# Ensemble prediction experiments with the Limited-Area Stochastic Pattern Generator

M Tsyrulnikov, E Astakhova, D Gayfulin

HydroMetCentre of Russia

Rome, 9 September 2019

# Outline

# Introduction

- Goal of the study
  A tool for model-error simulation in LAM EPS/EDA.

- Status
  1) The SPG works on 2-D and 3-D limited area spatial domains with meaningful and tunable spatio-temporal structure.
  2) AMPT implements the SPG in the additive mode with an automatically selected magnitude.
  3) AMPT works in COSMO and perturbs T,p,u,v,qv,qc,qi.

- Recent work
  Tuning and polishing of AMPT.

- Outlook
  Transfer from COSMO to ICON.

# SPG: formulation

The scheme was developed from the requirement that the 4D fields should obey the "proportionality of scales principle": large/small spatial scales should be associated with large/small temporal scales.

$$\left(\frac{\partial}{\partial t} + \frac{U}{\lambda}\sqrt{1-\lambda^2\Delta}\right)^3 \xi(t,\mathbf{s}) = \sigma\,\alpha(t,\mathbf{s})$$

– $t$ is time, $\mathbf{s}$ is the spatial vector

– $\alpha$ is the white driving noise

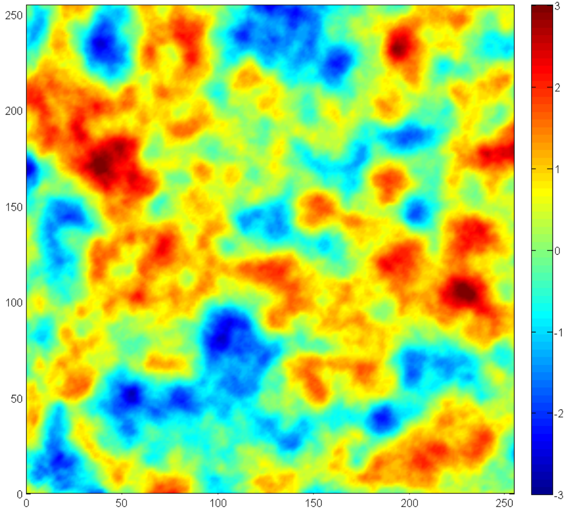
– $\xi$ is the output random field

Parameters: $\sigma$ controls the variance, $\lambda$ controls the spatial scale, $U$ controls the temporal scale

Numerics: spectral in spatial coordinates and finite-difference in time.

*Tsyrulnikov M. and Gayfulin D. A limited-area spatio-temporal stochastic pattern generator for simulation of uncertainties in ensemble applications. Meteorol. Zeitschrift, 2017, v.26, 549–566.*

# An example of the SPG random field (horizontal cross-section)



# AMPT: Motivation: Drawbacks of SPPT

1. SPPT is a multiplicative scheme and produces small perturbation whenever the physical tendency is small. But small physical tendency doesn't imply small *error*.

   $\Rightarrow$ An **additive** model-error component would resolve the problem.

2. SPPT perturbs only the magnitude of the **multivariate** physical tendency $\mathcal{P}$:
   $\mathcal{P}^* = (1 + \xi) \cdot \mathcal{P}$
   tacitly assuming that at each grid point the error is only in the *magnitude* of the vector $\mathcal{P}$, whilst the *relationships* between the physical tendencies of different variables are *error-free*, which is highly unlikely.

   $\Rightarrow$ Introducing **uncorrelated** perturbations in different *variables* can mitigate the problem.

# AMPT

The **AMPT model error perturbations** are the mutually uncorrelated spatio-temporal (SPG-generated) random fields scaled by the *area averaged* (in the horizontal) $|\mathcal{P}|$.

$|\mathcal{P}|$ is updated every hour at every level for every field.

Tapering in the lower troposphere is now switched off.

An upper-level **humidity** tapering is introduced.

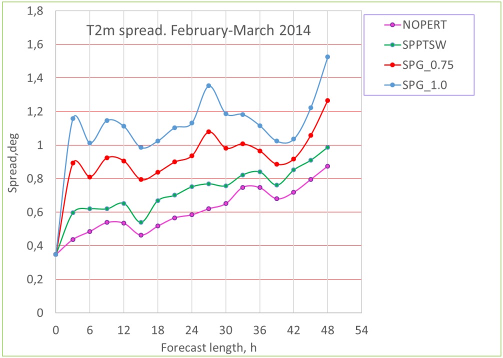
Hydrometeors: only at grid points with non-zero concentrations the perturbations are added.

# Numerical experiments

# Domain and cases

- 300*400 km area centered at Sochi (latitude 44N). Half of the domain is Black sea, another half is land with mountains.
- Resolution: 2.2 km, 50 levels.
- Ensemble size 10.
- Initial and lateral boundary conditions for ensemble members are taken from COSMO-LEPS adapted for a larger Sochi region (resolution 7 km) – made by the Italian colleagues.
- Time period: February – March 2014.

# $T_{2m}$: ensemble spread




Legend:

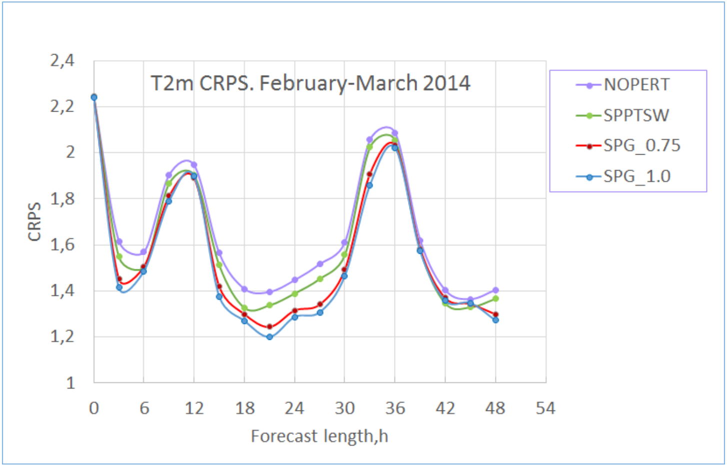
NOPERT: without model perturbations.

SPPTSW: SPPT with a "Swiss" setup.

SPG_0.75: AMPT perturbations multiplied by the factor of 0.75.

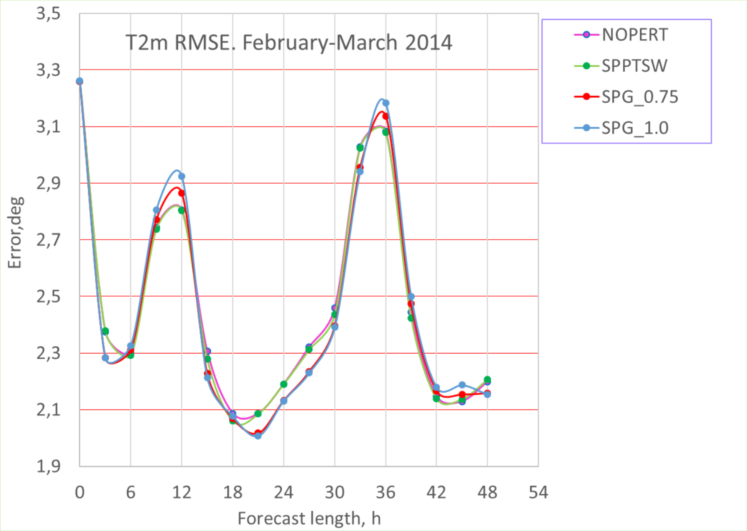
SPG_1.0: AMPT perturbations multiplied by the factor of 1.0.

# $T_{2m}$: CRPS



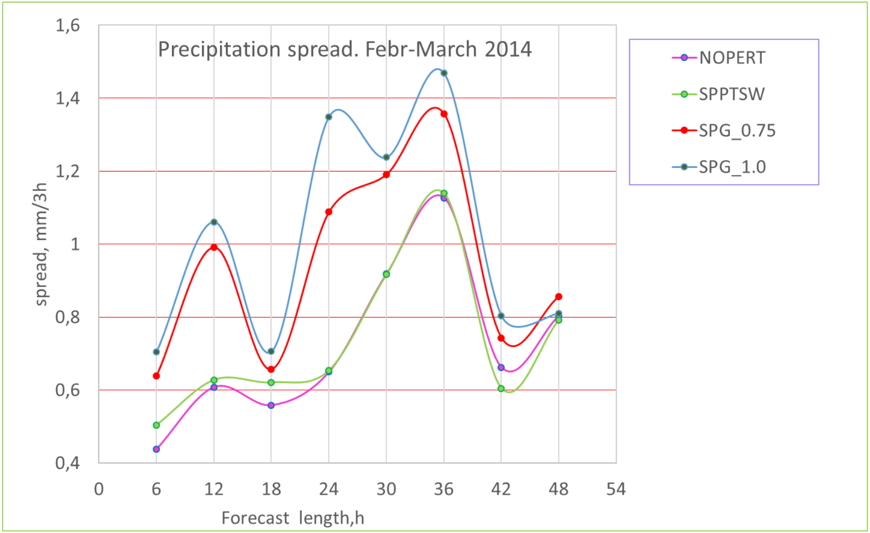
T2m CRPS. February-March 2014
CRPS
2,4
2,2
2
1,8
1,6
1,4
1,2
1
0
6
12
18
24
30
36
42
48
54
Forecast length,h
NOPERT
SPPTSW
SPG_0.75
SPG_1.0

# $T_{2m}$: RMSE

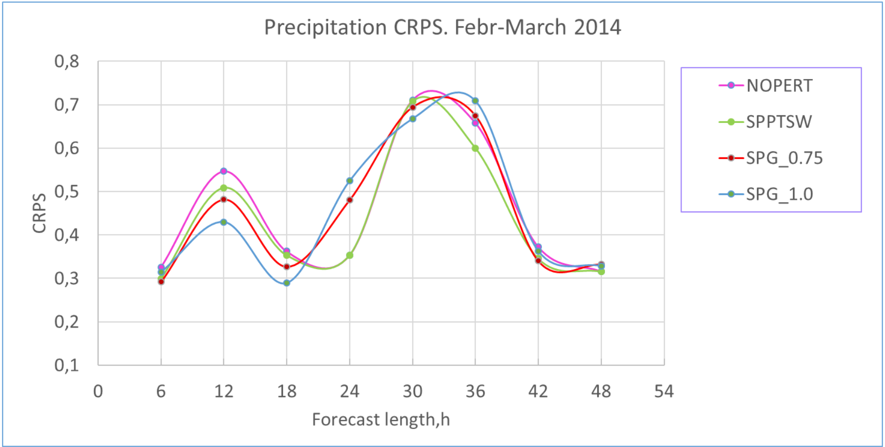


T2m RMSE. February-March 2014
3,5
3,3
3,1
2,9
2,7
2,5
2,3
2,1
1,9
Error,deg
0
6
12
18
24
30
36
42
48
54
Forecast length, h
NOPERT
SPPTSW
SPG_0.75
SPG_1.0

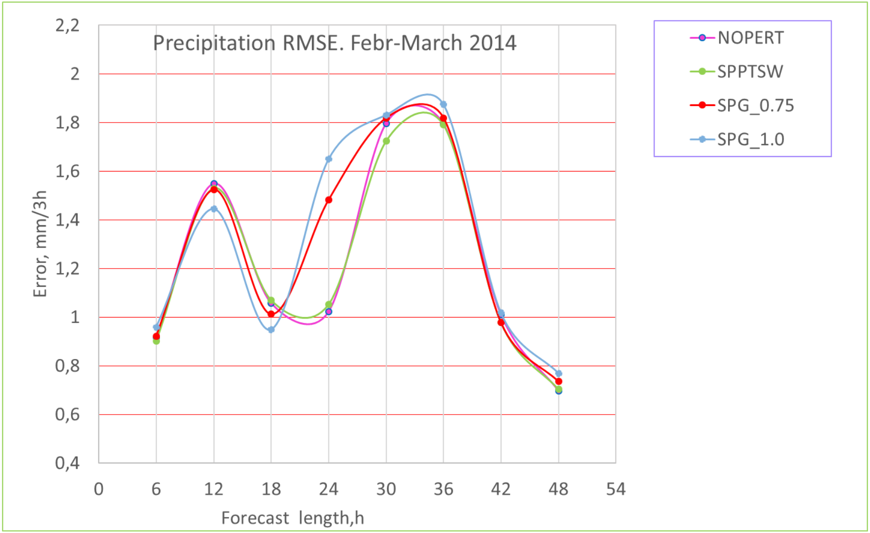
# Precipitation: ensemble spread



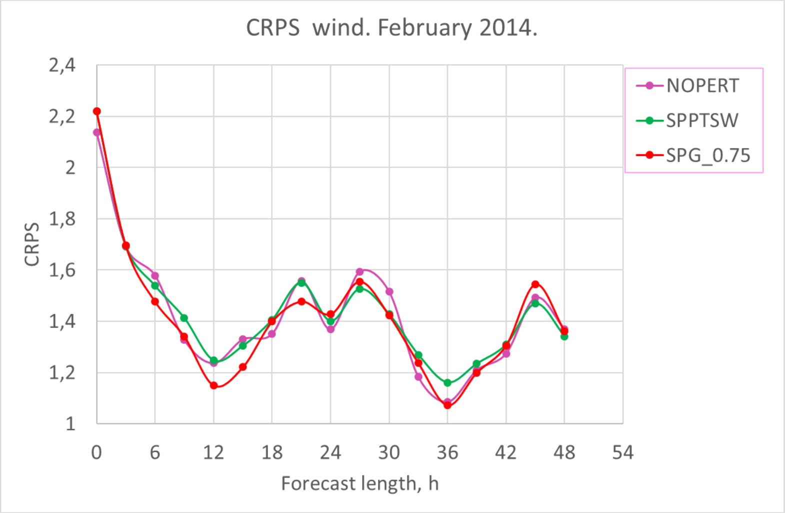
# Precipitation: CRPS



# Precipitation: RMSE



# Wind speed: CRPS



# Ensemble prediction with the *quadratic* toy forecast model

Forecast model: $f = a^2$.

$a$ is the analysis, $f$ is the forecast.

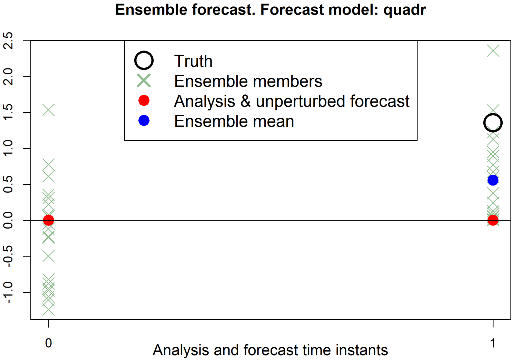
$a_i = a - \varepsilon_i$ is the analysis-ensemble member ($i = 1, \dots, N$; $N$ is the ensemble size).

$\varepsilon \sim N(0, 1)$ represents the error (analysis and model errors combined).

$f_i = (a - \varepsilon_i)^2$ is the forecast-ensemble member.

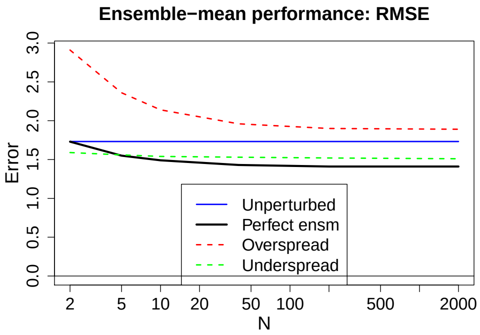
The truth is generated in the same way as an ensemble member.

Let $a = 0$.



# Performance of deterministic point forecasts: RMSE

Strongly nonlinear model: $a = 0$.




$\Rightarrow$ With a **nonlinear** model, the ensemble mean can be significantly better than the unperturbed forecast.

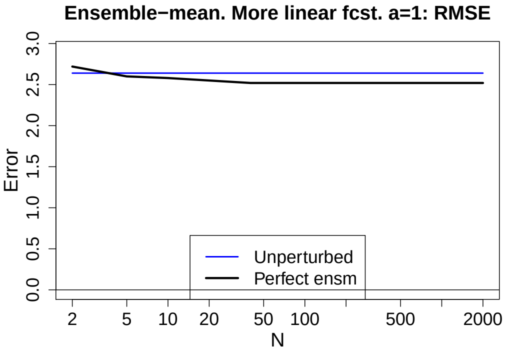
$\Rightarrow$ An underspread ensemble can perform better than the perfect ensemble because it is, actually, a mix of with the unperturbed forecast.

$\Rightarrow$ Overspread in the ensemble is more harmful than underspread.

# Performance of point forecasts: RMSE

A more linear model: $a = 1$.

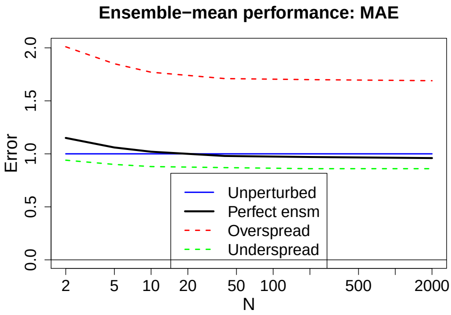
$truth = (a - \varepsilon)^2 = a^2 - 2a\varepsilon + \varepsilon^2$




$\Rightarrow$ If nonlinearity is weak, it's hard to beat the unperturbed forecast.

# Performance of point forecasts: MAE

Strongly nonlinear model: $a = 0$.




$\Rightarrow$ In the MEAN-ABSOLUTE sense:

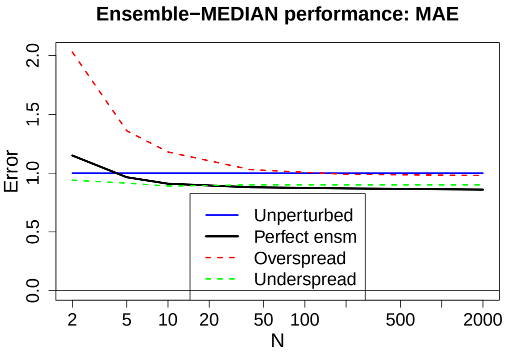
(i) Ensemble mean can be worse and is never significantly better than the unperturbed forecast.

(ii) The overspread ensemble performs very poorly.

(iii) The underspread ensemble performs uniformly better than the perfect ensemble (?)

# Ensemble MEDIAN: MAE

Strongly nonlinear model: $a = 0$.




$\Rightarrow$ It is the MEDIAN that is optimal in the MEAN-ABSOLUTE sense.

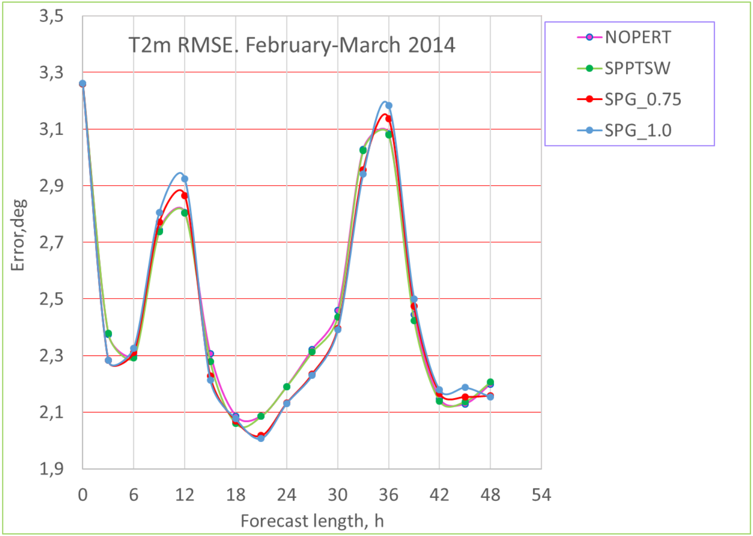
$\Rightarrow$ Ensemble MEDIAN is more robust to misspecification of error model.

$\Rightarrow$ The underspread ensemble is still competitive with the perfect ensemble.

# Take-away messages from the toy experiments

1. If the nonlinearity is **strong**, then the ensemble mean tends to be better than the unperturbed forecast in terms of both bias and RMSE.
2. If the nonlinearity is **weak**, then the unperturbed forecast may be the best choice.
3. Overspread ensembles perform poorly.
4. Underspread ensembles can be regarded as a mixture of the ensemble with the unperturbed forecast and perform much better than overspread ensembles and sometimes even better that the perfect ensemble.
5. In the mean-absolute sense (i.e. measured by MAE), the ensemble mean is no better than the unperturbed forecast.
6. If the verification score is MAE, then the ensemble median is to be considered as an alternative of the ensemble mean (in an EPS).
7. The ensemble median is more robust to misspecifications than the ensemble mean.

# $T_{2\mathrm{m}}$: RMSE



# Summary of the results

Two-month experiments (in winter-spring) with the debugged and tuned version of SPG/AMPT.

1. Probabilistic forecasts.
   - Reliability (measured by the proximity of spread to skill) was significantly improved for all tested elements ($T_{2m}$, Precip, $|V_{2m}|$).
   - Resolution (measured by CRPS) was significantly improved for $T_{2m}$, improved for $|V_{2m}|$, and slightly degraded for Precip.
2. Performance of the ensemble mean forecasts (RMSE).
   - $T_{2m}$: improvement at night and deterioration at the height of the day.
   - $|V_{2m}|$: neutral impact.
   - Precip: slightly negative impact.

# Further steps

- Implementation of SPG/AMPT in ICON (in the LAM setup).
- Setting up a new LAM-EPS in central Russia.
- Replacement of VERSUS with another verification tool.
- Improvement in the generation of AMPT wind perturbations (switching from $u, v$ to stream function and velocity potential).
- Further investigation into the role of humidity and hydrometeor perturbations.

Thank you!