SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Model parameter perturbations in ensemble prediction
Computer Project Account:	spfiolli
Start Year - End Year :	2012 - 2014
Principal Investigator(s)	Dr. Pirkka Ollinaho
Affiliation/Address:	Finnish Meteorological Institute P.O. BOX 503 FI-00101 HELSINKI FINLAND
Other Researchers (Name/Affiliation):	

The following should cover the entire project duration.

Summary of project objectives

(10 lines max)

We test an ensemble prediction system which is added with a functionality to perturb model parameters, on top of the initial state perturbations. Our ensemble prediction system is built around the ECHAM5 atmospheric general circulation model, with initial states taken from the ECMWF operational EPS. The parameter perturbations are handled via a parameter estimation algorithm, the Ensemble Prediction and Parameter Estimation System (EPPES). Focus is given to parameters related to physical parameterizations of clouds and precipitation.

The aim is to better comprehend the role of parametric perturbations in ensemble forecasting, and perhaps to make statistical inference of their posterior probability distributions. Furthermore, existence of possible seasonal variations in optimal parameter values is explored.

Summary of problems encountered

(If you encountered any problems of a more technical nature, please describe them here.)

……………………………………………………………………………………………………

……………………………………………………………………………………………………

Experience with the Special Project framework

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

……………………………………………………………………………………………………

……………………………………………………………………………………………………

Summary of results

This study is focused on applying an algorithmic method for studying possible seasonal variations of Numerical Weather Prediction (NWP) model closure parameters. This problem is approached by using the Ensemble Prediction and Parameter Estimation System (EPPES; Järvinen et al. 2012; Laine et al. 2012) to research whether the optimal closure parameter values vary in time. The EPPES algorithm has been previously successfully used in finding optimal values for a subset of parameters (Ollinaho et al. 2013a, 2013b, 2013c). Applying EPPES in the context of in-time varying parameters, however, had not been explored before the start of this project. Thus, before the algorithm was applied to a real NWP model case, we first experimented on the algorithm with the Lorentz-95 model (Lorenz, 1995; Wilks, 2005). Even though the basic settings for EPPES were defined via this experimentation, considerable amount of Special Project computer resource was spend on fine tuning the algorithm with the real model case due to the non-linear nature of the problem.

We use ECHAM5 (Roeckner et al. 2003) general circulation model (GCM) with T63 resolution and 31 vertical levels for the real model case. The motivation of choosing ECHAM5 comes from familiarity with the model from our previous work (Ollinaho et al. 2013a, 2013c). The model is used in an Ensemble Prediction System (EPS) emulator where the initial state perturbations are taken from the operational ECMWF Ensemble Prediction System (ENS). A 50-member ensemble is run every 36 hours, starting from 1st of January 2011 and ending 20th of December 2011. Thus 11900 sample points are generated in total. The optimisation target is set to be the atmospheric total Energy Norm (EN; see Ollinaho et al. 2013c for the formulation) at forecast day three. In addition, the length of the time period affecting the distribution updates of EPPES is chosen to be two months, which means that systematic parameter variations shorter than this period are damped. The subset of parameters used in this study is given in Table 1.

Table 1 The subset of ECHAM5 closure parameters used in parameter variations.

Parameter	Description
CAULOC	A parameter influencing the accretion of cloud droplets by precipitation (rain formation in stratiform clouds)
CMFCTOP	Relative cloud mass flux at the level of non-buoyancy (in cumulus mass flux scheme)
CPRCON	A coefficient for determining conversion from cloud water to rain (in convective clouds)
ENTRSCV	Entrainment rate for shallow convection

Figure 1 illustrates time evolution of the four chosen parameters during one year of simulations. Parameter distribution mean μ (blue line), width defined as $\mu\pm2$x *standard deviation* (red lines), default parameter values (black line), and tested parameter values (grey column of markers; dark triangles indicate which parameters had impact on the distribution update) are shown. Even after considerable amount of time tuning the algorithm, it is quite evident that three of the parameters (CMFCTOP, CPRCON, ENTRSCV) have variations on timescales shorter than the damping period of two months. This was due to an internal forcing term of the algorithm set to too large value.

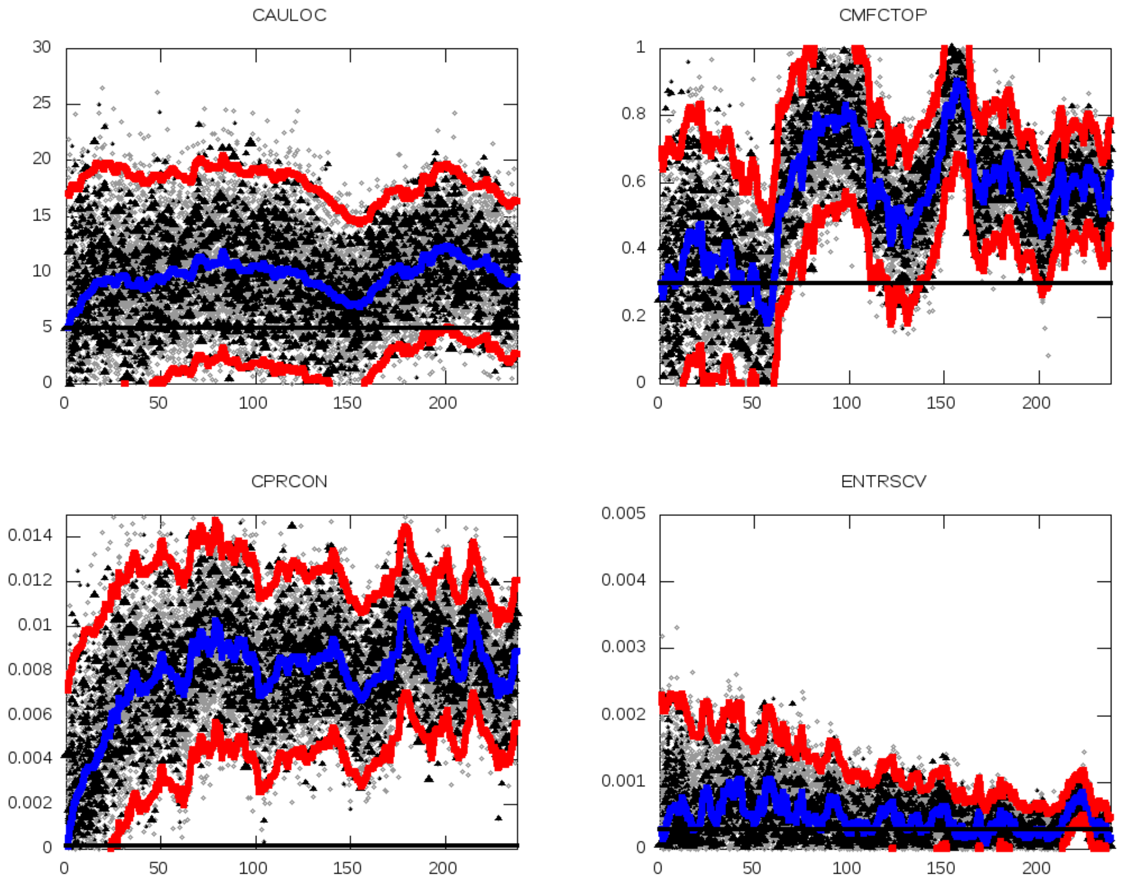

Figure 1 Time evolution of parameter distributions in 238 consecutive ensembles. A vertical column of markers represents parameter values of one ensemble. Values of high likelihood are indicated by black triangles. The parameter distribution mean μ (blue line), μ±2x standard deviation (red lines) and default parameter value (black line) are also shown.

Further fine tuning of the algorithm was done in order to lessen the effect of the internal forcing term in the parameter distribution variations. Figure 2 illustrates time evolution of the four chosen parameters in the modified algorithm experiment. The notation is the same as in Fig. 1. Of the parameter set, CAULOC and CPRCON show more stabile time evolution, and CAULOC values oscillate around the posterior mean value near the end of the sampling period. Values of ENTRSCV initially favour larger values, but converge back towards the initial values. On the other hand, CMFCTOP shows an approximately 120 day cycle in this sample.

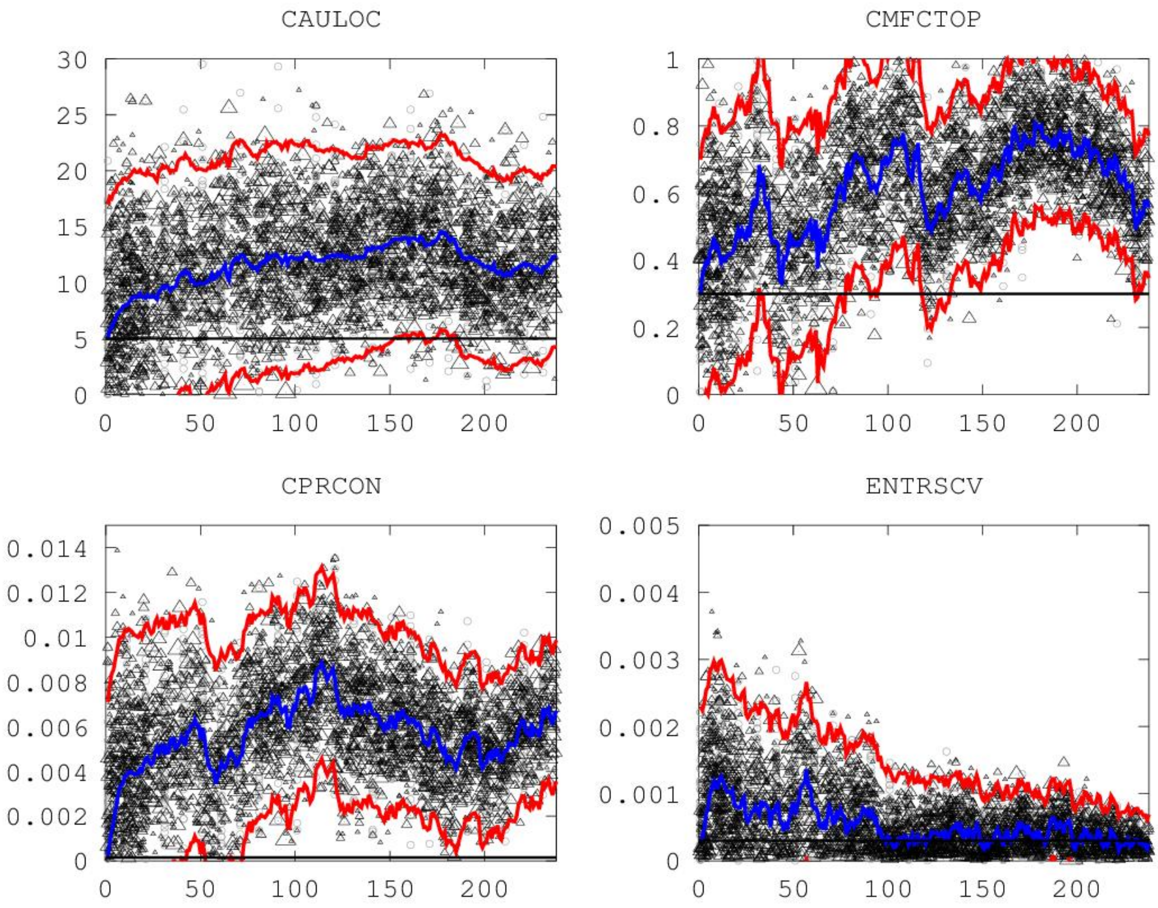

Figure 2 Time evolution of parameter distributions in 238 consecutive ensembles. A vertical column of markers represents parameter values of one ensemble. Values of high likelihood are indicated by black triangles. The parameter distribution mean μ (blue line), μ±2x standard deviation (red lines) and default parameter value (black line) are also shown.

Two validation runs were conducted covering the sampling period in order to explore the use of the parameter distribution evolution: In the first validation (Val1) the parameter posterior mean values (Table 2) were fixed as the parameter values. In the second validation run (Val2) parameter values were set to follow the evolving distribution mean with one day lag. This was done to see if the mean distribution values correspond to optimal parameter values also during the sampling period.

Table 2. ECHAM5 parameter values. Prior mean values correspond to the default model values. Posterior mean is the EPPES estimate after 238 estimation steps.

Parameter	Prior mean	Posterior mean
CAULOC	5.0	12.27
CMFCTOP	0.30	0.56
CPRCON	0.10×10^{-3}	6.70×10^{-3}
ENRSCV	0.30×10^{-3}	0.26×10^{-3}

Figure 3 shows the energy norm differences between model runs with the default and optimized parameters from Table 2 (Val1). Mean difference (continuous line) and 95% confidence bar of the mean value are shown for 10 day forecast range. The individual panels represent months in the sampling period, starting from January (top) and ending in June (bottom). All months show a statistically significant improvement for forecast lengths between 3.5 to 7 days. For most months this improvement can be observed up to forecast day 8. Forecast ranges beyond 8 days show a mostly neutral impact, but for April and June a statistically significant improvement is visible up to forecast day 10 (and likely beyond that).

The following six months (July to December) are represented in Figure 4. Again, a statistically significant improvement is visible for forecast lengths between 3.5 to 7 days. September shows an improvement up to forecast day 10, but not at 95% statistical significance level. For August and December the improvements are statistically significant up to the day 10 forecasts.

In Fig. 5 results for Val2 case are presented for January to June. Val2 results in a more mixed model response, and there is no clear statistically significant signal present. A positive impact can be observed from February to April, although the positive impact holds with 95% confidence level only for April.

In Figure 6 the following six months (July to December) are presented. Contrary to the first half of Val2 runs (Fig. 5), the latter part of the year shows a more positive overall impact: A statistically significant improvement is present for forecast lengths between 3.5 to 7 days. An improvement up to forecast day 10 is visible for September and October, although not at 95% statistical significance. For August and December the improvements are statistically significant up to the end of the forecasts.

Comparison of Val2 performance during the validation year indicates that the parameter values in the beginning of the parameter evolution are less skilful than in the end of the sampling period. Better scores in Val1 also suggest that the algorithm is able to find a more skilful parameter subspace near the end of the validation period. To confirm this an additional validation run was also conducted (not shown) where the whole sample was used to determine the optimal parameter values, i.e. an average of all the mean distribution values was used as optimal parameter set. Again the posterior mean (Val1) performed best. Note however that there is no degradation of EN scores for Val2 in January (Fig. 5) even though the EPPES algorithm is evolving the parameter distributions with large shifts to new regions. This indicates that there is no significant loss of model predictive skill coming from the larger parameter changes.

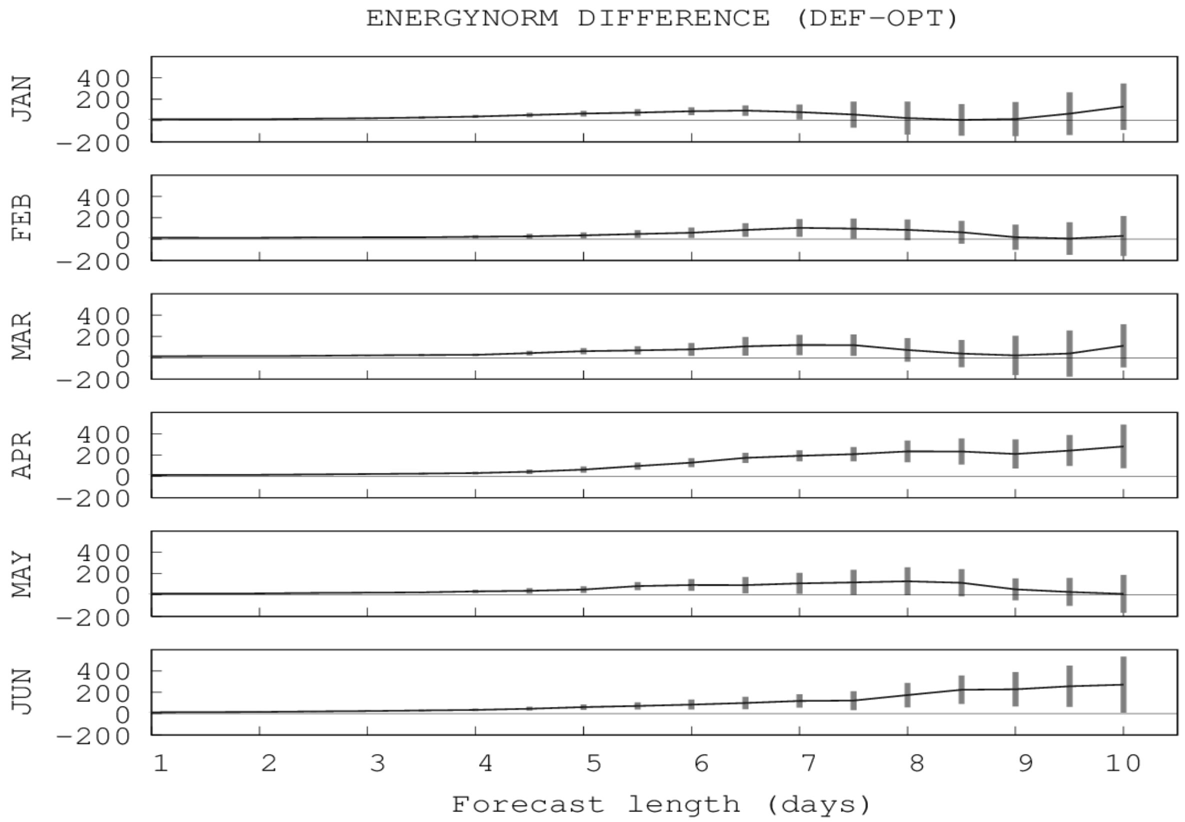

Figure 3 Energy norm differences between the default and optimized model for Val1-case. The rows represent individual months between January (top) and June (bottom). Mean forecast difference (continuous line), and the 95% confidence interval of the difference (vertical bars).

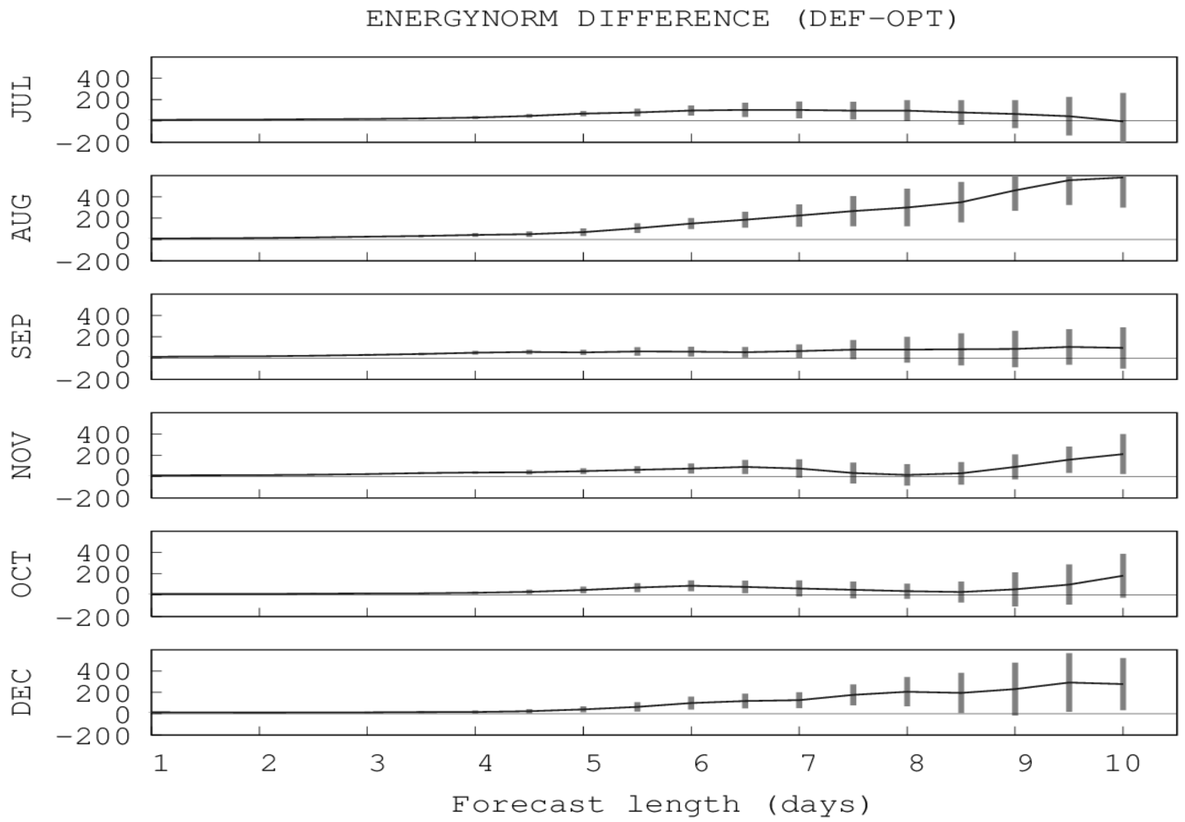

Figure 4 Energy norm differences between the default and optimized model for Val1-case. The rows represent individual months between July (top) and December (bottom). Mean forecast difference (continuous line), and the 95% confidence interval of the difference (vertical bars).

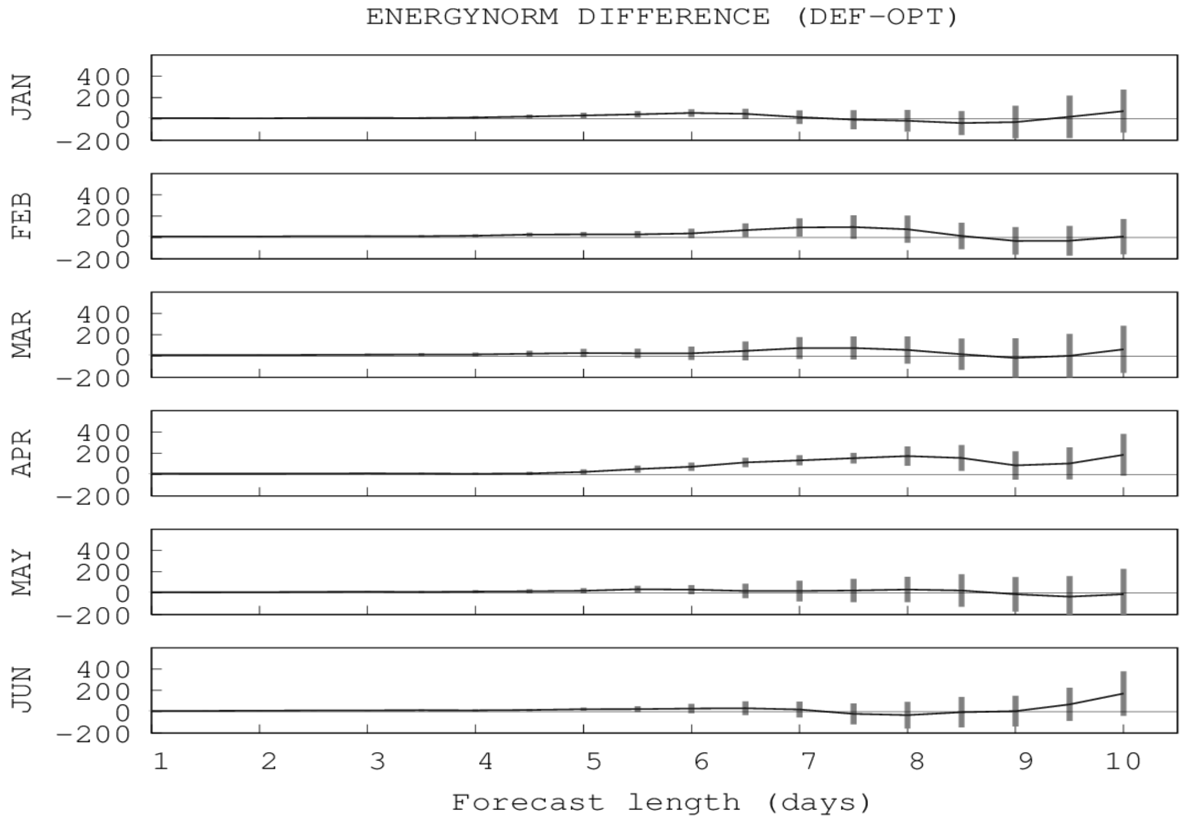

Figure 5 Energy norm differences between the default and optimized model for Val2-case. The rows represent individual months between January (top) and June (bottom). Mean forecast difference (continuous line), and the 95% confidence interval of the difference (vertical bars).

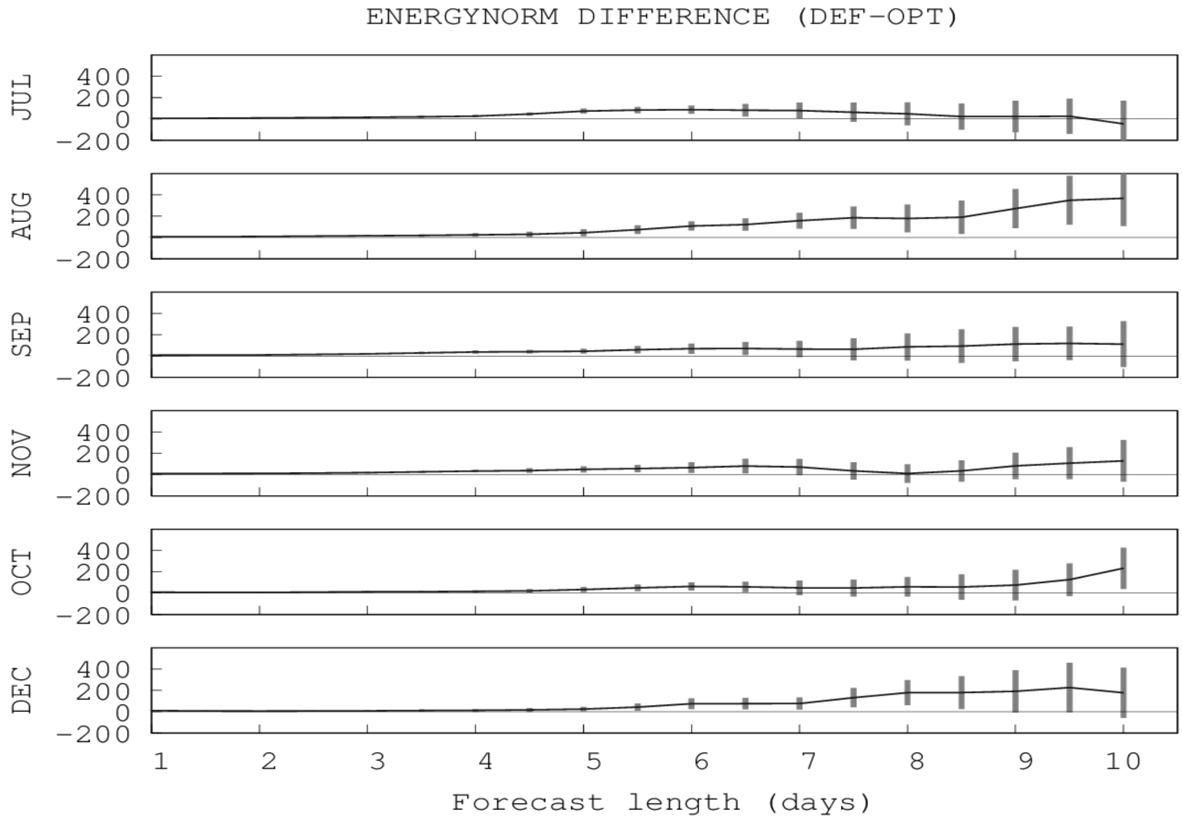

Figure 6 Energy norm differences between the default and optimized model for Val2-case. The rows represent individual months between July (top) and December (bottom). Mean forecast difference (continuous line), and the 95% confidence interval of the difference (vertical bars).

In this study seasonal variations in NWP model closure parameter values was studied. The hypothesis was studied by applying EPPES algorithm to estimate the optimal values of a subset of four closure parameter in ECHAM5 over a time period of one year. The summarizing results achieved through this ECMWF Special Project are:

i) Posterior mean values from EPPES parameter optimization lead to a more skilful forecast model, confirming what has been reported earlier by Ollinaho et al. (2013a, 2013b, 2013c).

ii) A more skilful model is also found when the model is ran with in-time varying closure parameter values which follow the evolving mean of the distribution from the EPPES estimation, although this improvement can only be observed in the later part of the validation period.

iii) The model using the final posterior mean values is more skilful than the one using in-time varying parameter values. This implies that either a) a fixed optimal parameter value is genuinely better for the average predictive skill of the model, i.e. to represent the varying conditions stemming from seasonal and weather regime changes, or b) as indicated by the very different signal in the scores of Val2 for the first and second half of the validation year, the algorithm is only able to find an optimal parameter subspace in the end of the estimation period. To verify this an additional year of EPPES estimation could be run starting from the posterior distribution. Additionally, it would be interesting to study would the Val2 scores improve if a smoothed mean was used instead of the now somewhat noisy distribution mean value. Lastly, in the progress of fine tuning the algorithm an additional one year long and 8-month estimation runs were conducted. The parameter distributions from these runs, or more specifically of the latter half a year of the estimation, could be used to see if simple data combination would result in more information about the system, and possibly lead to a parameter distribution which represents an even better model than that found now.

References:
Järvinen H, Laine M, Solonen A, Haario H. 2012. Ensemble prediction and parameter estimation system: The concept. Q. J. R. Meteorol. Soc. 138: 281–288

Laine M, Solonen A, Haario H, Järvinen H. 2012. Ensemble prediction and parameter estimation system: The method. Q. J. R. Meteorol. Soc. 138: 289–297.

Lorenz EN. 1995. 'Predictability: A problem partly solved.' Pp 1–18 in Proceedings of the Seminar on Predictability, Vol. 1, 4–8 September 1995, ECMWF, Reading, UK. Available at www.ecmwf.int/publications.

Ollinaho, P., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2013a. NWP model forecast skill optimization via closure parameter variations. Q.J.R. Meteorol. Soc., 139, 1520–1532, doi:10.1002/qj.2044.

Ollinaho, P., Bechtold, P., Leutbecher, M., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2013b. Parameter variations in prediction skill optimization at ECMWF. Nonlin. Processes Geophys., 20, 1001–1010, doi:10.5194/npg-20-1001-2013.

Ollinaho, P., Järvinen, H., Bauer, P., Laine, M., Bechtold, P., Susiluoto, J., and Haario, H., 2013c. Total energy norm in NWP closure parameter optimization. Geosci. Model Dev. Discuss., 6, 6717–6740, doi:10.5194/gmdd-6-6717-2013

Roeckner E, Bäuml G, Bonaventura L, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblueh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A. 2003. 'The atmospheric general circulation model ECHAM5, Part I: Model description.' Tech. Rep. No. 349, Max-Planck-Institut für Meteorologie.

Wilks DS. 2005. Effects of stochastic parametrizations in the Lorenz '96 system. Q. J. R. Meteorol. Soc. 131: 389–407.

List of publications/reports from the project with complete references

Manuscript in preparation, estimated time of publication 2016 (PI doing post-doc in 2015).

……

……

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.)

……

……