A continuous EDA based ensemble in MetCoOp

Technical Report · January 2020

Inger-Lise Frogner
Norwegian Meteorological Institute

Ole Vignes
Norwegian Meteorological Institute

Roohollaha Azad
Norwegian Meteorological Institute

A continuous EDA based ensemble in MetCoOp

Ulf Andrae[1], Inger-Lise Frogner[2], Ole Vignes[2], Andrew Singleton[2], Roohollah Azad[2], Mikko Partio[3], Niko Sokka[3]

1. Swedish Meteorological and Hydrological Institute (SMHI)
2. Norwegian Meteorolical Institute (MET Norway)
3. Finnish Meteorological Institute (FMI)

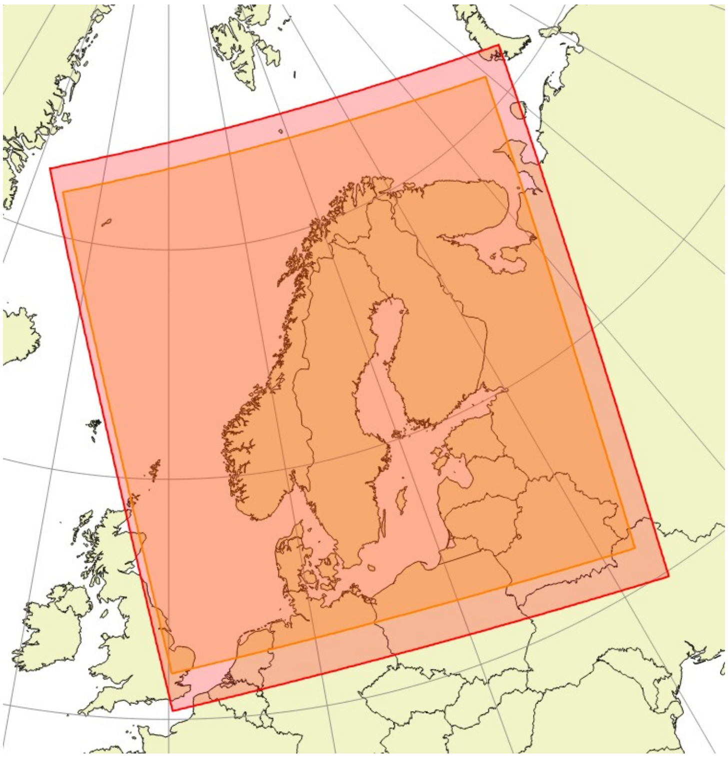

Figure 1. Current and future domain

1 Introduction

The MetCoOp ensemble prediction system, MEPS, has been operational since late 2016 and is today an important component in the daily forecasting at the institutes. The increased computational capacity made available in early 2019 offered the possibility to introduce a parallel, close to full scale, experimental ensemble system. In the following we describe the properties and limitations of the current system, the reasoning behind the design of the new setup and a comparison of the performance of the two systems. The new ensemble will replace the current one as the operational ensemble for MetCoOp in February 2020.

2 Properties of the current ensemble

The current operational system, MEPS (Andrae 2017), is based on the harmonie-40h1.1.1 version of the HARMONIE-AROME model (Bengtsson et.al. 2017) with perturbation methods as in HarmonEPS (Frogner et.al. 2019b). The model runs over the inner domain shown in Figure 1 with a horizontal resolution of 2.5km and 65 levels in the vertical. For the first three members, control and pseudo-control members, the assimilation cycle is 3h and a large variety of conventional and non-conventional observations are assimilated. For the additional members the assimilation cycle is 6h and only surface assimilation is applied. The generation of the ensemble members is distributed over three HPCF, two in Sweden and one in Finland, which runs 5, 4 and 1 members respectively. MEPS serves as the primary source of uncertainty information for the first two days at the MetCoOp institutes and has been shown to add forecast skill on e.g. precipitation for scales not described well by the ECMWF ensemble IFSENS (Frogner et.al. 2019a). We can however identify a few problems and limitations in the current system:

- The number of members are limited by two factors. The first is the total computational capacity available when we run all members simultaneously and constrained by a delivery time of about one hour. The second limiting factor is the SLAF methodology where we use two IFS HRES forecasts from 6h apart to generate the perturbations, with hourly nesting. In practise it means that the boundary file for member nine at +54h uses the difference from +84h and +90h. From +90h the data is only available every 3h, hence, there is no room for further expansion with the current method.
- Attempts to use MEPS for nowcasting of precipitation have failed due to an overestimation of precipitation during the first forecast hours. Further it has been noted that the wind field during the first hours has an undesirable high noise level. Both aspects suggests that the current initial perturbations are too imbalanced.
- A general feedback from duty forecasters is a lack of spread in e.g. clouds making the ensemble less useful for e.g. aviation forecasting.

3 Designing a new ensemble

The continuous approach

The most straightforward and common way to produce an ensemble is to launch all members at the same time and as in the current MEPS configuration perturb the initial state around a control analysis. This approach leads to a very unbalanced usage of the computer resources and, as mentioned above, a strong constraint on the numbers of members. To construct probability products by lagging of several forecasts with different initial time is not a new approach (Hoffman and Kalnay 1983). However, to distribute the generation of ensemble members in time as described by Yang et. al. (2017) and Porson et.al. (2020) is fairly new. The approach has several appealing properties. Spreading the generation of members in time not only evens out the workload on the computer, it also allows us to increase the number of members for a given maximum computer capacity. For the forecaster we are able to present frequent updates and an ensemble with less jumpiness, due to the lagging. In our setup we've distributed the members in time, and over our three HPCF as shown in table 1. The results presented here are based on 9 unique members whereas the operational setup will have 6 more members meaning that we will produce 5 new members each hour. Hereafter we will call this suite CMEPS. The ensemble related configuration differences between MEPS and CMEPS are summarized in table 3. The settings for IFSENS are included as a reference.

Table 1: Distribution of members in time and over the MetCoOp computers for the results described in the report. Numbers in parentheses are additional members in the operational setup

	Time (UTC)	**Cirrus**	**Stratus**	***Voima***
Stream 1	00,03,...,21	0, *(12)*	1,2	*(9)*
Stream 2	01,04,...,22	7,*(13)*	3,4	*(10)*
Stream 3	02,05,...,23	8,*(14)*	5,6	*(11)*

Ensembles of Data Assimilation (EDA) properties

Describing the initial uncertainty through perturbation of observations within their observations error limit is a well established method used e.g. by ECMWF (Isaksen et.al. 2010). The method is also implemented in the HarmonEPS system and has been shown to improve the spread of e.g. near surface variables without loss of skill (Frogner et.al. 2019b). For a system that is distributed in time and over several HPCs it's also appealing not to have to relate to a single control analysis. Another aspect is the potential usage of the operational ensemble to generate background error statistics for the data assimilation discussed later. Either as climatological statistics from different seasons or in a more continuous manner. Following Frogner et.al. (2019b) we have combined EDA with perturbations derived from IFSENS but scaled by 0.5 as compared to the SLAF based perturbations in MEPS. Running with EDA only does not provide a satisfactory spread (result not shown).

Distributing the different streams in time, like shown in table 1, naturally has an impact on the number of observations available for assimilation. The result is summarized in table 2 and we note that for most observations types the difference is small. The most striking difference is the number of radiosondes (TEMP T) which mainly launched at synoptic hours and which results in 85% less observations in stream 2 compared to stream 0. However, when comparing the mean absolute error of e.g. MSLP from perturbed member forecasts we see no differences in quality (not shown) suggesting that all members have similar quality despite the variations in observation density.

Table 2. Typical number of observations at different hours (UTC)

	Stream 0 (00,03,…,21)	**Stream 1 (01,04,…,22)**	**Stream 2 (02,05,…,23)**
TEMP T	250-300	150-200	40
AIRCRAFT T	105	130	120
SYNOP Z	530	530	445
GNSS	70	70	70
RADAR	1100	1100	1100
ASCAT	600	750	600
IASI (METOP1)	300	300	300
SYNOP SNOW	>650	>650	>650
SYNOP T2M	850	760	760

In the default HARMONIE-AROME setup snow is assimilated once a day at 06Z since this is when the majority of the snow observations are reported. This would of course leave the members not running at 06Z without any snow assimilation or at least with much less observations. To overcome this the snow assimilation has been shifted to be done at 07/08/09Z using a longer time window to maximize the number of snow observations used. An example of the impact in observation usage is shown in figure 2.

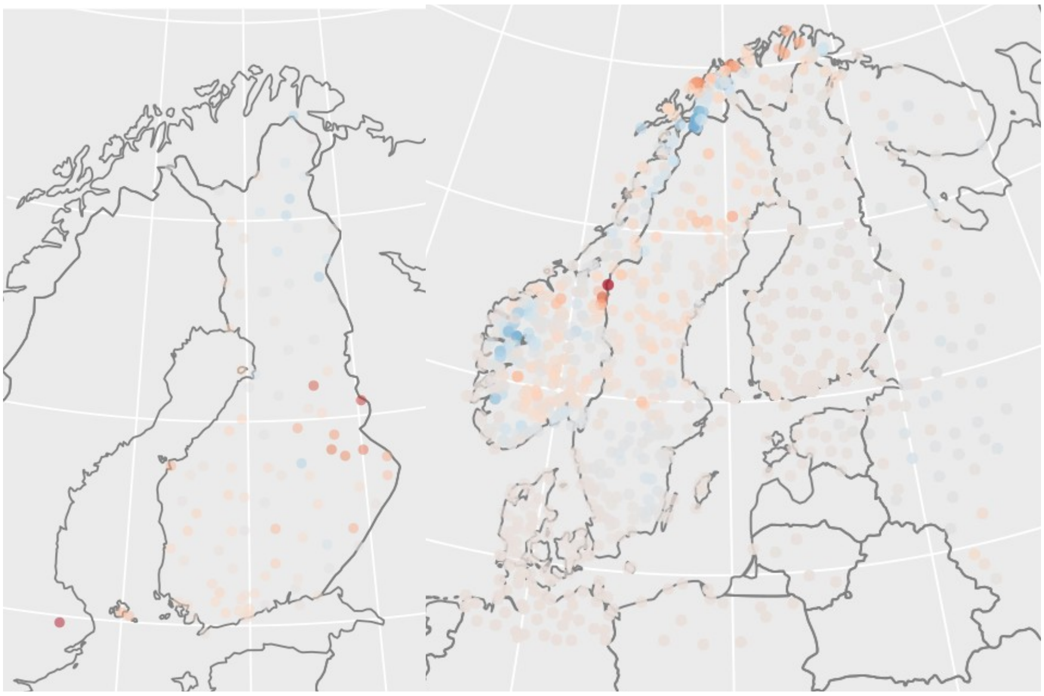

Figure 2. Snow increments at 2019-03-04 08Z before (left) and after (right) the extended observation usage.

Using IFSENS boundaries

Generating lateral boundary perturbations through the SLAF method has been an efficient way to introduce an operational ensemble without any extra requirements on the dissemination of data from the global model (IFS). However, as mentioned earlier we are, with the current method, limited by the maximum available forecast length of the forcing model. Frogner et. al. (2019b) showed that by using IFSENS it's possible to maintain a larger spread throughout the forecast compared to SLAF. It was also shown that with a clustering method maximizing the spread (Molteni et. al. 2001) it's possible to select IFSENS members and increase the spread even more. For simplicity this is not pursued further but we pick the first 28 members so that each CMEPS member uses different IFSENS members seen over a time window of 6h. The control member is still coupled to IFS HRES from which we also take the SST and ice cover for all members since the coarser resolution in IFSENS degrades T2M scores along coastlines. Note that the perturbation of SST maintains the variability in the lower boundary forcing over sea.

Technical aspects

Although the three streams in table 1 are in principle independent they only differ by a few configuration settings. Running a continuous setup of course poses challenges when it comes to updates and maintenance, especially for a setup running simultaneously on three different platforms. Using HARMONIE terminology we've therefore introduced the concept of STREAM to allow several forecast cycles to run in parallel in the same experiment setup. Each STREAM has an independent set of binaries, scripts and config files.

To allow us to grow the ensemble with a factor of three without increasing the cost for data transfer and storage too much we introduced GRIB2 using the lossless CCSDS algorithm with an accuracy of 16 bits for all fields. The algorithm provides a reasonable balance between reduced size and increased cost in packing/unpacking, figure 3. As a result the size reduced with 59% for a typical history file and 56% for a fullpos file.

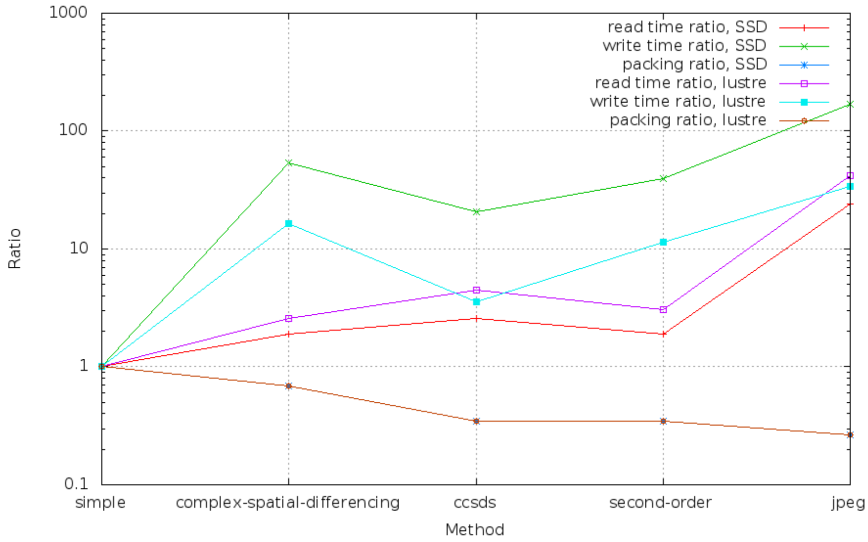

Figure 3. Comparison of various packing methods available in ecCodes.

Table 3: Summary of properties of the current and next generation ensemble. The ECMWF ensemble, IFSENS, is included as a reference

	MEPS	**CMEPS**	**IFSENS**
Availability and members	10 members up to 54h every 6h	5 members up to 66h every 1h	51 members up to 90h every 6h
Initial uncertainty	Perturbed control analysis using SLAF (+/-) Perturbed surface state variables	EDA + perturbation from IFSENS Perturbed surface state variables	EDA Singular vectors
Model uncertainty	Perturbed surface properties	Perturbed surface properties	SPPT
Lateral boundary uncertainty	Perturbed ECMWF HRES using SLAF (+/-)	Use the first 28 ECMWF IFSENS members. HRES is used for member 0.	

4 Results

The continous setup has been running in real time since February 2019 and in early May the domain was extended to the red area in figure 1. In the following we focus on the results from 15th of May to 13th of August 2019.

Ensemble performance

In figure 4 we compare only the first three members from MEPS and CMEPS. These members are the only ones running at the same hours and allows to examine the differences without any influence of lagging. It’s clear that CMEPS has a much smaller spinup with less excessive initial winds and precipitation, suggesting that the EDA based perturbations are more balanced than the ones based on SLAF used in MEPS. The bias signature for precipitation resembles the deterministic model for CMEPS whereas for cloud cover the spinup is larger.

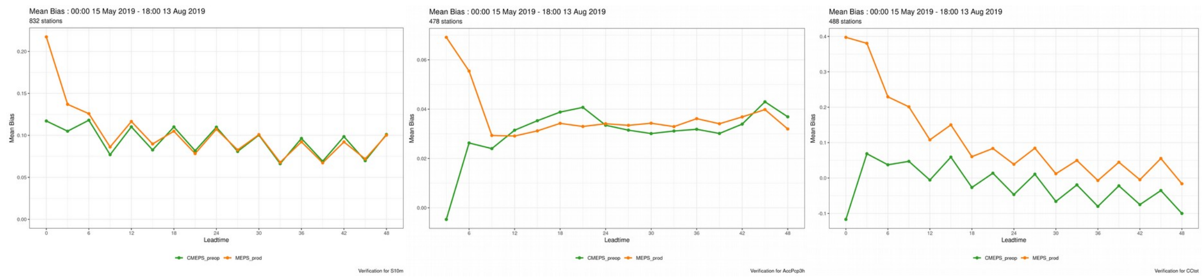

Figure 4: Bias for 10m wind, 3h precipitation and total cloud cover respectively. The comparison is done for member 0-2 only to avoid the influence of lagging. MEPS is shown in orange and CMEPS in green.

With the continuous approach we have the possibility to generate a new ensemble every hour. It’s therefore interesting to compare the maximum benefit of CMEPS compared to MEPS by including a 6h old version of MEPS in the verification. In figure 5 we see that for PMSL CMEPS is somewhere in between a fresh and an old MEPS. The penalty compared to a fresh MEPS comes from a combination of using older boundaries and from the larger error growth in IFSENS as compared to IFS HRES due to the lower resolution. This is in line with Parson et.al. (2020) and they suggest that a slightly larger ensemble is required in a continuous setup to compensate for this effect. The 10m wind speed shows an overall improvement although small. Both clouds and precipitation has an improvement in terms of CRPS although for precipitation the spread at longer lead times is a bit too large, figure 6.

In CMEPS we’ve managed to increase the spread for T2M although the impact on CRPS is not very large. RH2m is the only parameter where we see an overall degradation in CMEPS as compared to MEPS, figure 7. This is related to an increased dry bias in CMEPS as compared to MEPS, figure 8. In shorter sensitivity tests during the summer season, figure 8, we have seen that switching off the perturbation of soil moisture reduces the dry bias somewhat. However, switching off the perturbations for soil moisture over a 20 day period in December 2019 to January 2020 , figure 9, does not cure the dry bias compared to MEPS.

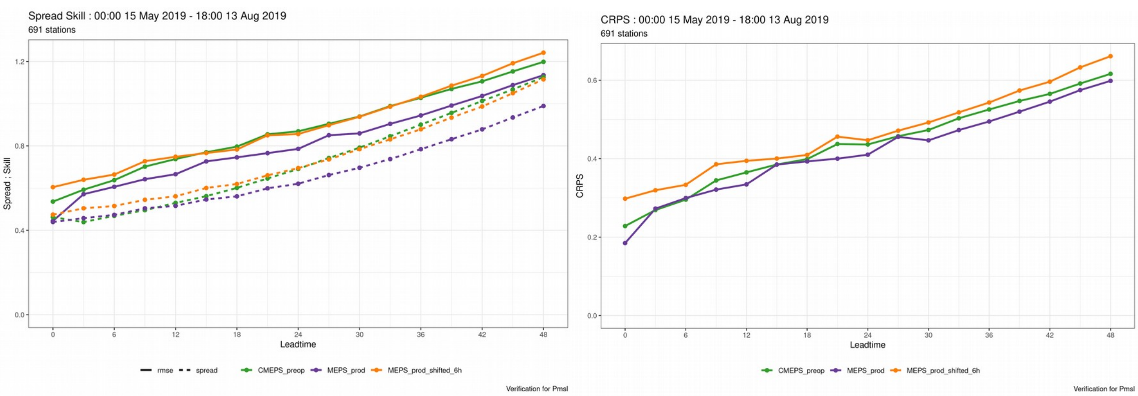

Figure 5: PMSL spread/skill (left) and CRPS(right) for CMEPS (green), MEPS (purple) and a 6h old MEPS (orange).

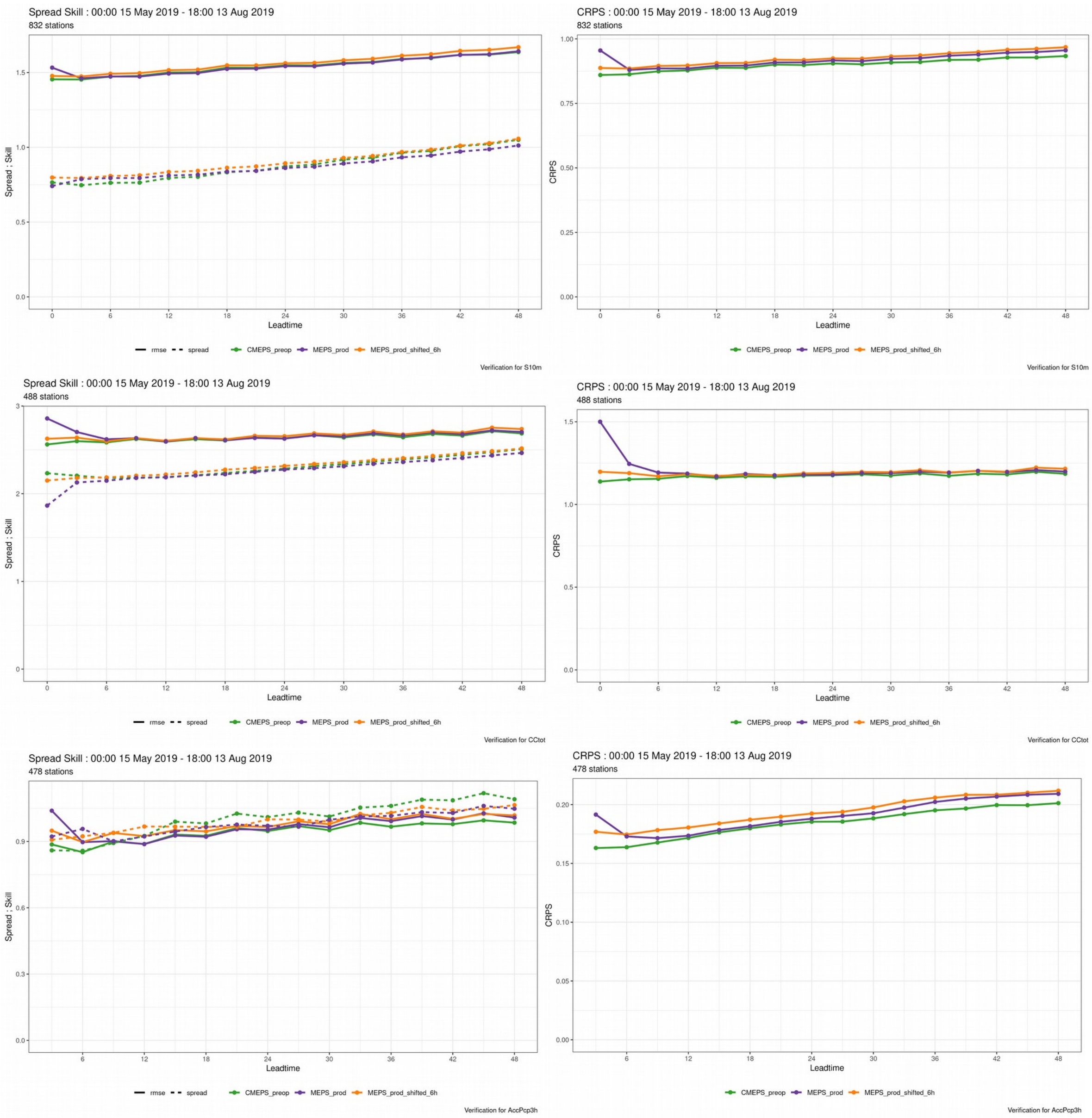

Figure 6: Spread/skill (left) and CRPS(right) for CMEPS (green), MEPS (purple) and a 6h old MEPS (orange). From top to bottom: 10m wind, total cloud cover and 3h precipitation.

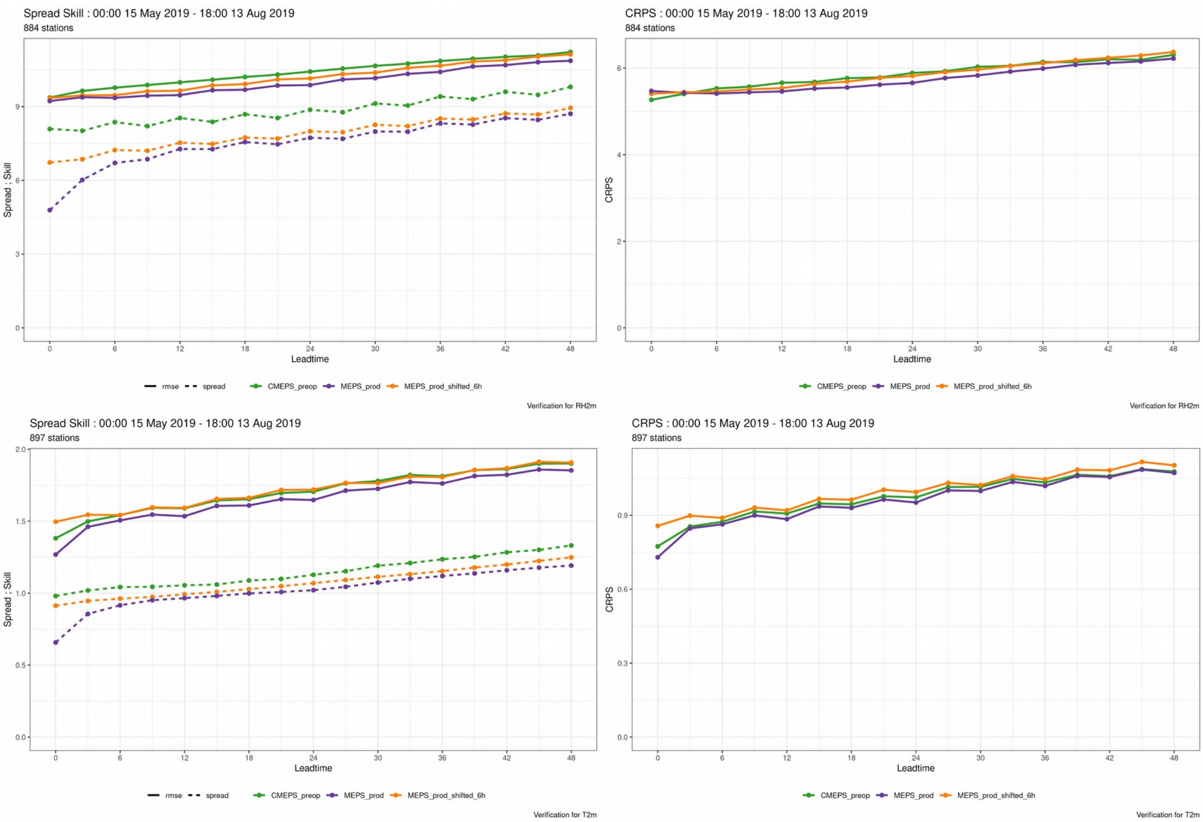

Figure 7: Spread/skill (left) and CRPS(right) for CMEPS (green), MEPS (purple) and a 6h old MEPS (orange). From top to bottom: RH2m and T2m.

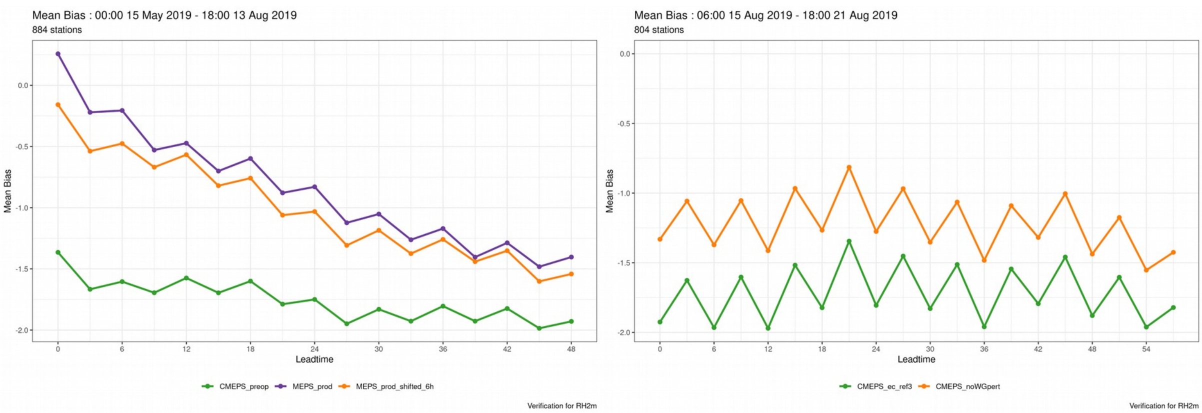

Figure 8: RH2m bias for CMEPS (green), MEPS (purple) and a 6h old MEPS (orange).

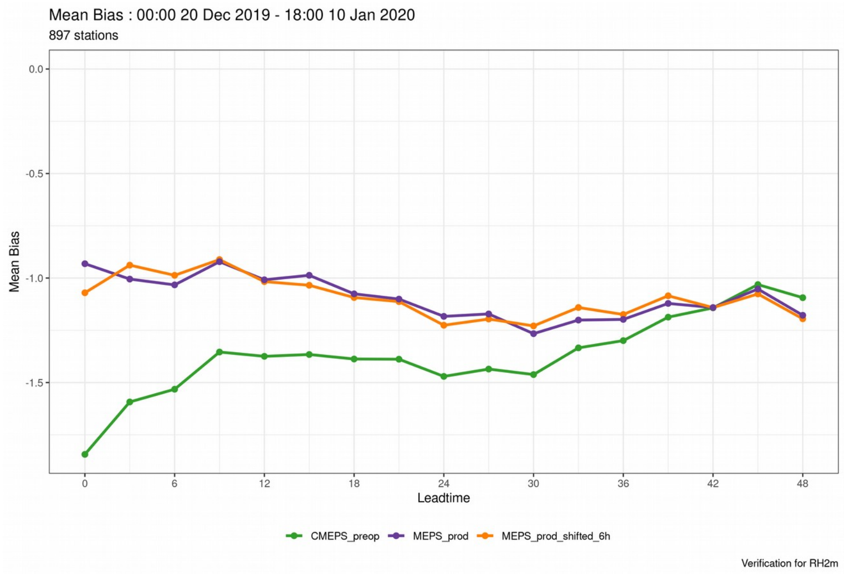

Figure 9: RH2m bias for CMEPS (green), MEPS (purple) and a 6h old MEPS (orange) for the 10 last days in December 2019.

Use CMEPS for generation of background errors

The background error statistics currently used for assimilation in MEPS are based on statistics derived from dedicated offline EDA experiments for a winter a summer season. Since then not only the forecast model characteristics have changed but also the model domain. Using data from the CMEPS archive is thus an appealing approach to generate new statistics. It would allow us to do more frequent updates without any extra cost. Sampling from 14 May 2019 to 29 Nov 2019 every 5 days and 3 times per day gives us 348 differences between pairs of archived 6h range forecasts. Comparing the spectral density with the currently used statistics we note that we have more energy on larger scales, figure 10. The result of a single temperature observation 1K warmer than the corresponding background value and with and observation error standard deviation of 1K placed in the center of the domain at 500 hPa (lev=24) is shown in figure 11. The response with the new structure function are slightly larger increments with a larger spatial scale. In addition to the pure EDA perturbations CMEPS has additional ones coming from the surface perturbations and the use of IFSENS member derived perturbations which may contribute to the larger scales. Further experimentation is required to see whether data from CMEPS is suitable for generation of background errors or not.

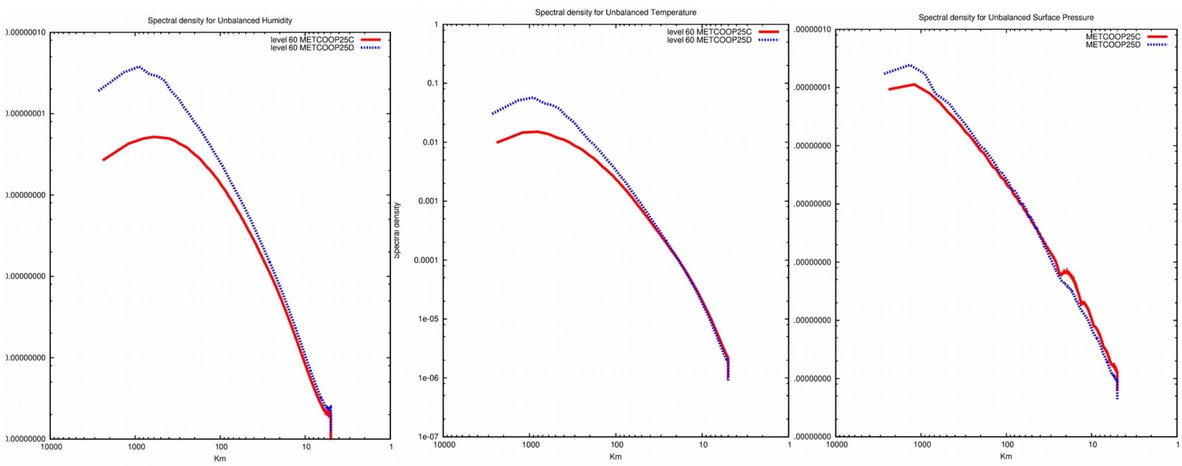

Figure 10. Spectral density functions for humidity, temperature and surface pressure. MEPS in red and CMEPS in blue.

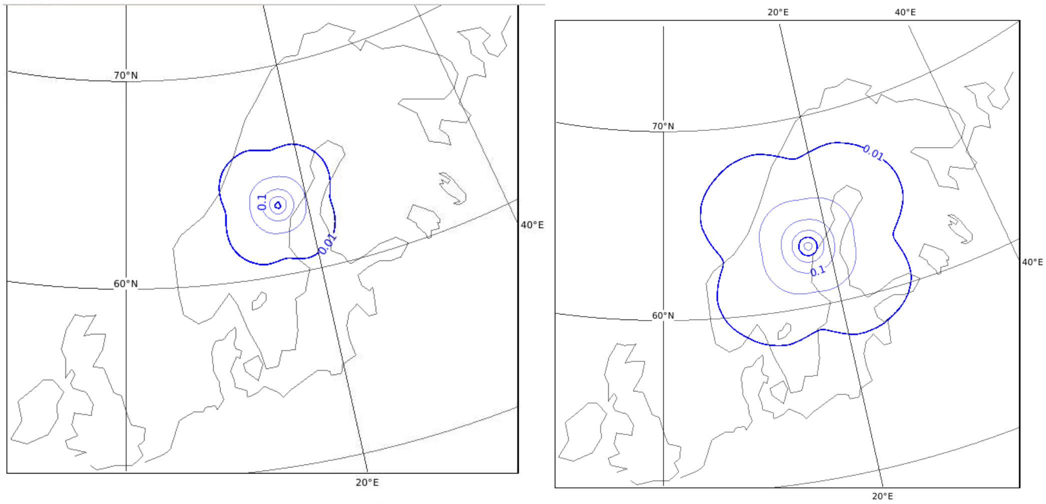

Figure 11. Impact of single obs experiment for old (left) and new (right) structure functions.

5 Conclusions and outlook

The MetCoOp ensemble have been redesigned to address two out of three mentioned shortcomings in the current setup. By a continuous EDA based approach driven by IFSENS on the boundaries we have been able to construct a three times larger ensemble with less spinup problems offering the forecasters hourly updates. There are however remaining problems such as a pronounced dry near surface bias for perturbed members. Neither is the lack of spread in cloud variables improved by the above mentioned changes and it's clear that a representation of the internal model uncertainty is still missing. In operational like setups HarmonEPS is implemented both with multiphysics and even multimodel (Frogner et. al. 2019b). Stochastically perturbed physical tendencies (SPPT) is also available and stochastic physics perturbations (SPP) is under development. The two latter options seems like the next natural steps to improve the uncertainty representation on short and longer development time scale. In Figure 12 the impact of adding model perturbations (SPP) on total cloud cover is shown for a one month long experiment in June 2019. The SPP experiment is compared to a reference experiment (black) using default HarmonEPS perturbations, SPP (orange) is the same except SPP perturbations are also used. SPP clearly improves the scores for total cloud cover, with higher spread and lower RMSE.

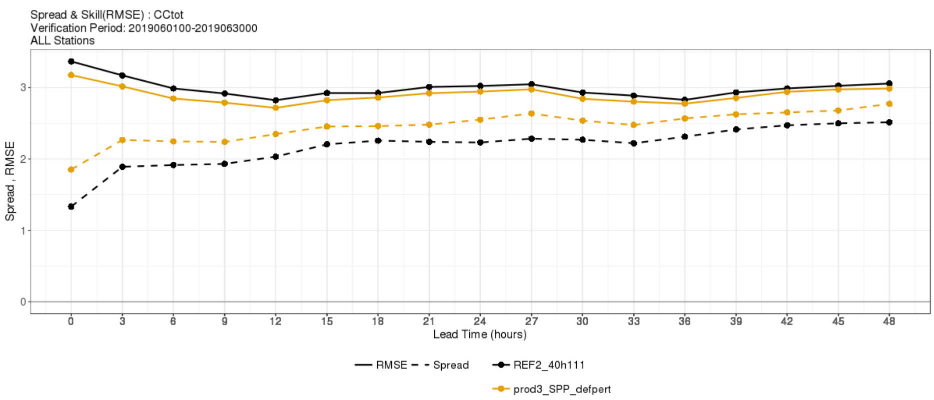

Figure 12: Spread and skill for total cloud cover for June 2019, reference experiment (black) and with SPP added (orange).

6 References

Andrae U., The MetCoOp ensemble MEPS, ALADIN-HIRLAM Newsletter No. 8, pp. 98–103, 2017. Available at: http://www.cnrm.meteo.fr/aladin/IMG/pdf/nl8.pdf.

Bengtsson, L., U. Andrae, T. Aspelien, Y. Batrak, J. Calvo, W. de Rooy, E. Gleeson, B. Hansen-Sass, M. Homleid, M. Hortal, K. Ivarsson, G. Lenderink, S. Niemelä, K.P. Nielsen, J. Onvlee, L. Rontu, P. Samuelsson, D.S. Muñoz, A. Subias, S. Tijm, V. Toll, X. Yang, and M.Ø. Køltzow, 2017: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System. *Mon. Wea. Rev.,* **145**, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1

Frogner, I-L, Singleton, AT, Køltzow, MØ, Andrae, U. Convection-permitting ensembles: Challenges related to their design and use. *Q J R Meteorol Soc*. 2019a; 145 (Suppl.1): 90–106. https://doi.org/10.1002/qj.3525

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

Hoffman, R.N. and Kalnay, E. (1983) Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus A*, 35A(2), 100–118. https://doi.org/10.1111/j.1600-0870.1983.tb00189.x.

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. Research Department, Tech. Memo. 636, 48 pp., https://www.ecmwf.int/en/elibrary/10125-ensemble-data-assimilations-ecmwf.

Molteni, F., R. Buizza, C. Marsigli, A. Montani, F. Nerozzi, and T. Paccagnella, 2001: A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments. *Quart. J. Roy. Meteor. Soc.*, **127**, 2069–2094, https://doi.org/10.1002/qj.49712757612.

Porson, A. et.al., 2020: Recent upgrades of the Met Office convective-scale ensemble: an hourly time-lagged 5-day ensemble, *Submitted to Quart.J. Roy. Meteor. Soc.*

Yang, X., Feddersen, H., Hansen Sass, B., Sattler, K., 2017: Construction of a continuous mesoscale EPS with timelagging and assimilation on overlapping windows, *ALADIN-HIRLAM Newsletter*, 8, 112-118