

SPECIAL PROJECT FINAL REPORT

Project Title:	Boundary layer model errors in the AROME ensemble prediction system
Computer Project Account:	spfrbout
Start Year - End Year :	2015 - 2017
Principal Investigator(s)	François Bouttier
Affiliation/Address:	CNRM, Météo-France 42 Av Coriolis F-31057 Toulouse francois.bouttier@meteo.fr
Other Researchers (Name/Affiliation):	Mihály Szűcs, Hungarian Meteorological Service (OMSZ)

Summary of project objectives

The objective is to research new parametrisations for the representation of low-level model error in ensemble prediction systems. The main intention is to test several strategies for stochastically perturbing parameters in the Bougeault-Lacarrère TKE-based vertical mixing scheme.

Summary of problems encountered

n/a

Experience with the Special Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

no problem

Summary of results

see following pages

List of publications/reports from the project with complete references

Mihály Szűcs, 2015: [Tests of possible SPPT developments](#) , Report on stay at ZAMG 28/09/2015 - 06/11/2015, Vienna, Austria

Mihály Szűcs, 2016: [Stochastic pattern generators](#) , Report on stay at ZAMG 23/05/2016 - 17/06/2016

Mihály Szűcs, 2017: [Implementation of Stochastic Pattern Generator \(SPG\) in ALADIN code](#), Report on stay at ZAMG, 12/06~21/07, 2017, Vienna, Austria

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

This activity will continue in the Hirlam/Aladin/Lace consortia, in the framework of the HARMONIE, ALARO and Arome-France ensemble prediction systems.

Study and experimentation of low-level model error in convection-resolving ensemble prediction systems

Project report

Mihály Szűcs and François Bouttier
(OMSZ and Météo-France, June 2018)

1. Motivation

The aim is to use the AROME EPS ensemble prediction system with various stochastic perturbation schemes, in order to gain insight and develop improvements of the representation model error. The algorithmic difficulties and physical processes behind variations in ensemble spread and probabilistic scores will be investigated in order to design an optimal perturbation scheme.

Most of the work has been performed by the Hungarian Met Service (Mihaly Szucs in particular), which is registered in this project.

2. Test of variations of the SPPT scheme

As a first step, improvements to the Arome SPPT stochastic physics scheme have been tested.

2a. Clipping of tendency perturbations

Current preoperational SPPT settings have a relatively high spread, which means that during any Arome-EPS run, SPPT clipping is active at many gridpoints. Active clipping means that physics tendencies are multiplied by zero or two, which seems a bit excessive from a physical standpoint. An SPPT version with reduced spread (standard error is divided by 5) has been tested in Arome-EPS on several cases, which somewhat reduces member forecast errors, but also reduces the ensemble spread, which is undesirable (the spread/skill relationship is degraded).

A reduction of SPPT tendency spread by a factor 5 was tested, which reduced ensemble spread at the expense of a degradation of the spread/skill relationship. It has been concluded that SPPT with the MF settings is rather unrealistic in terms of physical tendencies, but its benefits (in terms of improving the reliability of spread) outweigh this problem in the current system.

2b. Multivariate versions of SPPT

Multivariate versions of SPPT have been tested, in the hope that they can provide higher spread without degrading ensemble skill. 'Multivariate' means that the U,V,T,q variables are perturbed using independent stochastic noise. Ensemble spread is improved by this scheme (compared with a univariate SPPT with an equivalent tuning), particularly in the boundary layer, without an improved spread/skill relationship. Unfortunately, significant biases appear on upper-level fields (e.g. temperature, humidity, wind speed), which suggests that the univariate SPPT excites unphysical error structures.

Some possible improvements to SPPT have been designed and tested notably a 'multivariate elliptical' system by which a fraction of the SPPT perturbations is allowed to be different between the perturbed tendencies (U,V,T,q). In the original ECMWF and operational MF setting, all these variables use the same perturbations. The multivariate elliptical system is found to improve the ensemble spread/skill relationship in the boundary layer, at the expense of creating biases at upper levels. Nevertheless this is regarded as a promising option. It seems to be a good tool for improving spread in the boundary layer, where multivariate relationships between forecast errors might not be as strong as in the upper troposphere.

Operational experience with the default SPPT settings at Météo-France have revealed that SPPT can sometimes crash the model in severely convective situations. This is linked to instabilities in the planetary boundary layer (although SPPT is not directly acting at the lowest model levels). As a quick fix, the reductions to the SPPT perturbation amplitude have been successfully tested, and then activated in the operational system.

2c. Treatment of saturation

A prominent weakness of SPPT in Arome is that it tends to reduce tropospheric humidity, which causes dry biases with respect to observed screen-level humidity and precipitation. (This problem has been reported in other systems such as IFS-ENS). Changes to the treatment of humidity in SPPT have been tested in the hope of alleviating this problem. The

NQSAT_SDT supersaturation check option was originally developed at ECMWF, but not used in Arome for two reasons: (1) there is already a built-in supersaturation check in the Arome physics, which is believed to already play this role, and (2) the NQSAT_SDT code is designed for the IFS treatment of saturation, which is slightly different from Arome. Nevertheless, tests of several supersaturation treatments for Arome involving the NQSAT_SDT code have shown a successful reduction of the dry SPPT bias (available experiments are too short to conclude on the impact on the overall probabilistic performance of the modified ensemble). Clearly, this is not a long-term solution because it is slightly inconsistent with rest of the Arome model setup, but the results show that there is some potential to reduce the dry SPPT bias by a more careful treatment of humidity.

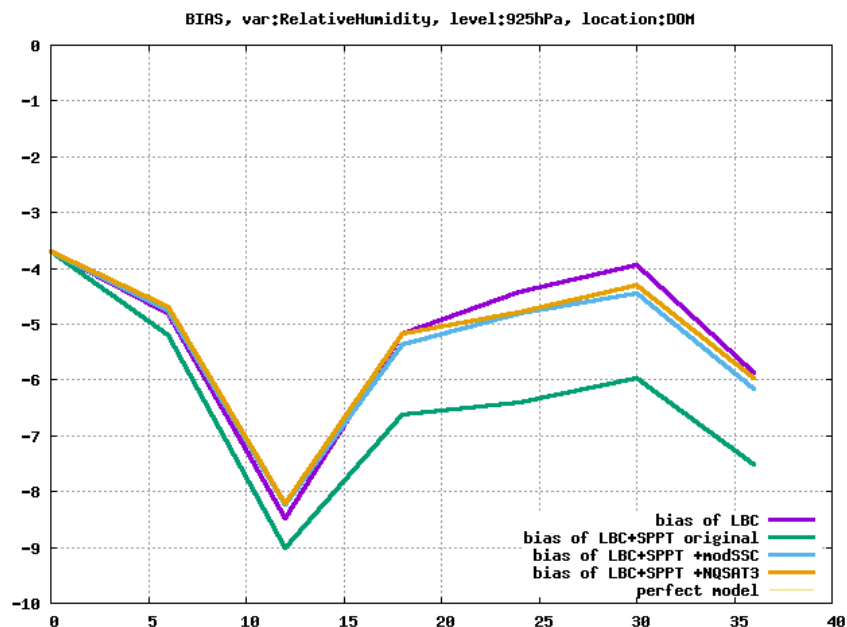


Figure: comparison of 925hPa relative humidity forecast biases averaged over 10 runs, with various supersaturation treatment strategies. The green curve is the setup with the original SPPT, and it has the largest dry bias.

3. Implementation of stochastic pattern generator (SPG) in ALADIN code

3.a. Introduction

The global version of the Stochastically Perturbed Parameterized Tendencies (SPPT) has been successfully used by ECMWF (Buizza et al., 1999) and tuned with a spectral pattern generator during a main revision (Palmer et al., 2009). The limited area version and AROME extension was implemented by François Bouttier from Météo France (Bouttier et al., 2012). A detailed examination and further extension to the ALARO model took place in a framework of a LACE stay, 2014 (Szűcs, 2014). In the following year some possible developments were investigated and tested as problems were also reported (Szűcs, 2015). One of these problems is the unsatisfying behaviour of the random pattern generator in its LAM version, which motivated its revision and start of the implementation of a new Stochastic Pattern Generator (SPG, Tsyrlunikov and Gayfulin, 2016) that was the focus of a LACE stay last year (Szűcs, 2016a). Some details about these investigations were presented in ALADIN-HIRLAM (Szűcs, 2016b), SRNWP-EPS (Szűcs, 2016c) and HIRLAM-EPS (Szűcs, 2016d) workshops. More information about this work can be found in Szűcs (2017).

3. b. Problematic issues with the current spectral pattern generator

Some problems of the current random pattern generator have been detailed in a previous report (Szűcs, 2016a). Now only the main points are being highlighted to give a motivation for the implementation of the new Stochastic Pattern Generator (SPG).

Theoretically the main disadvantage of the current pattern generator is that the same time correlation is applied to all the spatial scales. Since forecast error is connected to atmospheric motions it would be beneficial to have various scales with separated spatial and time correlations for error representation purpose. With the current pattern generator the only possible way to meet this goal is if several random patterns are defined and applied during the same time (Palmer et al., 2009). This kind of solution can handle multiple scales in a discrete way but cannot represent a continuous spectrum of motions.

In practice the LAM version of the current pattern generator is not implemented in a fully satisfactory way: its standard deviation should be controlled from the namelist but in practice it is larger than the specified value. At the same time horizontal correlation is much smaller than its namelist-defined value. This difference is domain-size dependent.

3.c. Description of a limited-area spatio-temporal stochastic pattern generator

The two main requirements for a new stochastic pattern generator are the following:

- when representing the model error at various scales there should be different time correlation values connected to different spatial correlation values. This feature is called as the “proportionality of scales”.
- The new pattern generator should be correctly tunable: the namelist-defined values should be identical (or at least close enough) to the statistical values calculated from the generated fields. The theoretical background of the SPG scheme is well-described by its inventors in an article (Tsyrlunikov and Gayfulin, 2016). The corresponding FORTRAN code can be freely downloaded from github with additional technical documentation:

<https://github.com/cyrulnic/SPG>

In this report the focus is not on the theoretical details but on the technical implementation and on properties which can affect the results and their usage. Here are some interesting features of this SPG:

- as already mentioned, the spatio-temporal covariances should obey the “proportionality of scales” principle: larger (shorter) spatial scales should be associated with larger (shorter) temporal scales.
- the SPG should produce Gaussian univariate pseudo-random fields that are stationary in time and homogeneous and isotropic in space. The authors are also interested in non-Gaussian noise which can be better suited for some meteorological variables.
- it is possible to produce 2D and 3D random fields, as well.
- a 3rd order in time spectral-space based solver is used. It makes it easy to implement in the ALADIN (i.e. Alaro, Arome and Harmonie) code where the current pattern generator is applied in spectral-space as well.
- the SPG used here was tuned with some changes to save a large amount of computational cost without significantly affecting the statistical behaviour of the fields.

A big part of the original SPG code is devoted to the Gaussian noise generation and to FFT transforms. Those algorithms are also available in the ALADIN code, so that it was found easier to implement the rest inside the ALADIN model code, than to generate thousands of fields on file from an external program and read them during model integration.

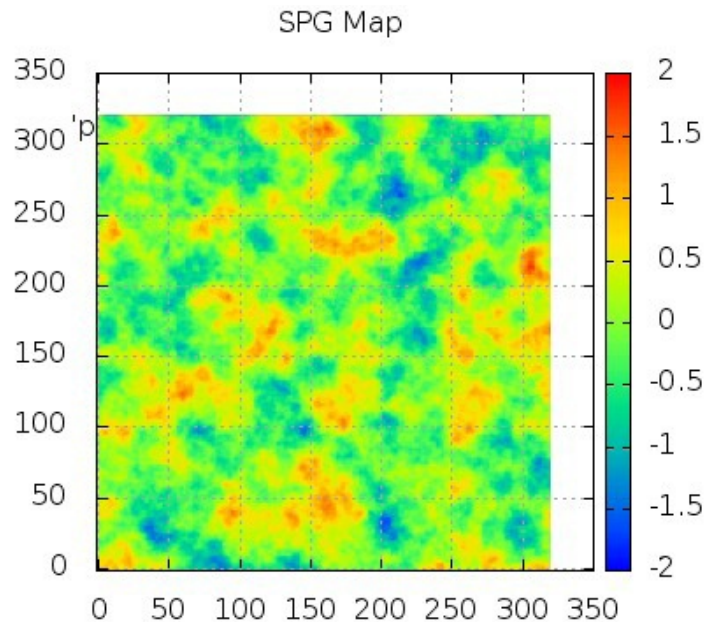


Figure: random field generated by SPG (as an external program)

3.d. Implementation of the SPG into the ALADIN code

In a first test we wished to feed the random fields of the SPG into SPPT as it is done with the current random pattern generator. So the easiest way was to implement everything at the same part of the model where the current pattern generator works. Technically speaking it means that an additional switch can enable the SPG method in the initialization (suspsdt routine) and in the calculations in spectral space (functions of spectral_arp_mod module file). These calculations are called from the very-high-level stepo routine. The above-mentioned switches can make it easy to the user to decide if the current pattern generator or SPG would be applied.

A very tricky part of the external code implementation is the initialization. To set the time and spatial correlation length in the external program two config parameters needed. Parameters L05 and T05 respectively set the value where and when the spatial and time correlation functions are 0.5. Via an iterative process λ and μ values are calculated which are actually used in the SPG equations. During this iterative process, FFT calculations are called several times. Normally in ALADIN code, FFT and IFFT are called during the integration when we switch from spectral to physical-space or back. It did not seem straightforward to use transformation immediately in the setup routine so in practice the following approach was used: First the L05 and T05 are defined in the external program together with the domain and timestep information and then it calculates the correct λ and μ values. This step is necessary only once for a given model configuration and for a given L05 and T05 setting. After that the evaluated λ and μ can be set as a namelist parameter of ALADIN implementation. This is not a really nice solution but needs only short time at the beginning of a test.

The structure of the pattern generation needed some massive reorganization because of the order of the loops. Of course in our case all the eps members are independent model runs so their loop has to come on the highest level. What is even more interesting is that the loops on wavenumbers and on timesteps have to be switched.

An extra problem is that the SPG works with a different timesteps than the NWP model itself. Additionally this timestep is wavenumber dependent and also effected by the tuning which makes the code faster. To handle this problem for every wavenumber there is an extra calculation which defines a number of substeps and their length which is used by the SPG. The solver is 3rd order in time which means that for calculation of a new value, we need the value of the previous three substeps. That means that the storage of the last three substep fields also had to be handled because they are needed to evolve values over model timesteps. We can also note that independent Gaussian-noise is necessary for every substeps which means that the vector size (which is filled by these random values at the beginning of the model timestep) has to be increased in accordance with the wavenumber-dependent substep number.

An additional challenge was that the external SPG program works with rectangular truncation while ALADIN uses elliptic truncation. It needed a careful revision on the total wavenumbers, as well. It has to be noted that in this implementation of the SPG only a 2D version became available. Doing it this way was more simple and in accordance with the way how we (and also ECMWF) currently use random patterns. Of course, a possible direction of developments could be to implement the 3D version, as well.

3. e. Fields and their statistical behaviour

First it can be demonstrated that fields are good-looking and qualitatively better than the ones with the current pattern generator (see previous reports). That means that there are no strange spots filled by -1 or +1 values, the spatial and

time structures look reasonable and the “proportionality of scales” feature is visible. The following figures can be compared with the ones which are in the SPG inventor's publications and documentations. To get the following results Hungarian AROME domain was used. We used standard deviation $\sigma=0.5$, $L05=100\text{km}$ and $T05=1\text{hour}$. If we visualize the x-oriented cross-section of the values the small scale structures can be even better recognizable which is an advantageous feature of SPT in comparison with the current pattern generator.

Of course such a qualitative comparison can not be absolutely satisfactory. The statistical behaviour of the pattern was investigated over 10 runs which length was set to 6 hours. The standard deviation calculated over such a relatively big sample was 0.502 which is close enough to the namelist defined value. Note that with the current pattern generator it is around 1.2 if we omit clipping which can significantly decrease that at the end. As mentioned in previous reports the histogram of random numbers did not have a Gaussian shape with the current pattern generator. With SPG it looks much better from this perspective (not shown).

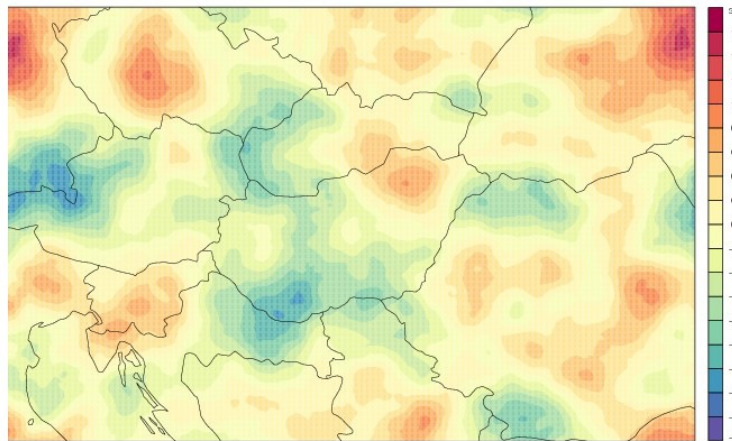


Figure: example of a random field generated by SPG (in ALADIN code implementation) for the Hungarian AROME domain.

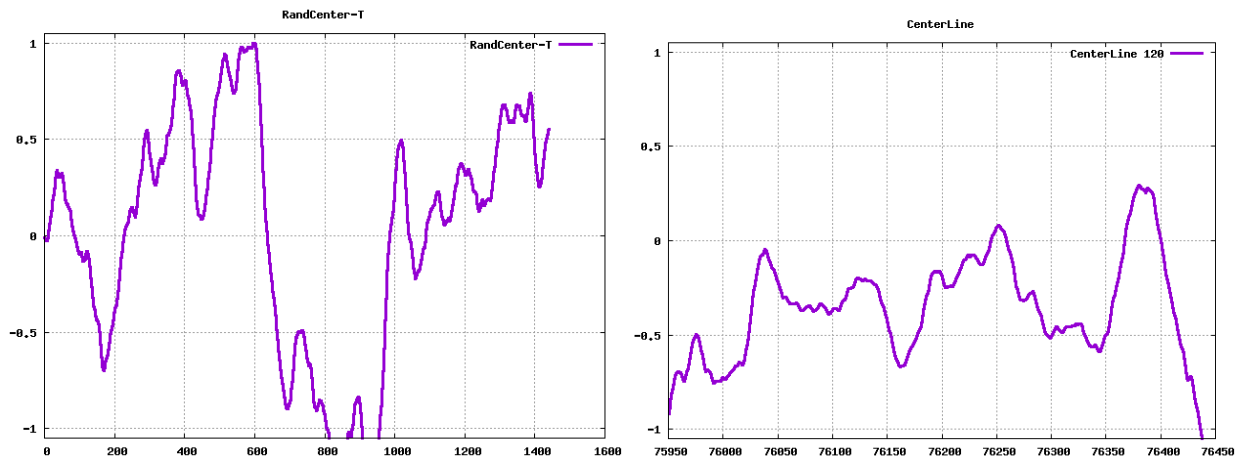


Figure: left: x-oriented cross-section of the random pattern generated by SPG (in ALADIN code implementation). right: time evolution of the random value of a given gridpoint in the center of the domain.

The spatial and temporal correlation functions have been also checked. They give very important clues about the correctness of the implementation. We can compare the functions calculated from the ALADIN code and from the fields made by the external SPG. It is also possible to check if the decreasing correlation functions reach 0.5 value as set by $L05$ and $T05$, or not.

The following figure shows the spatial and temporal correlation functions. It is obvious that the lines belonging to ALADIN and to external program calculations are quite close. The ALADIN version crosses 0.5 level between 107.5 and 110km if function is calculated in direction of x and between 102.5 and 105 in direction of y. These values are getting closer and closer to the theoretical value if the sample size is increased so these differences look acceptable. Note that in the current pattern generator we can get quite similar values if we set 4000km as horizontal correlation

length for Hungarian AROME domain. In such case correlation functions are crossing 0.5 level between 110 and 112.5km in direction x and between 87.5 and 90km in direction y. So SPG looks also better from the aspect of isotropy.

If we examine temporal correlation function the situation looks even better. It reaches 0.5 between 59 and 60 minutes. Note that in the current pattern generator this value is around 4 hours if we set 6 hours as decorrelation time length in the namelist.

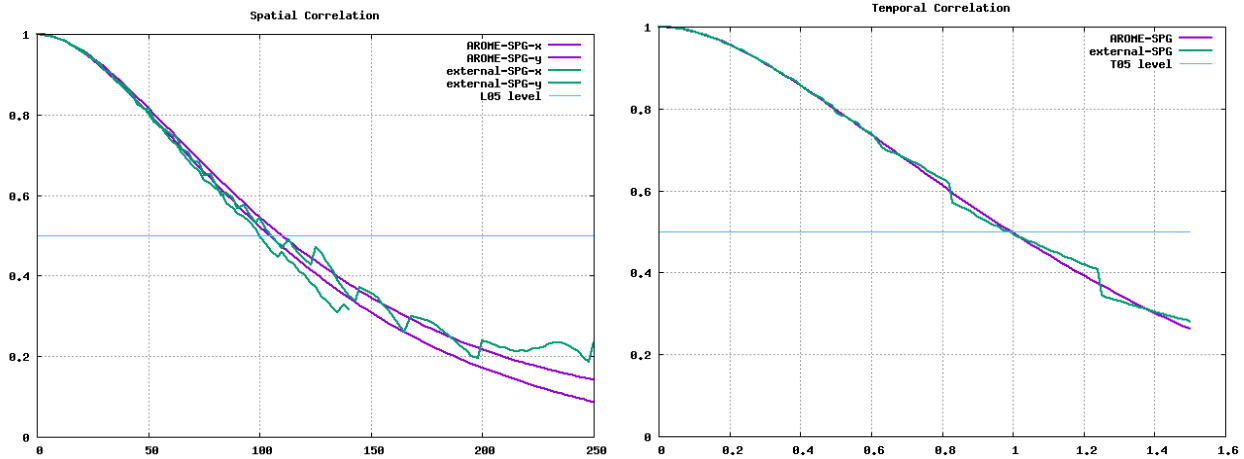


Figure: Spatial (left) and temporal (right) correlation functions calculated from fields of ALADIN code implemented SPG fields(purple) and from external SPG program results (green). The blue line shows the 0.5 level.

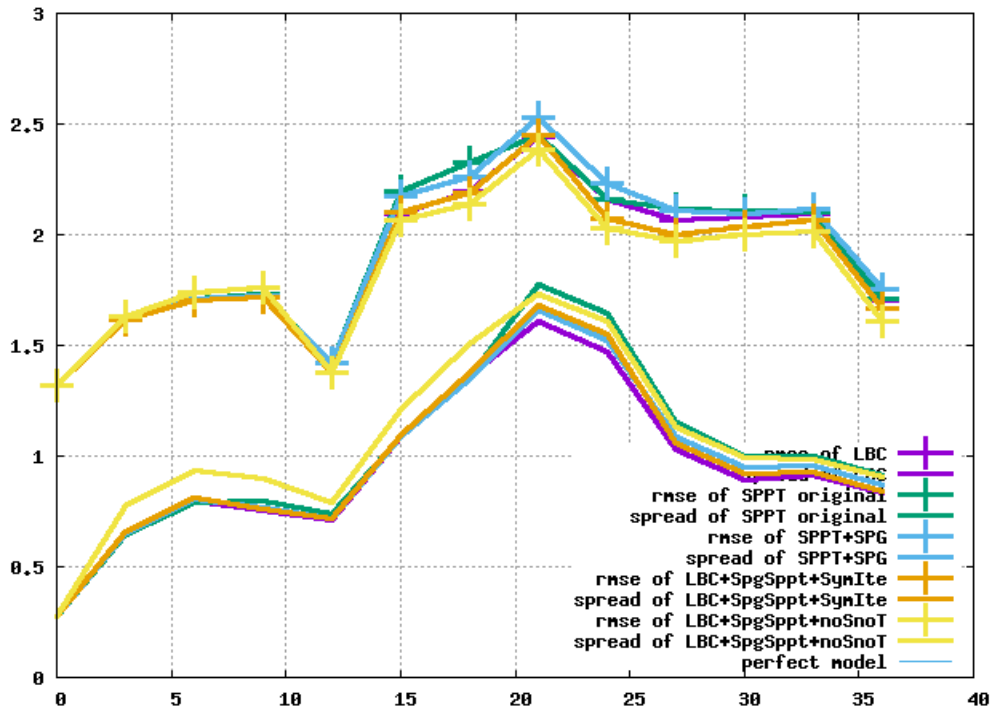
4. Using SPG without additional filtering techniques

In part 2 it was underlined that the tested “supersaturation check” methods are not really consistent with AROME physics and they can easily cause biases (e.g. drying effect with NQSAT_SDT=0) as well as a degradation of model quality which is hardly compensated for by some spread improvement. Moreover the tapering function was also active in all the tests which was found necessary in IFS to decrease perturbations in boundary layer and avoid numerical instabilities. In part 3 it has been shown that a motivation for using SPG was that in the current random pattern generator the effective amplitude of perturbations could not be precisely controlled using the namelist. We can summarize that the existing SPPT system inserts large perturbations and then filter them strictly to avoid model crashes and quality problems.

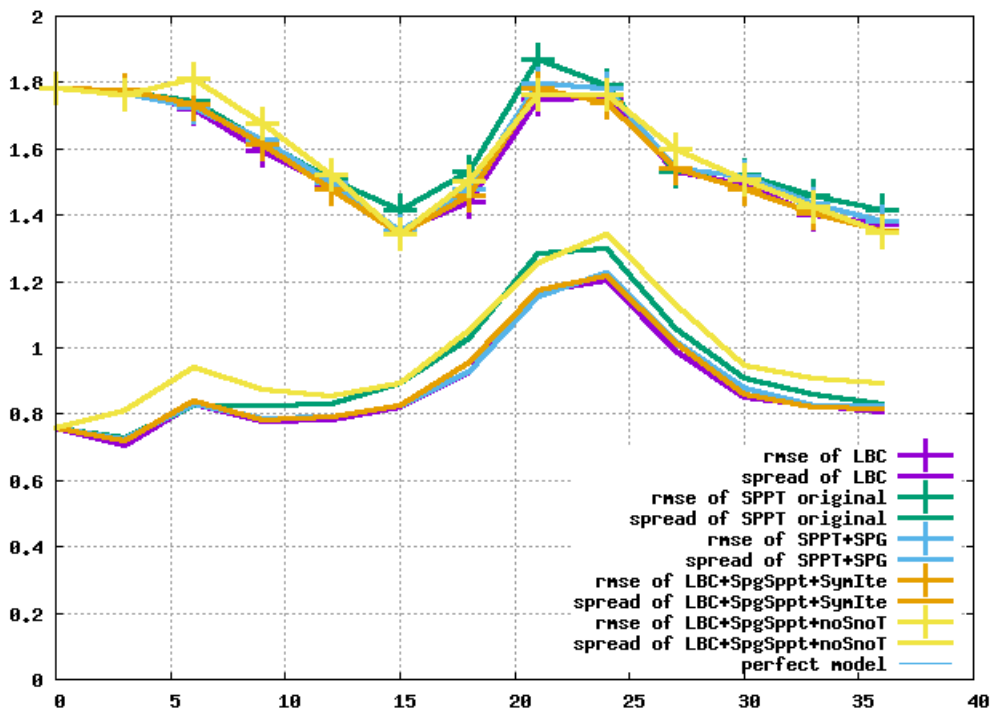
In this part we show the results of an experiment with the following motivation:

- to use good quality random fields with smaller amplitude (SPG with $\sigma=0.4$);
- to switch-off additional filtering methods, namely the tapering function and supersaturation check.

Spread-Skill relationship, var:2nTemperature, level:Surface, location:allstation



Spread-Skill relationship, var:10mWindSpeed, level:Surface, location:allstation



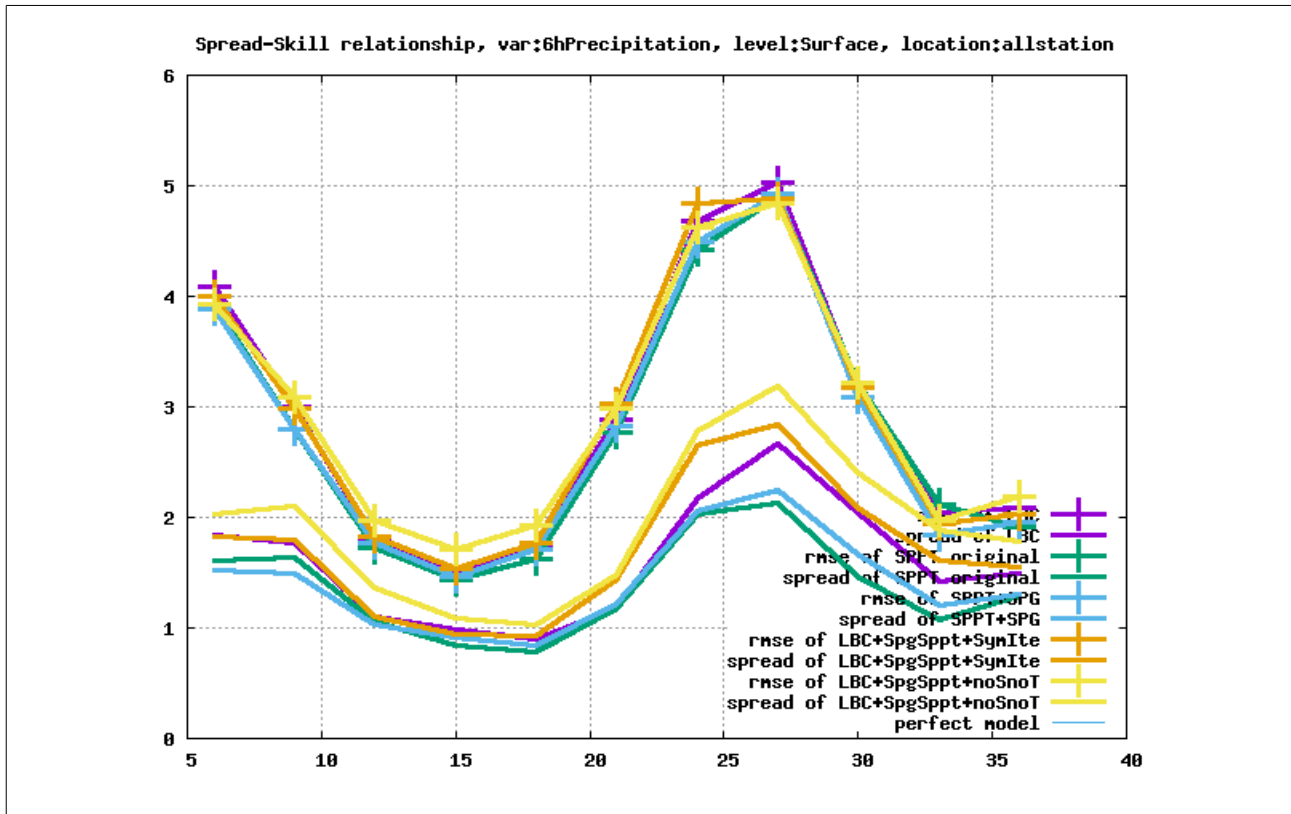


Figure: Spread-skill relationship for 2meter temperature, 10m wind speed and 6-hour precipitation amount. Purple line is the reference when perturbations arrive only via downscaled ICs and LBCs; green line is SPPT with default settings; blue and orange lines are SPG based SPPT with active tapering function and different supersaturation check methods; yellow line is SPG based SPPT without tapering function and supersaturation check.

This development removes the need for artificial filters and vertically inconsistent perturbations. It also reduces model biases. Generally, spread can decrease in higher atmosphere because of the smaller amplitude but this effect is compensated by the activity of the scheme in the boundary layer. It is important to note that during the test period no model crashes were detected.

5. Diagnosing model error in the boundary layer

A fundamental problem of model error representation is to identify the correct sources of errors in ensemble prediction systems. Since many atmospheric processes constantly interact at multiple scales, one expects the effects of model error on ensemble spread to be mixed with (say) deficiencies in the representation of initial condition error, surface conditions or large-scale condition of limited area models. In the free troposphere, at scales where chaotic processes such as cyclogenesis dominate error growth, it is reasonable to hope that any reasonable noise source that will excite chaotic amplification of forecast differences, will lead to a sensible spread in ensemble prediction. This is not so in model parts where physics play an important and flow-dependent part. One expects such problems to manifest themselves as a lack of ensemble spread, when compared with ensemble average error. We have focused on the representation of wind speed in the dry boundary layer, over plains.

The Arome-France ensemble wind speed spread has been compared with the ensemble average error over a large number of cases (every day over two winter months) and observation points (Southwestern Europe). The average spread/skill ratio, which is a measure of reliability, has been stratified as a function of height above ground. The main difficulty was to locate observations of vertical wind profiles with sufficient vertical resolution, extent and geographical coverage to produce robust statistics. Experimentation with instrumented masts, wind farms, doppler radars and lidars, sodars, UHF and VHF wind profilers revealed that they all had data quality and/or availability issues that prevented this type of statistical analysis. The best dataset found were AMDAR/ACARS aircraft reports, which are abundant near European airports. Their main weakness is the lack of data at night time, which may bias the result by hiding model errors specific to the nocturnal boundary layer. In the future it is planned to expand this procedure to Mode-S aircraft data, which is even more abundant. The potential of instrumented masts will also be revisited in specific weather conditions such as fog.

The results so far (see figure below) show a clear impact of boundary layer processes on ensemble reliability: as one would expect, turbulent wind dissipation near the ground means that both ensemble spread and forecast errors decrease sharply at low levels (typically, below 2000m above ground). The key result is that the spread/skill ratio decreases too, from an upper-level ratio of about 0.8 (which is representative of windspeed ensemble reliability in the free

troposphere, and consistent with radiosonde verification) to less than 0.4. Since near-surface wind is rather insensitive to initial conditions beyond a few minutes, and is mainly driven by the same large-scale processes that drive upper-level wind, one concludes that the current model error representation in Arome-EPS (based on SPPT with stochastic multiplication of model physics tendencies) is missing some important error sources that are specific to the surface boundary layer. Previous work by Bouttier et al (2015) on surface ensemble perturbations showed that not much wind spread can be gained by just perturbing surface parameters. The conclusion is that an additional model error source needs to be developed near the surface, and we have constructed an observation-based framework to guide its tuning. This will be the topic of future work. These conclusion are consistent with independent results by e.g. Berner et al 2015 who used a multiphysics approach to treat PBL-specific lack of ensemble spread.

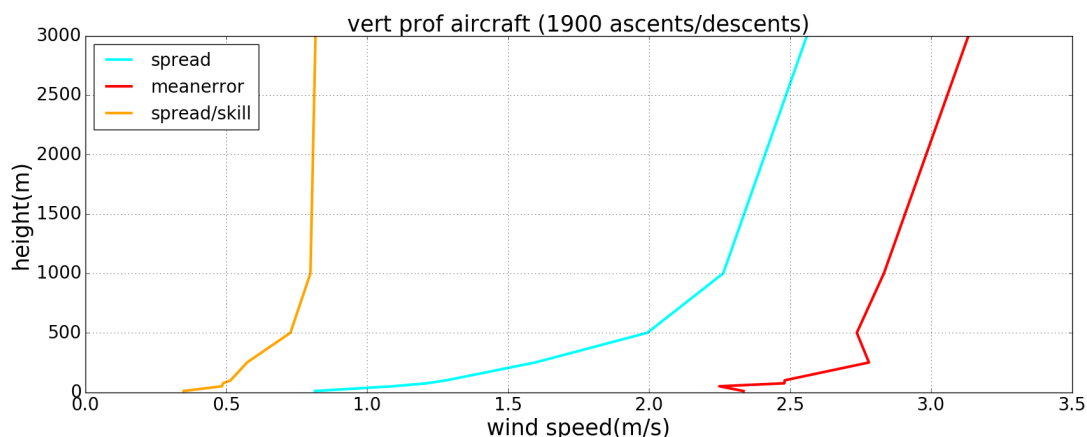


Figure: vertical average profiles of wind spread, average ensemble forecast error and spread-skill ratio, using the Arome-France-EPS system and aircraft observations.

5. Conclusions and future plans

It was found that the current spectral pattern generator can be improved in LAM. In particular its tunable parameters (standard deviation, horizontal correlation length) do not produce the expected results. This motivated the implementation of the Spectral Pattern Generator (SPG). Some of its theoretical (eg. “Proportionality of scales”) and practical attractive properties have been highlighted. Random pattern fields have been visualized and their statistical behaviour was investigated. As a conclusion we found that the ALADIN implemented version of SPG is able to give results that are very similar to the external program version. SPG works better than the current pattern generator in terms of many aspects which can be highlighted as the follows:

- SPG is better tunable (control parameters settings are more consistent with the actual behaviour);
- SPG has the “proportionality of scales” property;
- SPG is closer to be really isotropic.

SPG is attractive to test as an input to SPPT. Even better schemes than SPPT are designed in the future, it is very likely that uncertainty representations in NWP models will need random patterns. SPG can also be useful also for surface perturbations.

The current SPG implementation could evolve as follows:

- Using L05 and T05 namelist parameters directly from the namelist. This be just a technical improvement which does not impact the performance of SPG, but would make life of users easier.
- Implementation of 3D version. It would be possible to give some vertical structure to random patterns, and it would be interesting to see its effect.
- Implementation of non-Gaussian noise: this could be useful to perturb parameters that do not have Gaussian distributions.

Some extension to SPPT will be developed to handle proven weaknesses in current low-level ensemble spread, using observation-based methods.

Acknowledgements

The whole modeling group of ZAMG has been supportive. We also grateful to Martin Belluš for inspiring discussions and mailings and contribution as LACE Area Leader. ECMWF's special project 'spfrbout' provided SBU computer units for this work. We are really thankful to Michael Tsyrlunikov and Dmitry Gayfulin from the Russian HydroMetCenter. They gave access to their source program and documentations. Additionally they were very helpful and cooperative.

References

- Belluš, M., 2014: Stochastically perturbed physics tendencies of surface fields in ALADIN-LAEFsystem, Report on stay at ZAMG 12/05 - 20/06/2014, Vienna, Austria
- Berner, J., K. Fossell, S.-Y. Ha, J. Hacker and C. Snyder, 2015: Increasing the Skill of Probabilistic Forecasts: Understanding Performance Improvements from Model-Error Representations *Mon. Wea. Rev.*, 153: 1295-1320.
- Bouttier, F., Vié, B., Nuissier, O., Raynaud, L., 2012: Impact of Stochastic Physics in a Convection-Permitting Ensemble. *Mon. Wea. Rev.*, 140, 3706–3721.
- Bouttier, F., L. Raynaud, O. Nuissier, B. Ménétrier, 2016: Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Quart. Jour. Roy. Meteor. Soc.*, 142, 390-403.
- 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/.](http://www.ecmwf.int/publications/)]
- Szűcs, M., 2014: Stochastically Perturbed Parameterized Tendencies in ALARO and AROME , Report on stay at ZAMG 06/10/2014 - 31/10/2014, Vienna, Austria
- Szűcs, M., 2015: Tests of possible SPPT developments , Report on stay at ZAMG 28/09/2015 - 06/11/2015, Vienna, Austria
- Szűcs, M., 2016a: Stochastic pattern generators , Report on stay at ZAMG 23/05/2016 - 17/06/2016
- Szűcs, M., 2016b: Test of the SPPT scheme, Joint 26th ALADIN Workshop & HIRLAM All Staff Meeting 2016, 4-8/04/2016, Lisbon, Portugal
- Szűcs, M., 2016c: The convection-permitting ensemble system of the Hungarian Meteorological Service, SRNWP-EPS II Workshop “Probabilistic prediction of severe weather phenomena” 17-19 May 2016, Bologna, Italy
- Szűcs, M., 2016d: SPPT in AROME and ALARO: test results and open questions, HIRLAM EPS Working Week 2016-2 https://hirlam.org/trac/attachment/wiki/Meetings/EPS/EPS2016-2/SPPT_SzucsMihaly.pdf
- Szűcs, M. 2016: Arome-EPS of the Hungarian Meteorological Service. presentation to the 2016 Eumetnet-EPS project workshop, Bologna, May 2016. available at http://www.arpae.it/dettaglio_evento.asp?id=2400&idlivello=1530
- Szűcs, M. 2017: Implementation of Stochastic Pattern Generator (SPG) in ALADIN code, 2017. Report on LACE stay at ZAMG, 11pp.
- Tsyrlunikov, M. and Gayfulin, D., 2016: A Limited-Area Spatio-Temporal Stochastic Pattern Generator for ensemble prediction and ensemble data assimilation, *Meteorologische Zeitschrift*, .