SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	2022
Project Title:	Operationalization of SPP and further improvements of EDA, boundary and surface perturbation in MEPS
Computer Project Account:	spseandr
Principal Investigator(s):	Ulf Andrae
Affiliation:	SMHI
Name of ECMWF scientist(s) collaborating to the project (if applicable)	None
Start date of the project:	2020-01-01
Expected end date:	2022-12-31

Computer resources allocated/used for the current year and the previous one
(if applicable)
Please answer for all project resources

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	16M	11.3M	16M	12.3M
Data storage capacity	(Gbytes)	60000	>90000	90000	>120 000

Summary of project objectives (10 lines max)

The aim of the project is to improve ensemble related components in the MetCoOp ensemble system MEPS. Last year we focused on a possible new cycling strategy where the EDA and ENS streams are separated, and getting a first setup of the model uncertainty scheme SPP (the stochastically perturbed parameterizations scheme) ready for operational use.

Summary of problems encountered (10 lines max)

The general feeling is that the throughput of experiments on cca continues to be slow both due to slow MARS retrievals and long queuing times on cca. The slow access to data on ECFS is also always a challenge.

Summary of plans for the continuation of the project (10 lines max)

Test different cycling strategies by separating DA and EPS streams, and at the same time revisit choices for initial perturbation sizes in presence of SPP.

Include more parameters in SPP.

List of publications/reports from the project with complete references

None

Summary of results

If submitted **during the first project year**, please summarise the results achieved during the period from the project start to June of the current year. A few paragraphs might be sufficient. If submitted **during the second project year**, this summary should be more detailed and cover the period from the project start. The length, at most 8 pages, should reflect the complexity of the project. Alternatively, it could be replaced by a short summary plus an existing scientific report on the project attached to this document. If submitted **during the third project year**, please summarise the results achieved during the period from July of the previous year to June of the current year. A few paragraphs might be sufficient.

Introduction

The MetCoOp ensemble (MEPS) serves the participating countries (Estonia, Finland, Norway and Sweden) with high resolution ensemble forecasts over the domain shown in Figure 1. MEPS is based on HarmonEPS developed within the ACCORD consortium (Frogner et al. 2019). A redesigned setup for MEPS was introduced in operations in February 2020. The new ensemble is based on Ensembles of Data Assimilation (EDA), using members from IFSENS on the boundaries and is running in a continuous mode as described in Andrae et al. (2020). The upgrade also included a bigger domain. The new larger domain requires a new set of background error statistics (BES) for the data assimilation and it is normally a costly procedure to generate representative statistics for all seasons. A lot could be gained if (pre) operational data from the ensemble forecasts could be used and it would also allow us to explore a more continuous update of statistics representing "errors of the day". The first part of this work was reported last year, in this report we describe the work done after the last report, which includes testing the new BES in MEPS forecasts.

Currently there is no model uncertainty representation in upper air in MEPS. The stochastically perturbed parameterizations scheme (SPP) is being developed in HarmonEPS (Frogner et al. 2022). The second part of this report consists of an analysis of the behaviour of a reduced parameter setup for SPP in MEPS.

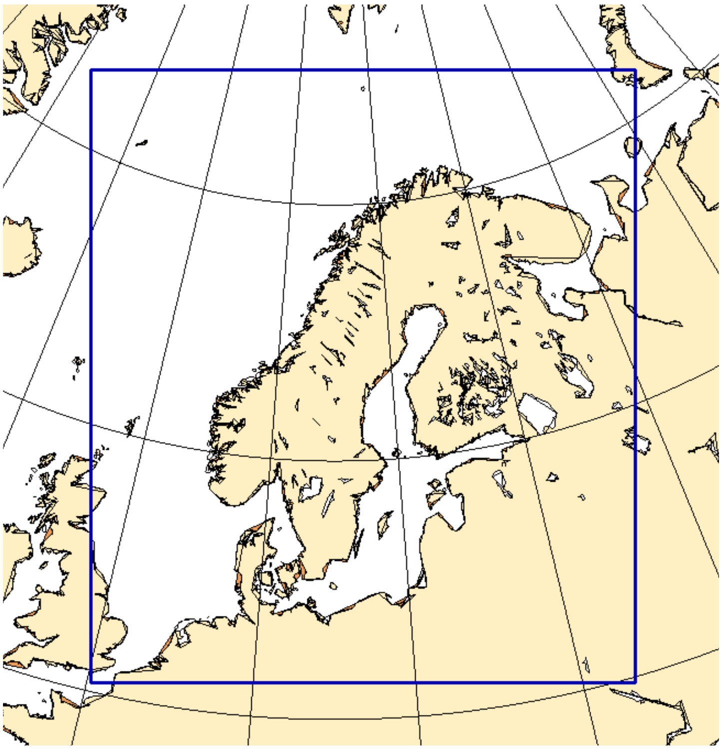

Figure 1: MEPS area.

Effect on EPS scores and background error statistics from separated EPS and DA perturbations

MEPS currently runs EDA and the ensemble in one suite, see Figure 2. This results in the perturbations being cycled. We have seen that cycling all perturbations leads to a bias in the ensemble members compared to the control, and also excessive initial spread. We have therefore tested to separate the EDA and EPS streams, as seen in Figure 3. The effect this has on the scores is shown in Figure 4. The spread is reduced when we separate EDA and EPS streams, but the spread came partly from a bias introduced by the cycling of the perturbations (not shown). The growth of the spread is also more steady when this separation is done, and does not show the undesired drop in spread for the first forecast hours.

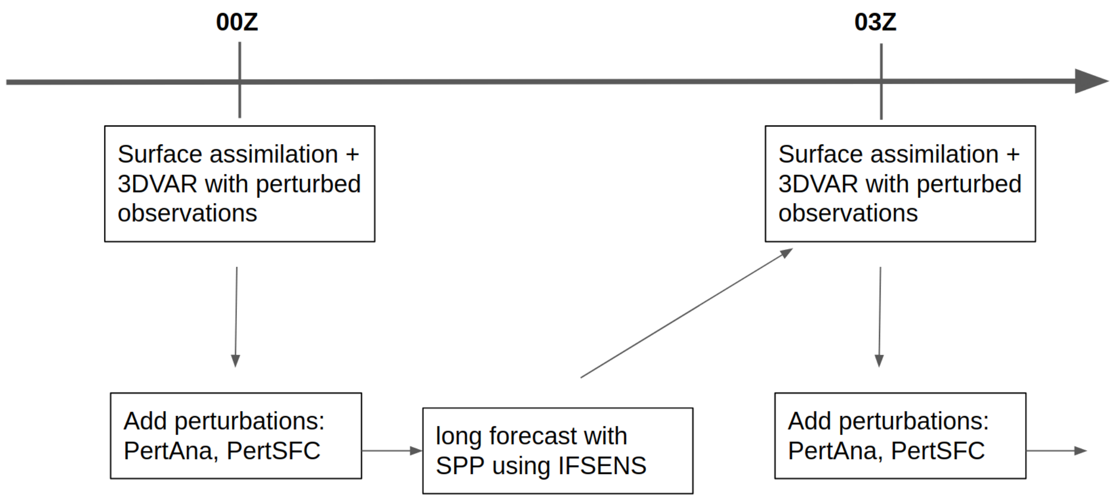

Figure 2: Schematics of how MEPS is currently run. PertAna and PertSFC are the initial and surface perturbations, respectively (see Frogner et al. 2019 for details).

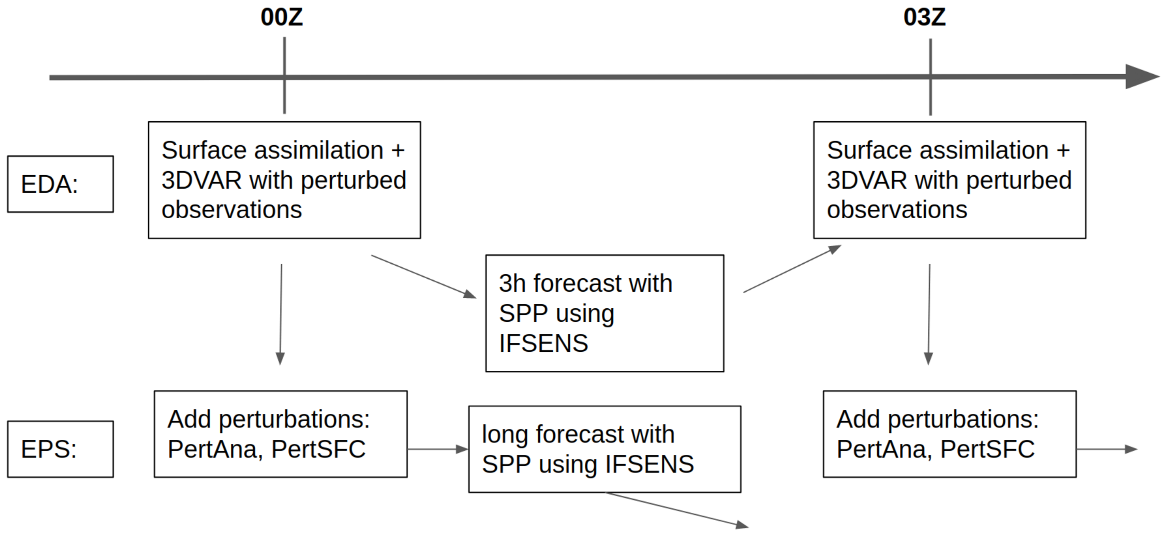

Figure 3: Schematics of how EDA and EPS can be separated.

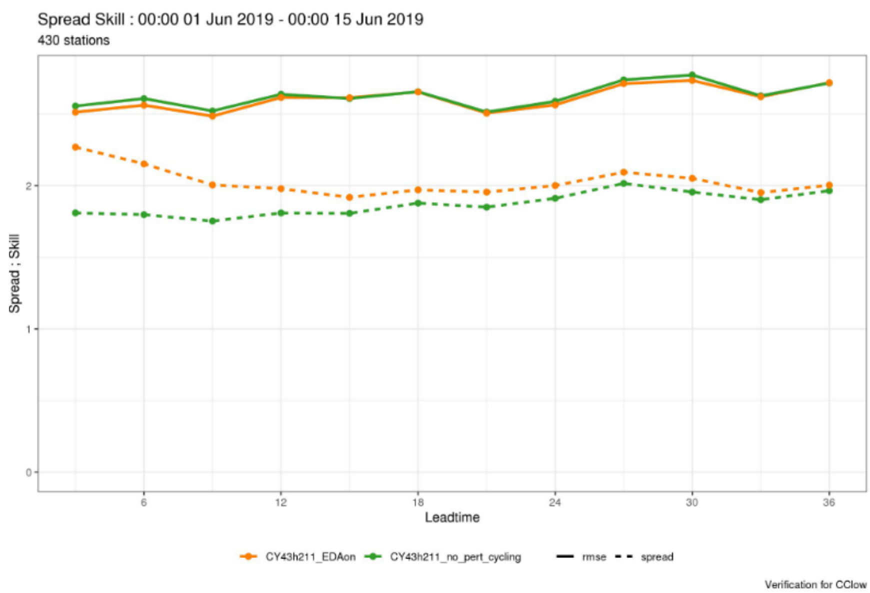

Figure 4: Spread and skill for low clouds for 14 days in June 2019. Spread in dashed lines and skill in solid lines, orange is when the perturbations are cycled (as in Figure 2) and green when they are not (as in Figure 3).

When EDA was introduced in MEPS it was hoped that the ensemble could be used also for generating BES. However, it was discovered that the EPS perturbations in combination with using IFSENS at the boundaries were introducing too large scales. When EDA and EPS are separated like in Figure 3 the BES are comparable to the ones used operationally independent of using IFSENS boundaries or ELDA boundaries. The operationally used BES were derived in 2017 using boundaries from the ECMWF ELDA suite available at the time. ELDA boundaries however are not suitable for MEPS forecasts as boundaries as they are not available in dissemination and have too short forecasts. There are some tiny differences in characteristics when background error statistics have been constructed using ELDA or IFSENS (not shown), with e.g. more energy on larger scales and slightly broader vertical correlations for the IFSENS case. However, the differences in forecast scores between the two are very small, see Figure 5. More extensive and in depth evaluation is required though before we can conclude that it's appropriate to replace ELDA with IFSENS for generation of BES.

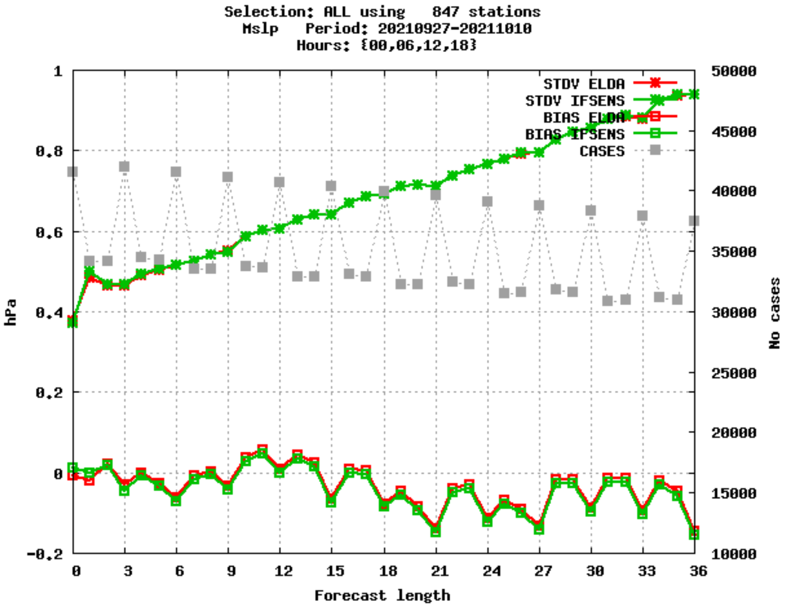

Figure 5: Bias (lower curves) and standard deviation (upper curves) for a forecast using BES generated with ELDA (red) and IFSENS (green). For MSLP for 14 days in autumn 2021.

First SPP setup ready for operations

The stochastically perturbed parameterizations scheme (SPP) is implemented and tested in HarmonEPS. SPP introduces stochastic perturbations to values of chosen closure parameters representing efficiencies or rates of change in parameterized atmospheric (sub)processes. SPP in the first version in HarmonEPS (Frogner et al. 2022) perturbs 11 parameters, active in different atmospheric processes and under various weather conditions. The main motivation for developing SPP was the lack of variability seen in cloud products in HarmonEPS. SPP in this first version is able to increase variability in a range of weather variables, including the cloud products. However, for some weather variables the root-mean-squared error of the ensemble mean was increased and a mean bias was introduced, especially in winter. This indicated that (some) parameter perturbation distributions were not optimal in the first configuration, and a further sensitivity analysis was required. To speed up operationalization in MEPS, it was decided to focus on a subset of the parameters in SPP; five parameters were chosen that had proven to be effective in creating ensemble spread. In the first version of SPP it was seen that using a lognormal distribution for some parameters resulted in biassed response to the perturbations drawn from the distribution. This was both because of the long tail at the right hand of the distribution that can result from a lognormal distribution, and from the accumulation of values near the left end of the distribution. A new distribution was therefore introduced; uniform. The uniform distribution was also implemented with the possibility to shift the distribution to both smaller and larger values. A careful adjustment of the standard deviation of parameter pdf's was done, one at a time, including testing the two different distributions. An example is in Figure 6. In Figure 7 the resulting spread obtained from perturbing this parameter with these three distributions can be seen. The largest spread was obtained with a lognormal distribution with a standard deviation of 0.6, but this resulted in the ensemble members having less clouds than the control. It is not desirable that the perturbations introduce biases compared to the unperturbed model. Both a lognormal distribution with a standard deviation of 0.3 and a uniform distribution with standard deviation of 1.2 leads to an unbiased ensemble compared with the control (see Figure 6), however, the spread is clearly improved from using the uniform distribution. Similar investigations were done for the four other parameters chosen. The list of the five parameters and their settings is in table 1. In addition to changing distribution, the new version also includes the possibility to correlate or anti-correlate the perturbations of two parameters. This is done when two parameters are influencing the same process(es) and when it is physically meaningful that they vary together. Two of the parameters in the setup described here are correlated: the Stable conditions length scale and the Asymptotic free atmospheric length scale.

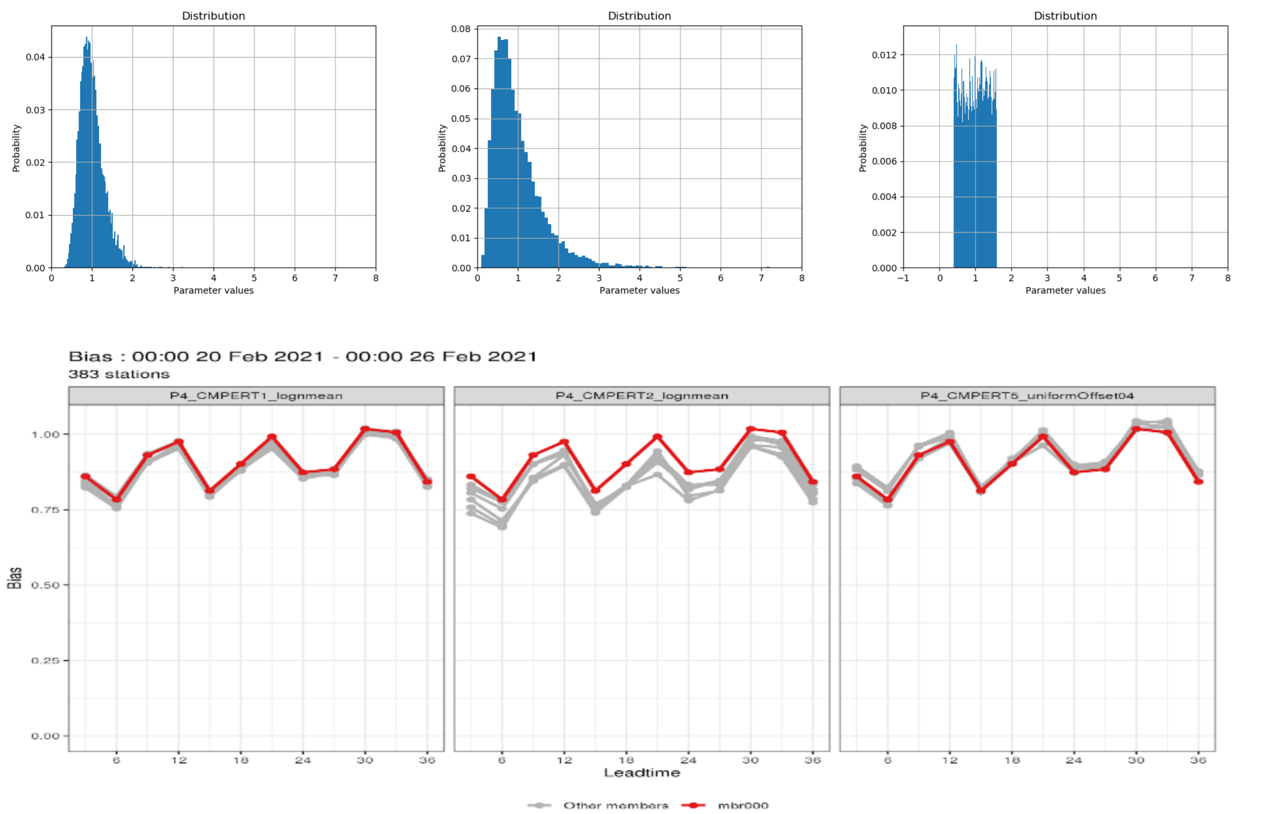

Figure 6: Top: three different distributions for a parameter that controls the cloud ice content impact on cloud thickness (ICE_CLD_WGT, deterministic value is 1.0). Left: lognormal distribution with standard deviation of 0.3. Middle: lognormal distribution with standard deviation of 0.6. Right: uniform distribution with standard deviation of 1.2. Bottom: the resulting bias for total cloud cover from utilising the distributions above. Red is for the unperturbed control members, grey for the ensemble members.

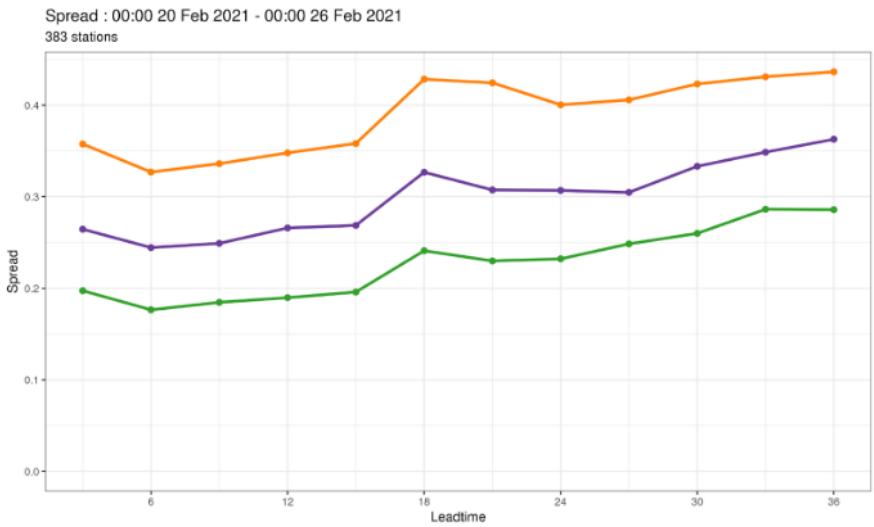

Figure 7: Spread for total cloud cover, perturbing ICE_CLD_WGT with lognormal distribution with standard deviation of 0.3 (left in figure 6, green), with lognormal distribution with standard deviation of 0.6 (middle in figure 6, orange) and with uniform distribution with standard deviation of 1.2 (right in figure 6, purple).

	Distribution	Standard deviation	Correlation	Shifted
Saturation limit sensitivity	Lognormal	0.6	no	no
Threshold cloud thickness used in shallow/deep convection decision	Lognormal	0.6	no	no
Ice content impact on cloud thickness	Uniform	1.2	no	no
Stable conditions length scale	Uniform	1.05	Yes, with asymptotic free atmospheric length scale	Yes
Asymptotic free atmospheric length scale	Lognormal	0.45	Yes, with stable conditions length scale	no

Table 1: The parameters used in the reduced SPP setup and the perturbation characteristics.

After testing each parameter separately, the combination of all five parameters was tested in the full ensemble setup, with all other perturbations active (initial condition perturbations including EDA, surface perturbations and lateral boundary perturbations, see Frogner et al. 2019 for details). An example on how it performs compared to a reference without SPP is in Figure 8. Including this reduced set of well tuned SPP parameters clearly results in higher spread and without changing the bias of the ensemble.

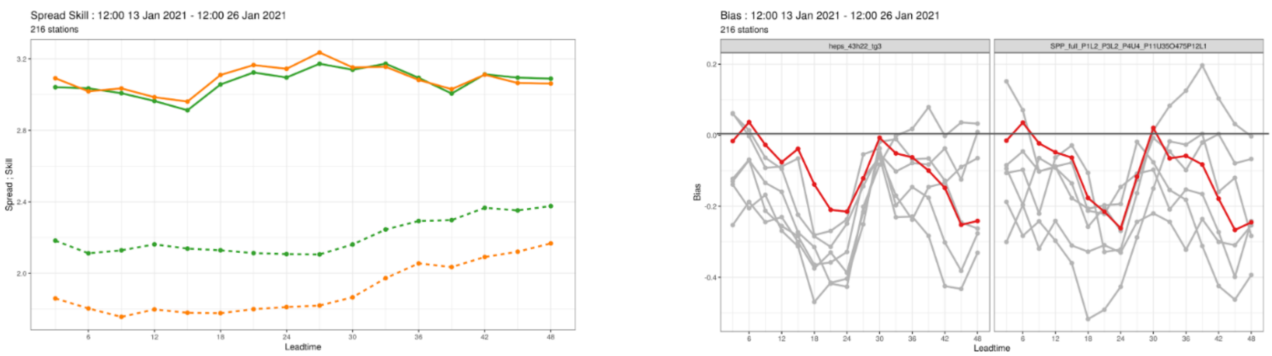

Figure 8: spread and skill (left) and bias (right) for low cloud cover. Reference without SPP in orange in the spread-skill plot and to the left in the bias plots, with SPP in green in the spread-skill plot and to the right in the bias plots. Red curves are bias for the control runs, the members in grey.

The reduced SPP setup is also tested in operational settings for cases with forecast failure of the current system. In Figure 9 a satellite picture from 9th of March 2022 at 12 UTC is shown, low clouds are clearly visible over the southern parts of Finland, and stretching southwards. In Figure 10 a 24 h forecast for the same time is shown (the control and 6 members) from a MEPS-like run without SPP. The area with low clouds is too small, only covering a small band in the southern most part. The same is shown in Figure 11, but now with SPP on. SPP is able to extend the area with low clouds in some members, particularly members 3 and 5, more in agreement with the satellite picture.

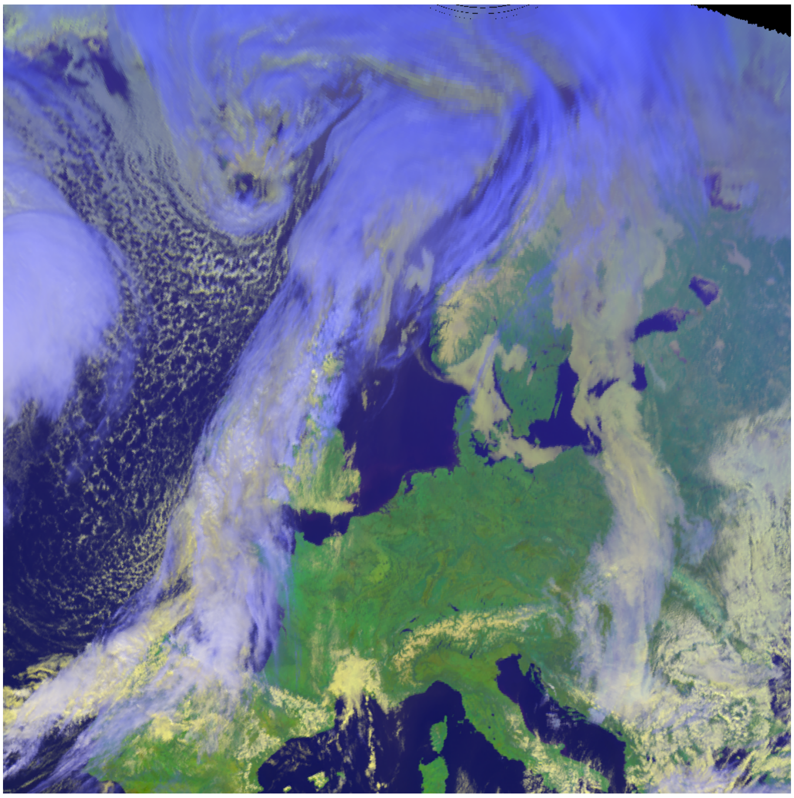

Figure 9: Satellite picture from 9th of March 2022 at 12 UTC.

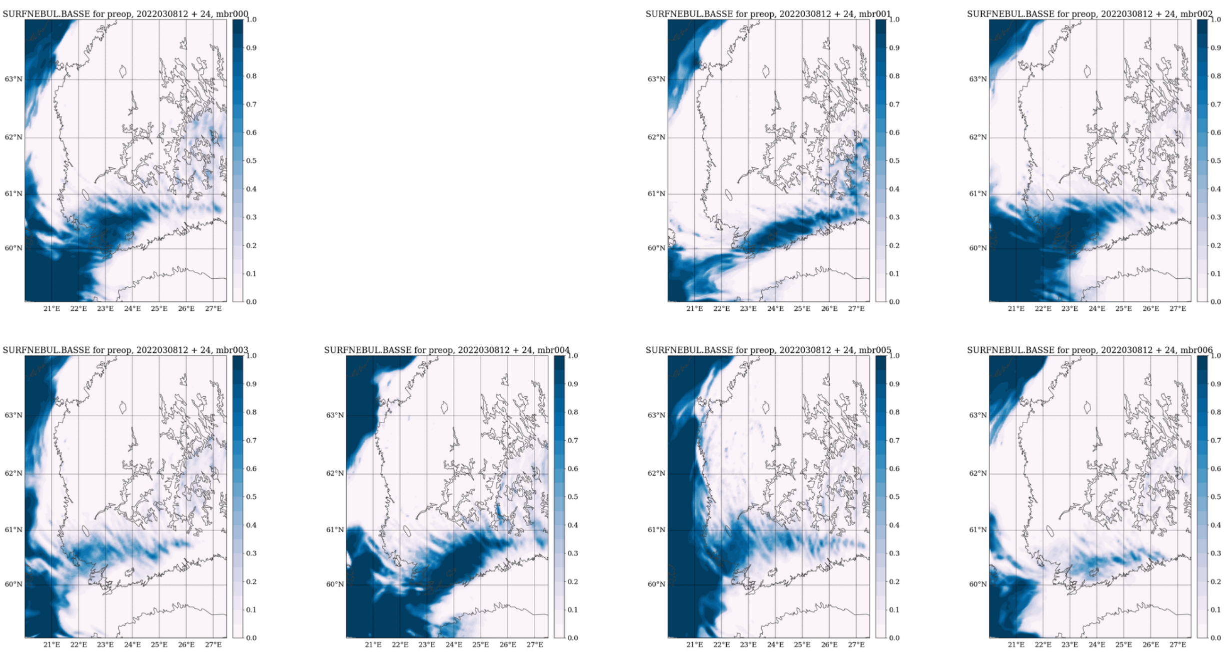

Figure 10: The control and 6 members from a run without SPP, zoomed in over the area of interest over southern Finland.

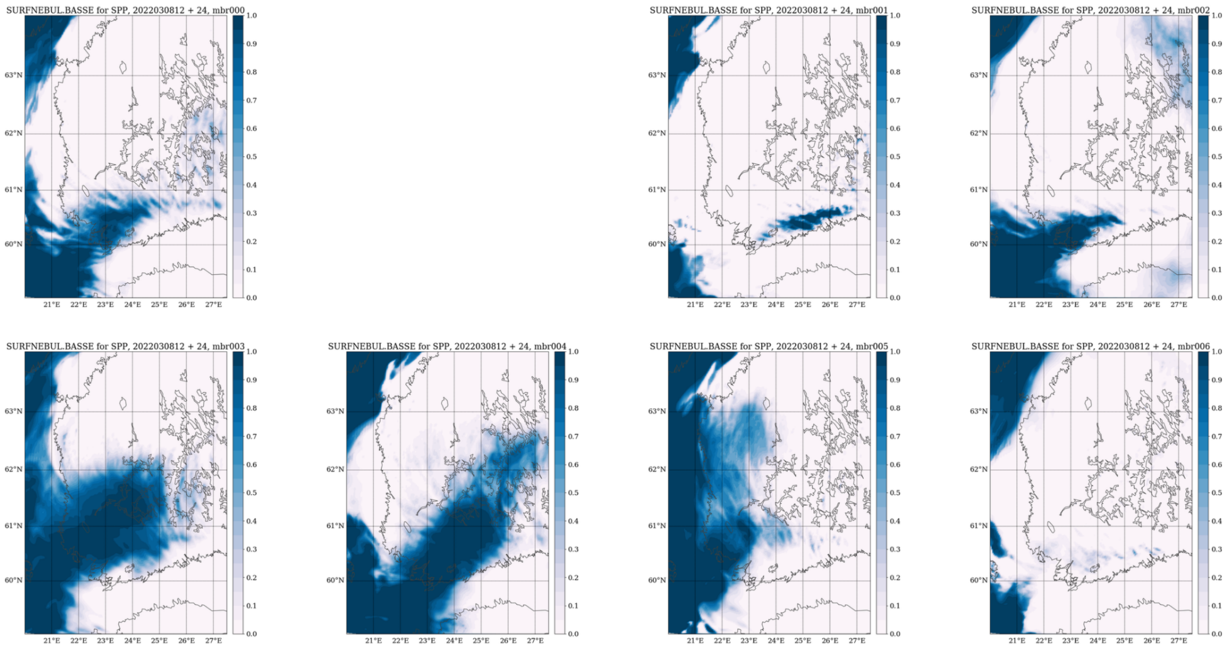

Figure 11: The control and 6 members from a run with SPP, zoomed in over the area of interest over southern Finland.

This reduced set of 5 parameter SPP described here entered MEPS preoperational run on 21 June 2022. If behaving as expected it will enter into MEPS in the third quarter of 2022. The cost of adding SPP is less than 1%, this is achieved by applying the perturbations every hour instead of every timestep as in Frogner et al. (2022).

References:

Andrae, U. et.al, 2020, A continuous EDA based ensemble in MetCoOp, ALADIN-HIRLAM Newsletter Nr 14, 189-198

Frogner, I., Andrae, U., Bojarova, J., Callado, A., Escribà, P., Feddersen, H., Hally, A., Kauhanen, J., Randriamampianina, R., Singleton, A., Smet, G., van der Veen, S., and Vignes, O. (2019). HarmonEPS - The HARMONIE Ensemble Prediction System. Weather and Forecasting 34, 6, 1909-1937,
https://doi.org/10.1175/WAF-D-19-0030.1

Frogner, I., Andrae, U., Ollinaho, P., Hally, A., Hämäläinen, K., Kauhanen, J., Ivarsson, K., & Yazgi, D. (2022). Model Uncertainty Representation in a Convection-Permitting Ensemble - SPP and SPPT in HarmonEPS, Monthly Weather Review 150, 4, 775-795,
https://doi.org/10.1175/MWR-D-21-0099.1