical models. This is the reason why the model error representation methods which are used currently in NWP community usually focuses on perturbing ***P*** (e.g. multi-physics method, stochastic physics methods, parameter perturbation methods).
According to the above-mentioned reasons in case of the *j*-th perturbed member of an ensemble system Eq. 2 can be modified as follows:

$$x_j(T) = \int_{t=0}^{T} F'(x_j;t)\,dt = \int_{t=0}^{T} \left(A(x_j;t) + P'(x_j;t)\right) dt \tag{3}$$

In Eq. 3 ***P′*** represents the perturbed contribution (tendency) of the parameterized processes while ***A*** stays untouched. In case of using SPPT for a given gridpoint of the model the perturbed tendency of the prognostic variables (temperature(*T*), humidity(*H*), wind components(*U;V*)) can be defined as follows:

$$P'_{X_j}(\lambda;\phi;h;t) = \langle 1+\alpha(h)*r_j(\lambda;\phi;t)\rangle * P_{X_j}(\lambda;\phi;h;t)$$
$$X \in \{T;H;U;V\} \tag{4}$$

In Eq. 4 *r* is random number picked from a Gaussian distribution which has 0 mean and small standard deviation. Although the values of *r* are not constant horizontally and in time but they are not independent either. In ALADIN model family (similar to IFS) the so-called spectral pattern generator is responsible for generating these values. Section 2 will focus on the theoretical details and on the actual performance of this pattern generator.

In Eq. 4 the same *r* is used in every gridpoint but it is modified with a height-dependent function (*α*) level-by-level. This function is called as tapering function and Section 3 describes some tests around its possible modifications.

In Eq. 4 the same r is used in every gridpoint for all the prognostic variables. Section 4 proposes two possibility to change this practice and perturb the tendencies of different variables with different random numbers.

2. Random numbers of the spectral pattern generator

If the SPPT scheme is applied in spectral models, then r_j fields can be generated in spectral space and then transformed to grid point space where the actual parametrization computations are performed. Therefore r_j is described by spherical harmonics in a spectral global model (Palmer et al. 2009) and by bi-Fourier functions in a spectral limited area model (Bouttier et al. 2012). The r_j field is evolved by a so-called spectral pattern generator where its spectral coefficients $(r_j')_{mn}$ are described by a first order auto-regressive (AR(1)) process which ensures the temporal correlation.

$$(r_j')_{mn}(t+\Delta t)=\varphi(r')_{mn}(t)+\sigma\mu_{mn}(t)$$
$$\varphi=\exp(-\Delta t/\tau) \qquad (5)$$

In every timestep new $(r_j')_{mn}$ values are calculated from an AR(1) process described by Eq. 5. φ denotes the one-timestep correlation set by τ decorrelation-timescale. μ values are independent random numbers picked from a Gaussian distribution with 0 mean, 1 variance and bounded into the [-2; 2] interval. These values are multiplied by the σ parameter which is responsible for the size of the perturbations and it is (most commonly and in every references) set to 0.5. In the spectral pattern generator the so-called space correlation length (L) can control the "smoothness" of r_j fields. Following the theory of SPPT references spatial and temporal correlation are set according to the characteristic scale of the errors in the atmospheric processes which is represented by the scheme. In practice it means that time correlation (τ) varies between hours and spatial correlation (L) between hundreds of kilometres depending on the model resolution.

Although the spectral pattern generator was originally designed to work on global model it was accepted as working well in LAM model during previous tests (Szűcs 2014, Belluš 2014). It was admitted that horizontal correlation length (L) does not deserve its original meaning in LAM framework but it was empirically tuned and set to a more or less reasonable value. (It meant that ~10 times bigger values have to be set to get the similar-looking pattern than in global case.) Finally the LAM random patterns were plotted (Szűcs 2014, Belluš 2014) and thought to be qualitatively similar than in global models (Palmer et al. 2009).

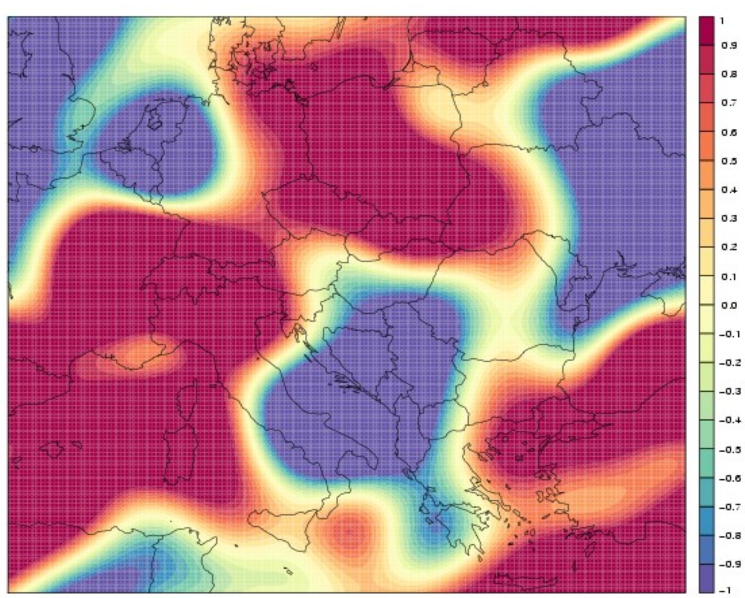

Fig.1: Spectral pattern in ALADIN model if L=500

But after a second look it can be realized that it reaches its minimum and maximum values on bigger areas than it is expected (Fig.1). If the gridpoint values of the random pattern are printed and plotted on a histogram, this feature looks even more problematic. With the original parameter choice of SPPT: $\sigma=0.5$ and the values of r are bounded into the [-1; 1] interval (which means that in ALADIN code clipping ratio=2.0). With the above described settings it can be expected that ~2.3% of the the random field is set to be 0 and ~2.3% of it is set to be 2 (Fig.2).

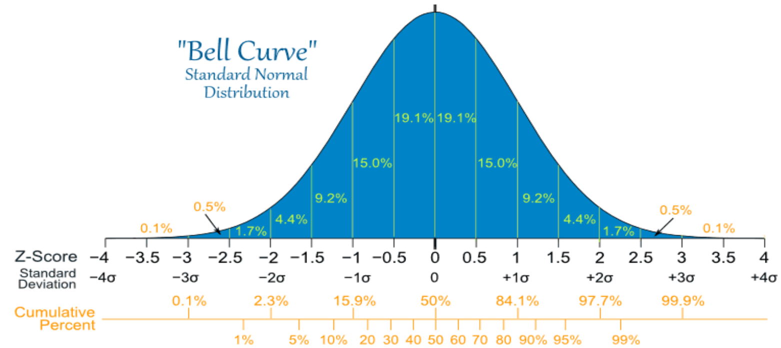

Fig.2: The density function of the Gaussian distribution (illustration from here: http://www.is-math.com/)

The histogram of the real random numbers looks definitely different (Fig.3). Of course the visualization of the histogram can be a bit tricky if it is compared with a density function but for Fig.3 ~160000 gridpoint values were printed and ~25% of them had the value of -1 and same ratio had the value of 1. It means that in quarter of the gridpoints physics was simply switched-off while in another quarter the parametrized tendencies were simply doubled. This feature sounds quite drastic and it is not in accordance with the theoretical design of the scheme.

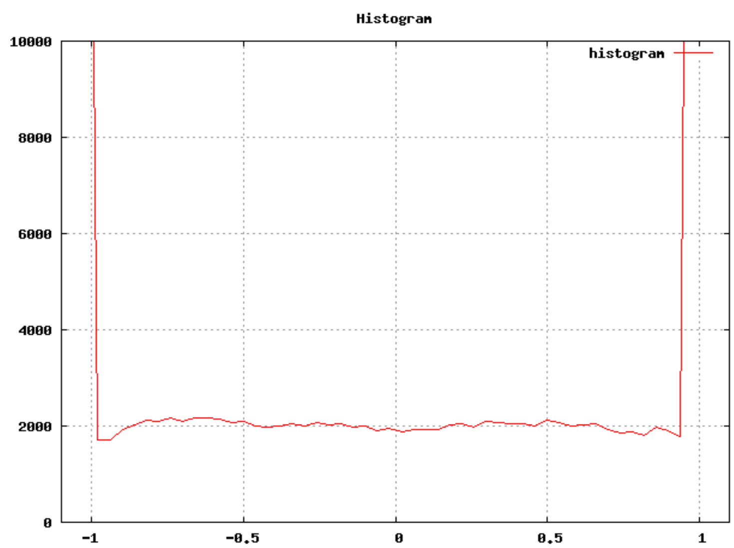

Fig.3: Histogram of random number values of the spectral pattern generator (**σ=0.5; clipping ratio=2.0**)

These results indicate that not only the correlation length can be misleading in the LAM representation of the SPPT, but also the standard deviation. It was tried to find empirically some settings of the scheme which can result more reasonable random gridpoint values. It was found that if σ=0.25 and clipping ratio=4.0 then the look of the histogram (Fig.4) and the ratio of the clipped values are closer to that was originally expected.

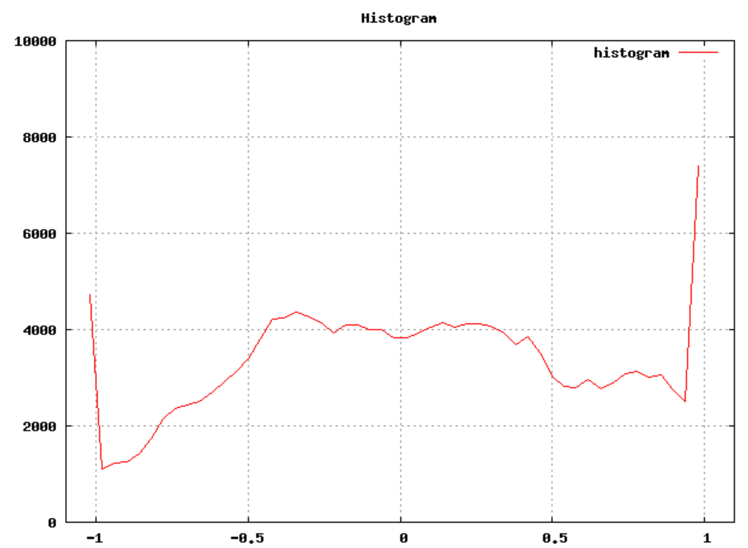

Fig.4: Histogram of random number values of the spectral pattern generator (**σ=0.25; clipping ratio=4.0**)

The above-mentioned two settings and a third one (σ=0.2 and clipping ratio=5.0) were tested on 10 cases of spring and summer 2015 (see more details in the introduction). Here let me focus on the negative impact of SPPT on BIAS if the “original” settings were used and how it is fixed if the two types of modified settings were used (Fig.5a-c).

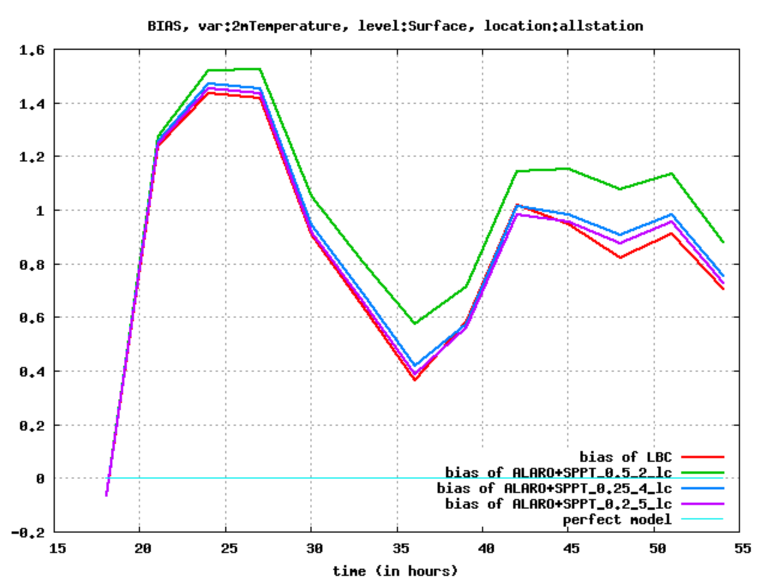

Fig.5a: BIAS of the 2meter temperature. Red line shows the performance of LAMEPS when

perturbations are only from the global model (PEARP). Green line belongs to the “original” settings of SPPT (***σ*=0.5; clipping ratio=2.0**). Blue line is the one which is closer to the theoretically optimal settings (***σ*=0.25; clipping ratio=4.0**). Purple belongs to the even more moderate settings (***σ*=0.2 and clipping ratio=5.0**).

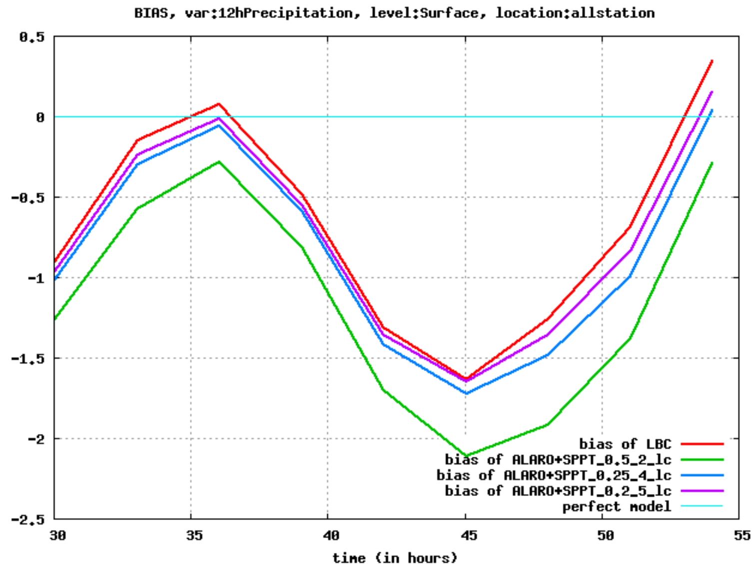

Fig.5b: Same as a) but for the 12-hours precipitation.

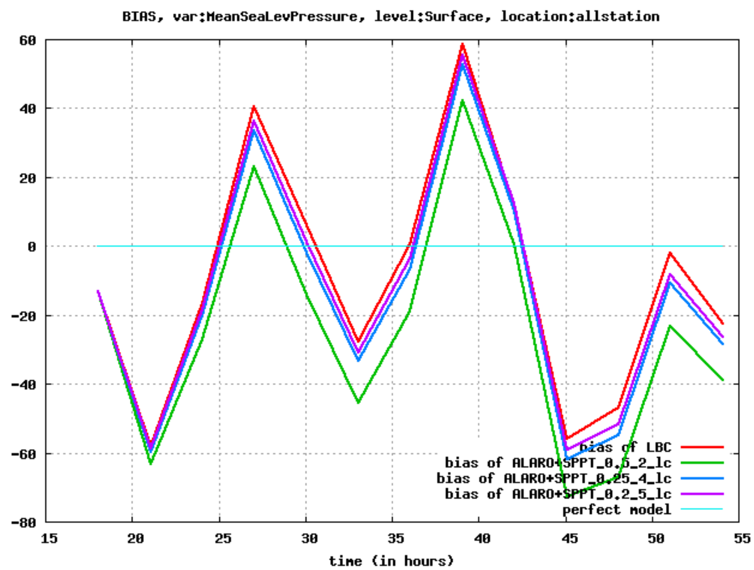

Fig.5c: Same as a) but for the mean sea level pressure.

One can argue that “original” settings are more drastic and they can increase the spread which is an important positive feature. Although it can be true but we can not neglect that it caused massive model degradation and a strong impact on BIAS which is against the theory of stochastic schemes. (They do not want to modify the average behave of the model.)

3. Impact of the vertical tapering function

In Eq. 4 the so-called vertical tapering function *α(h)* appears. In ALADIN code representation of SPPT this function is activated as default. It means that perturbations are multiplied with a height dependent value varying between 0 and 1 (Fig.6).

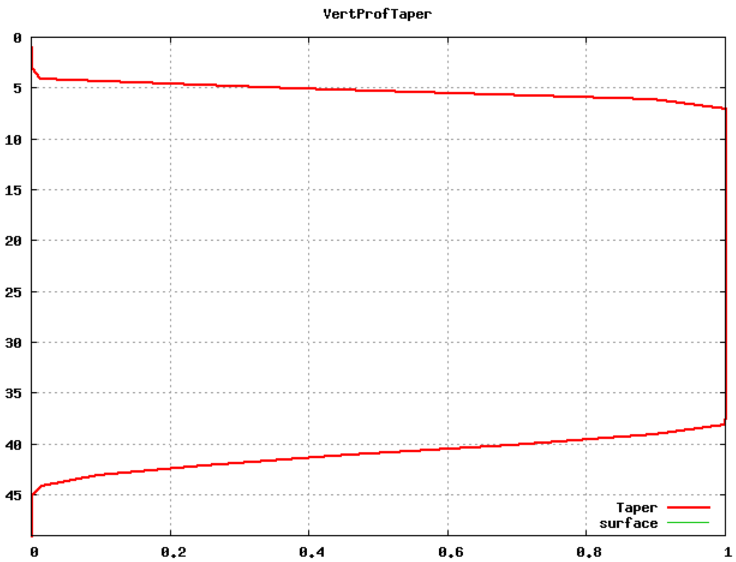

Fig.6: The value of the tapering function on different ALARO model levels

It was reported by ECMWF as a necessary solution to deserve the delicate balance near the surface and at the top of the atmosphere and to avoid model crashes. However it was not really clear what type of problems can appear in ALARO and AROME if tapering is switched-off.

Investigating around this question first tapering was disabled (α=1 as a constant) for all the tendencies and ALARO tests were launched. With such settings 10 perturbed members run on 10 days for 36 hours without model crashes. However error outputs were full of "SMILAG TRAJECTORY UNDERGROUND" warnings. Additionally verification results showed an enormous increase on error of wind speed (Fig.7a). While other verification looked acceptable it was decided to create a new pack where tapering can be activated separately (and its values can be flexibly set) for the tendencies of the different variables. With such a pack it was possible to run a test where tapering was applied (it deserved its shape shown by Fig.6) only for the two components of wind.

The results showed that such a modification can neglect the near-surface problems of wind speed and additionally increase the ensemble spread (Fig.7a-b) especially near the surface. That meant motivation for some further tests to have an idea what happens if tapering is not simply enabled or disabled but the shape of the function is carefully modified (e.g. bigger transition layer between 0 and 1). Before such an investigation AROME tests were also lunched to see if model instabilities appears in case of different model physics or not. Unfortunately 7 have crashed from 11 model runs which indicates that something is very different and unstable in AROME cases. These crashes need further investigation which was over the limitation of my LACE stay.

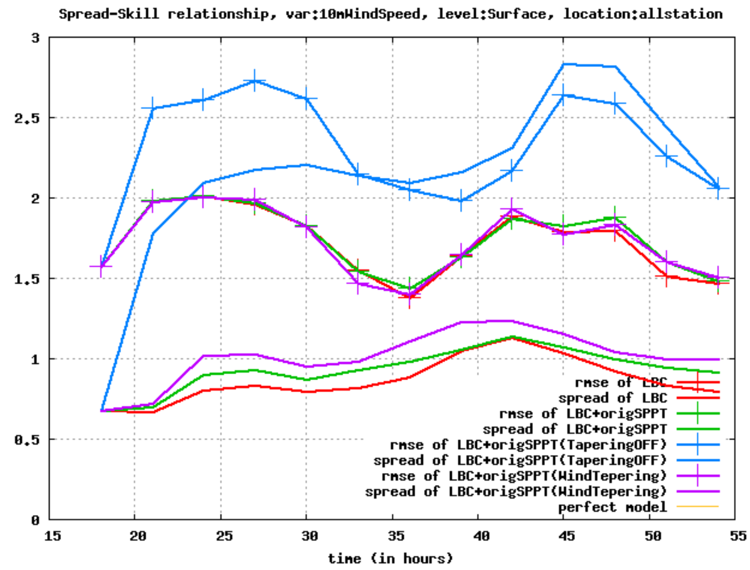

Fig.7a: Root mean square error of the ensemble mean and the spread of the ensemble for the 10meter wind speed. Red line shows the performance of LAMEPS when perturbations are only from the global model (PEARP). Green line belongs to the “original” settings of SPPT (tapering function was active for all the tendencies). Blue line belongs to the tests when tapering function was generally disabled (α=1) Purple shows the performance of LAMEPS if tapering is applied only for wind components.

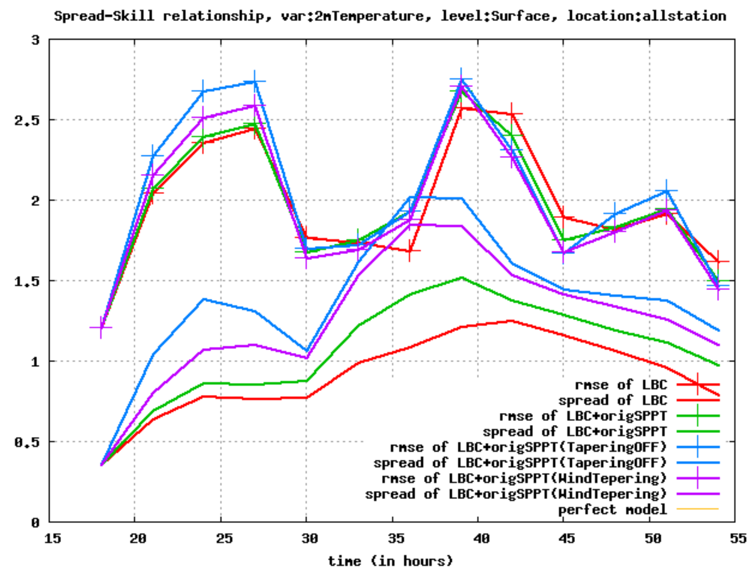

Fig.7b: Same as a) but for 2meter temperature

4. Dimensional extension of SPPT

How it was described by Eq. 4 the same random number is used for all the tendencies in a given column. Regarding to for instance the two wind components on a given level it means that both of them are multiplied with the same number. It means that the length of the wind tendency vector is changed but the direction is not. This 2 dimensional feature can be extended to all the 4 dimension of the space of the perturbed prognostic variables. Parameterizations define a 4 dimensional tendency vector which length is perturbed by SPPT but its direction is untouched. It can be noted that this perturbation is quite large (see section 2.) but again, it is done only in the original direction of tendencies.
There was an idea that various random numbers can be used for the 4 prognostic variables in every gridpoint. This idea can be described with a slight modification (X appears as a bottom index of r) of Eq. 4:

$$P'_{X_j}(\lambda;\phi;h;t)=\langle 1+\alpha(h)*r_{X_j}(\lambda;\phi;t)\rangle * P_{X_j}(\lambda;\phi;h;t)$$
$$X\in\{T;H;U;V\}$$
(6)

The most simple realization of this idea is the generation of four separated random patterns and their separated application on the four perturbed tendencies. This method will be referred as 4DSPPT in this section.
Such a realization can ensure bigger variability because not only the size of the tendency vector but also its direction is perturbed. It can be positive if it ensures bigger ensemble spread but this method does not trust the original direction of changes which is parameterized by model physics and which thought to be more likely than other directions. A proposed solution of this problem can be if the original direction is deserved as the most reasonable one and perturbed stronger than other directions. Other directions can be defined as orthogonal ones and perturbed with smaller amplitude. In this proposal also four independent patterns are generated which ensures four independent random numbers in every gridpoint (r_1,r_2,r_3,r_4). Four different linear combination of these patterns can give the four random numbers which are really used to perturb tendencies. It can be described in the simplified version of Eq. 6 as follows:

$$P_T'=\langle 1+\alpha r_T\rangle * P_T \qquad r_T=r_1(0;\sigma_1)+r_2(0;\sigma_2)+r_3(0;\sigma_3)+r_4(0;\sigma_4)$$
$$P_H'=\langle 1+\alpha r_H\rangle * P_H \qquad r_H=r_1(0;\sigma_1)-r_2(0;\sigma_2)+r_3(0;\sigma_3)+r_4(0;\sigma_4)$$
$$P_U'=\langle 1+\alpha r_U\rangle * P_U \qquad r_U=r_1(0;\sigma_1)+r_2(0;\sigma_2)-r_3(0;\sigma_3)+r_4(0;\sigma_4)$$
$$P_V'=\langle 1+\alpha r_V\rangle * P_V \qquad r_V=r_1(0;\sigma_1)+r_2(0;\sigma_2)+r_3(0;\sigma_3)-r_4(0;\sigma_4)$$
(7)

In Eq. 7 r_1, r_2, r_3, r_4 are still Gaussian distributed random numbers with 0 mean but their standard deviation is not necessarily equal. If in Eq. 7 $\sigma_1>\sigma_2=\sigma_3=\sigma_4$ then the original direction of the tendency will be perturbed more than the orthogonal directions. This version will be referred as 4DdiagSPPT in this section.
To explain its effect clearer let me visualize it just in 2 dimension (like these dimensions were wind components again). In 2 dimension r can be thought as a vector which two coordinates are random number. In original SPPT these coordinates are equal so if the endpoints of many vectors are visualized they will be alongside a line like the red dots on Fig.8a. In 4DSPPT these coordinates are independent, but they are random numbers from the same distribution so after visualization they will form a circle like green points on Fig.8a. In 4DdiagSPPT coordinates are linear combination of two random numbers. One of them is from a distribution where the standard deviation is double than for the other

one. This way the visualized dots can form an ellipse which major axis will be on the line of the red dots.

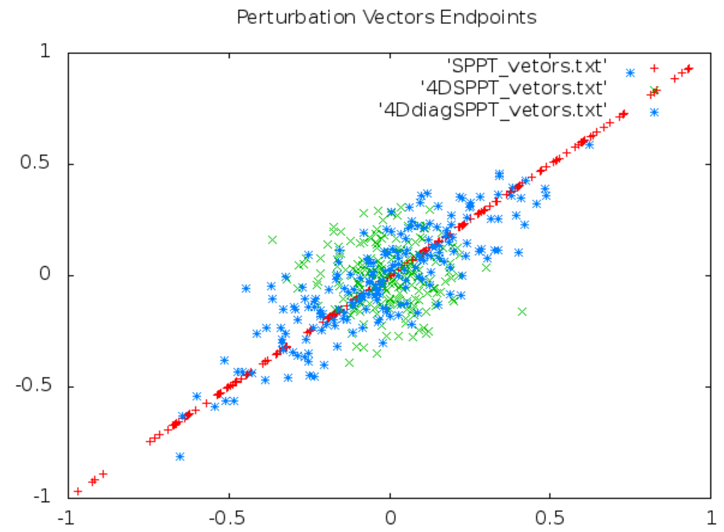

Fig.8a: Endpoints of the perturbation vectors. The coordinates of red points are always equal like in case of SPPT. The coordinates of green dots are independently picked from the same distribution like in case of 4DSPPT. The coordinates of blue points are the linear combination of two random numbers (one of them are picked from a distribution where σ is double) like in case of 4DdiagSPPT.

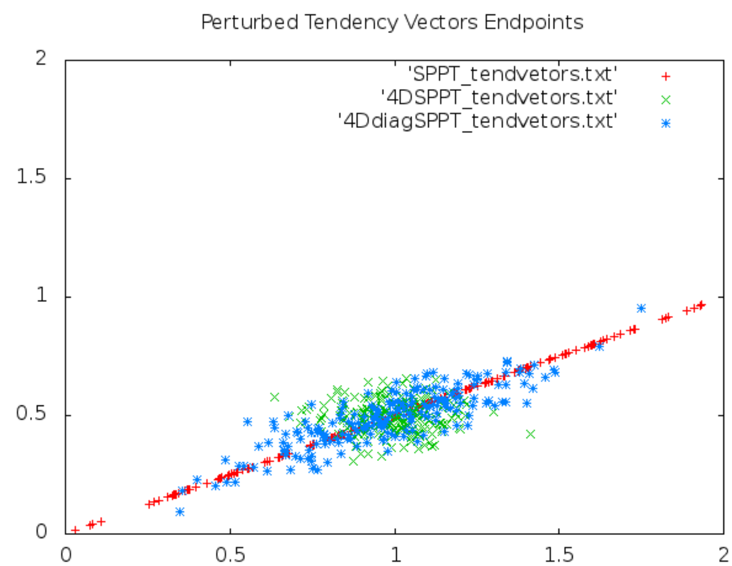

Fig.8b: Same as a) but vectors are multiplied with a presumed tendency vector which is [1;0.5].

If the vector of original tendency is let say [1;0.5] then the endpoints of the perturbed tendency vectors can be visualized as it is on Fig.8b.

Several tests were run to find the best setting of the above-described 4DSPPT and 4DdiagSPPT methods. Not all of them can be presented here but one of each to compare the two new methods with each other and with the original version. The choice of the settings can be justified as follows:

- Original SPPT was used with σ=0.2 and clipping ratio=5.0 to avoid the large number of bounded values on the edge of the distribution.
- 4DSPPT was used with the same settings for the better comparison (The four independent random numbers were picked from the same distribution).
- 4DdiagSPPT is a bit more complicated because σ of the final random number is calculated from the σ values of four independent random numbers ($\sigma_1=2\sigma_2=2\sigma_3=2\sigma_4$). Clipping is applied for this final value. For the better comparison it was decided not to deserve the σ value of the other two methods but the volume of the object where perturbation endpoints can take place in case of 4DSPPT. (In the previous 2-dimensional example this logic would be to keep the area of the circle where green dots are and make it equal with the area of the ellipse where blue dots are. In 3D we could use words like: volume, sphere, ellipsoid. In 4D the author do not know the valid expressions.)

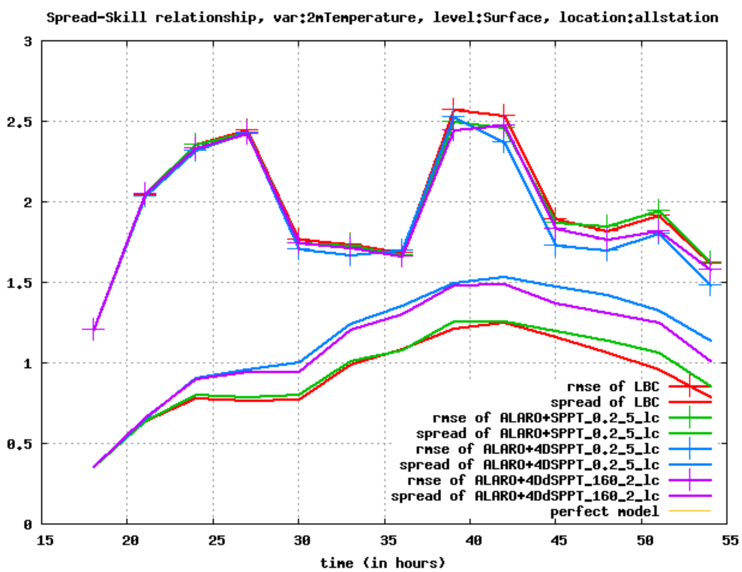

Fig.9a: RMSE of the ensemble mean and the **SPREAD** of the ensemble for the **2meter temperature**. Red line shows the performance of LAMEPS when perturbations are only from the global model (PEARP). Green line belongs to the "original" settings of SPPT (σ=0.2 and clipping ratio=5.0). Blue line shows the performance of the 4DSPPT which uses the same distribution than SPPT. Purple line belongs to 4DdiagSPPT which perturbations take place in an object with the same volume than in case of 4DSPPT.

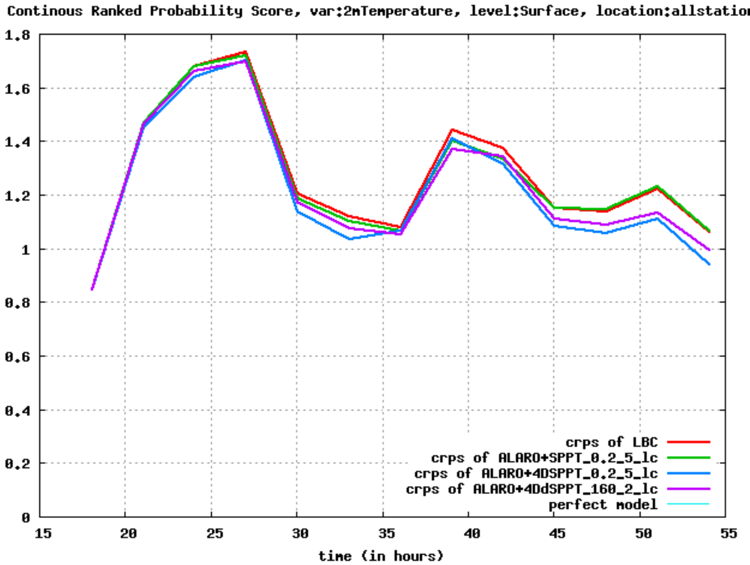

Fig.9b: CRPS for the **2meter temperature**. The meaning of the lines are the same than on a)

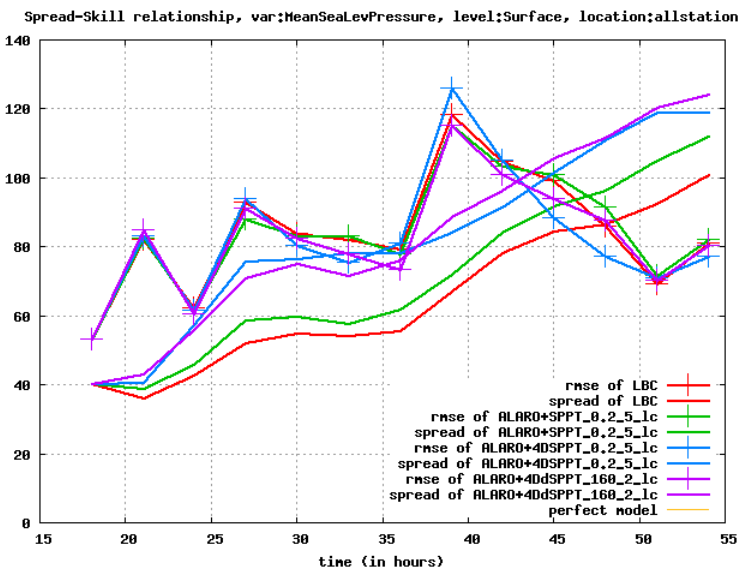

Fig.9c: RMSE of the ensemble mean and the **SPREAD** of the ensemble for the **mean sea level pressure**. The meaning of the lines are the same than on a)

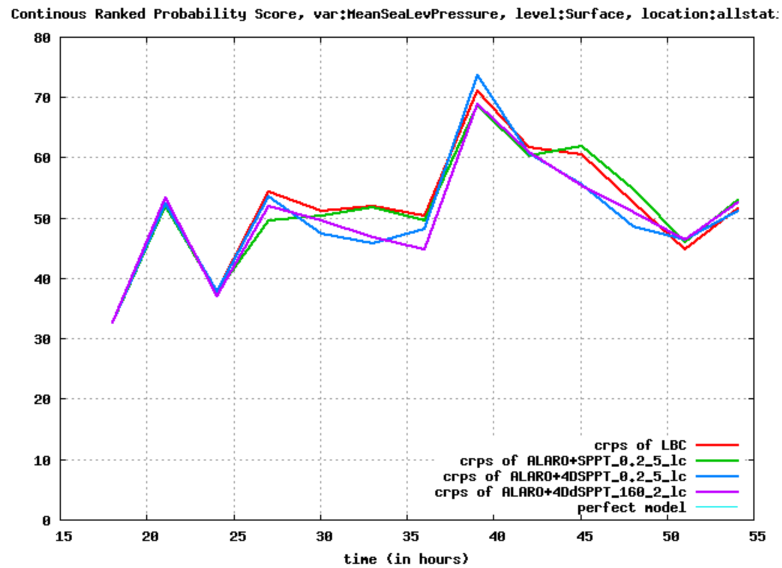

Fig.9d: CRPS for the **mean sea level pressure**. The meaning of the lines are the same than on a)

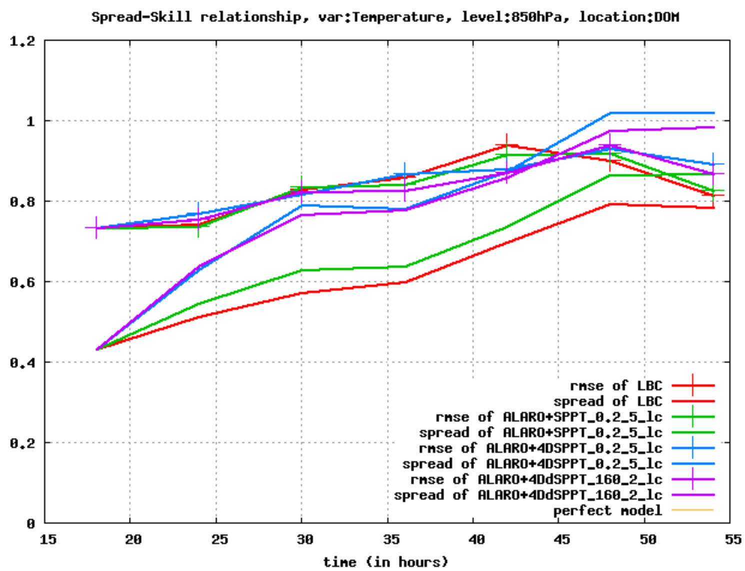

Fig.9e: RMSE of the ensemble mean and the **SPREAD** of the ensemble for the **850hPa temperature**. The meaning of the lines are the same than on a)

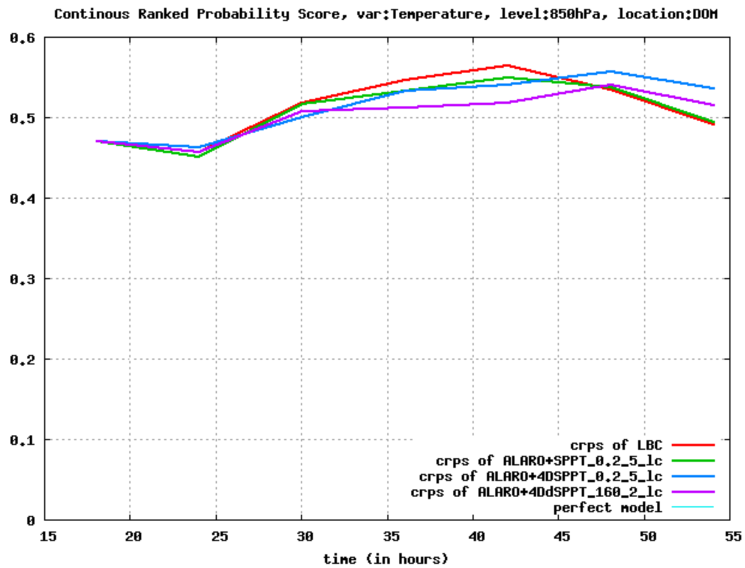

Fig.9f: CRPS for the **850hPa temperature**. The meaning of the lines are the same than on a)

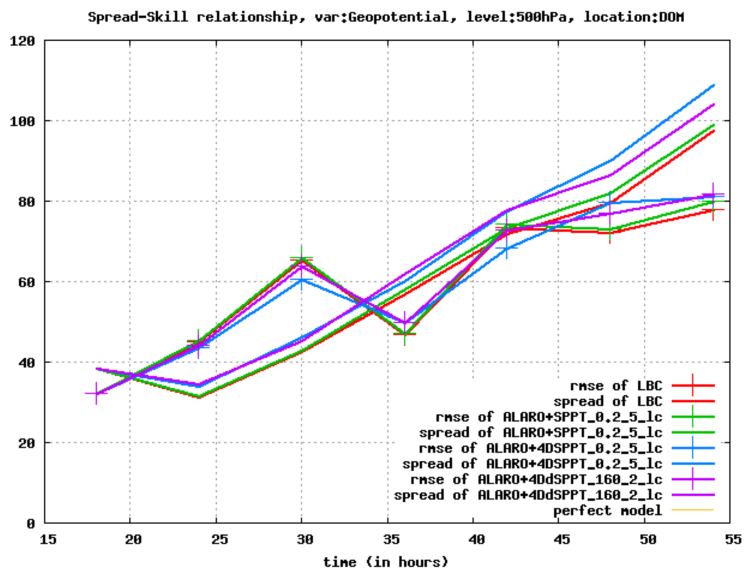

Fig.9g: RMSE of the ensemble mean and the **SPREAD** of the ensemble for the **500hPa geopotential**. The meaning of the lines are the same than on a)

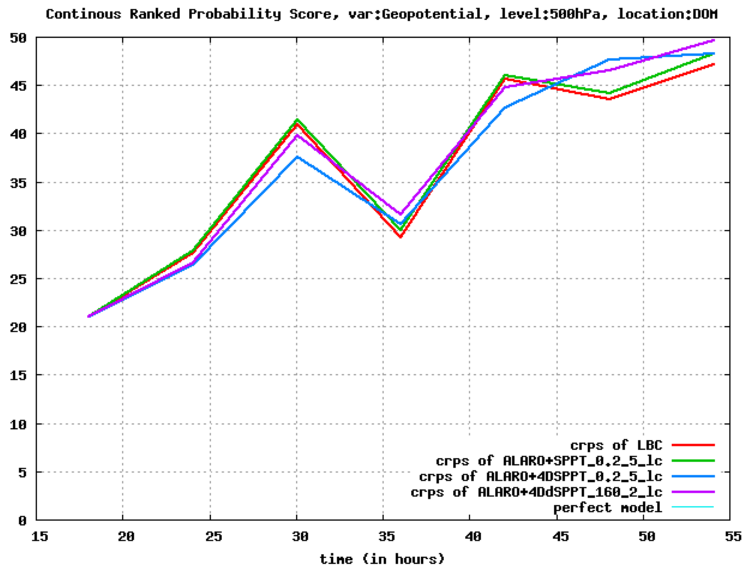

Fig.9h: CRPS for the **500hPa geopotential**. The meaning of the lines are the same than on a)

Fig.9a-h show some results from the comparison of these methods. Spread-skill relationship and CRPS values are presented. The spread of the 4DSPPT and 4DdiagSPPT versions are always bigger than the spread ensured by the original SPPT or the simple downscaling. In many cases new methods can also improve the RMSE of the ensemble mean. CRPS (smaller the better is) figures can give an overview about the quality of the ensembles which use the two new methods: Differences are not too big but usually the impact is positive especially in case of 4DdiagSPPT. It looks essential to run similar tests for longer period (e.g. for 60 hour like in the operational LAMEPS of Hungarian Met Service) because the impact of model perturbations can affect more on longer period.
It can be underlined that in present paper all the results are based on fixed value of τ and L. The search of their ideal tuning can cause further complications and modify the conclusion of the present results.

5. Future plans and possibilities

- In section 2 it was underlined that spectral pattern generator does not work properly in LAM. It means that its settings (standard deviation, horizontal correlation length) can not give back the expected results (and time correlation has not been really investigated yet). The best solution would be a detailed revision of the spectral pattern generator and make it work properly on LAM. This challenge was beyond my possibilities during my stay but it would be an important task in the future.
- In section 3 the importance of vertical tapering was confirmed. The switch-off of the tapering for temperature and humidity can give bigger spread near the surface in case of ALARO physics but causes problems in case of AROME physics. These problems should be further investigated in the future.
- In section 4 two modified versions of SPPT was proposed: 4DSPPT and 4DdiagSPPT. These methods were able to ensure bigger spread than the original SPPT without obvious model degradation. The impact of the schemes should be examined on a bit longer model runs (e.g. 60 hours) and on longer test periods.
- All the presented developments were realized under cycle 38. They should be phased under cycle 40 soon.

6. Acknowledgement

Finally I would like to say thank you for the whole modeling group of ZAMG, for their supportive attitude. They helped me in all my scientific, technical and personal troubles which made my work really smooth during my stay. I am also grateful to Martin Belluš for the inspiring discussions and for the patience what he showed as LACE Area Leader. I am appropriate to Francois Bouttier, who supervises ECMWF's special project, called 'spfrbout', which SBU was used by me during my stay.

7. References

Belluš, M., 2014: Stochastically perturbed physics tendencies of surface fields in ALADIN-LAEF system, Report on stay at ZAMG 12/05 - 20/06/2014, Vienna, Austria

Bouttier, F., Vié, B., Nuissier, O., Raynaud, L., 2012: Impact of Stochastic Physics in a Convection-Permitting Ensemble. Mon. Wea. Rev., **140**, 3706–3721.

Buizza, R., Miller, M., Palmer, T.N., 1999, Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, Quart. J. Roy. Meteorol. Soc., **125**, 2887–2908.

Palmer, T., Buizza, R., Doblas-Reyes, F., Jung, T., Leutbecher, M., Shutts, G., Steinheimer, M., Weisheimer, A., 2009: Stochastic parametrization and model uncertainty. Tech. Rep., ECMWF Tech. Memo. **598**, 42 pp. [Available online at http://www.ecmwf.int/publications/.]

Szintai, B., Szűcs, M., Randriamampianina, R., Kullmann, L., 2015: Application of the AROME non-hydrostatic model at the Hungarian Meteorological Service: physical parameterizations and ensemble forecasting. Időjárás **119**, 241-266

Szűcs, M., 2014: Stochastically Perturbed Parameterized Tendencies in ALARO and AROME , Report on stay at ZAMG 06/10/2014 - 31/10/2014, Vienna, Austria