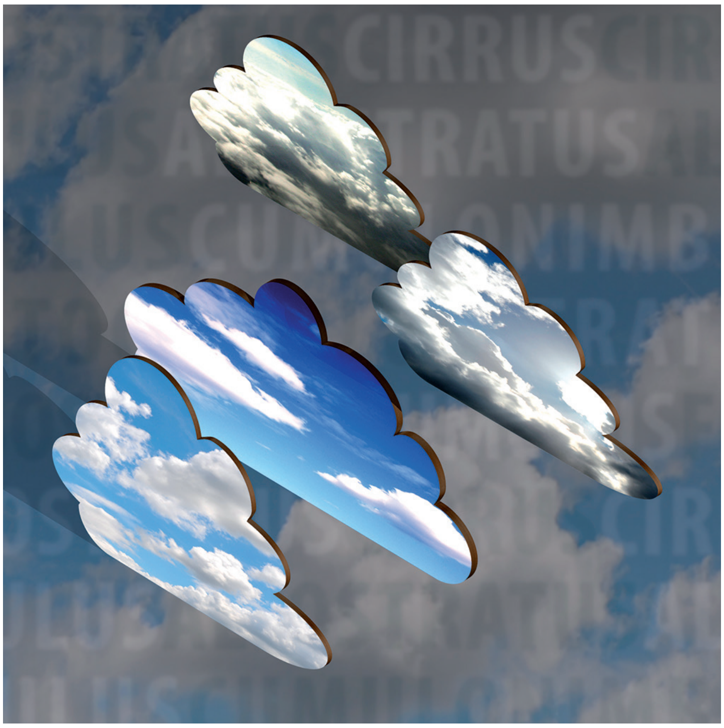
ECMWF Feature article

from Newsletter Number 129 – Autumn 2011

METEOROLOGY

Representing model uncertainty: stochastic parametrizations at ECMWF



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/fbqmkhv7

This article appeared in the Meteorology section of ECMWF Newsletter No. 129 – Autumn 2011, pp. 19–24.

Representing model uncertainty: stochastic parametrizations at ECMWF

Glenn Shutts, Martin Leutbecher, Antje Weisheimer,
Tim Stockdale, Lars Isaksen, Massimo Bonavita

When it was introduced, the ECMWF Ensemble Prediction System (EPS) was based on the assumption that errors in medium-range forecasts are mainly associated with errors in initial conditions. Later it was recognised that uncertainties in the model formulation may also be a significant factor. In particular, the physical parametrizations can be a significant source of random error. This led to the development of a stochastic representation of physical parametrization uncertainty (now known as the Stochastic Perturbed Parametrization Tendency (SPPT) scheme). The SPPT scheme has been used in the operational EPS since October 1998 – this version will be referred to as SPPT-98. Through the judicious application of random number multipliers to forecast tendencies, there has been an increase in ensemble spread in the EPS and improved probability skill scores (*Buizza et al.*, 1999).

Investigation of the performance of the EPS has shown that there would be benefits in enhancing the representation of model errors. This resulted in recent operational changes. The impacts of those changes are outlined in this article.

The Ensemble of Data Assimilations (EDA) was introduced in June 2010 to generate initial perturbations for the ensemble from ten independent 4D-Var assimilations with representations of observation and model errors (see, for example, *Bonavita*, 2011 and *Isaksen et al.*, 2010). EDA-based perturbations replaced the evolved singular vectors and the initial singular vectors were retained to ensure sufficient spread in the medium-range. In the EDA, SPPT plays an essential role in providing different realizations of the physics tendencies within each ensemble member. Importantly, EDA provides a means by which stochastic model error representations may be confronted with observational reality.

Improved representation of model errors

The SPPT scheme has been significantly revised in September 2009 and was further refined in November 2010. The revisions provided substantial improvements in ensemble spread: reduction in the error of the ensemble-mean and improved skill scores (e.g. Brier Skill Score and Continuous Ranked Probability Skill Score). These improvements followed from a more realistic version that uses a single spatially-smooth random pattern generator to perturb all parametrization tendency variables rather than independent, piecewise constant patterns for each variable as in the original scheme. The latest version of the SPPT scheme (November 2010) will be referred to as SPPT3 – for more detail see Box A.

Random error in physical parametrization is not the only source of model uncertainty so deficiencies in the dynamical component of the forecast model also need to be addressed. The upscale cascade of energy from sub-grid scales (or those scales that are poorly resolved in the model) is thought to be such a source of model error and considerable effort has been expended on formulating a Stochastic Kinetic Energy Backscatter (SKEB) scheme (also known as SPBS in earlier documentation; see *Shutts*, 2005 and *Berner et al.*, 2008).

The SKEB technique randomly forces vorticity perturbations into the model flow in such a way that the average energy input is a fraction of some measure of the local energy dissipation rate. Numerical dissipation and the implicit energy dissipation in the mountain wave drag and convection parametrization schemes are all regarded as sources of kinetic energy to be backscattered upscale. The SKEB scheme was introduced into the operational EPS in November 2010 to be used alongside the revised version of the SPPT scheme already implemented. More detail about the SKEB scheme is given in Box B.

doi:10.21957/fbqmkhv7

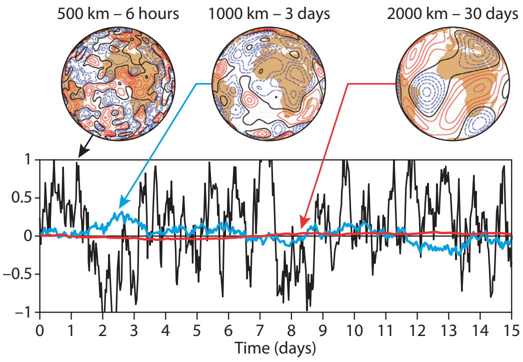
The Stochastic Perturbed Parametrization Tendency (SPPT) scheme

A

In the revised SPPT scheme the total physical parametrization tendency is multiplied by a randomly-evolving, global pattern field whose average value at any point is unity and whose standard deviation is prescribed. The pattern field is composed of three independent patterns, each generated from triangularly-truncated spherical harmonic expansions which have the property that their spatial auto-correlation function is independent of position on the sphere. Each spherical harmonic mode in each expansion is evolved in time using a first-order autoregressive process with fixed decorrelation time and wavenumber-dependent noise term. This three-pattern version of SPPT is referred to as SPPT3.

In the current operational implementation, the three patterns have quasi-Gaussian mean power spectra with horizontal correlations scales of 500 km, 1000 km and 2000 km, standard deviations of 0.52, 0.18 and 0.06, and decorrelation times of 6 hours, 3 days and 30 days respectively (see the figure in this box).

The above nine numbers characterizing SPPT3 are meant to span the uncertainty at mesoscale, synoptic scale and planetary space and time scales with pattern 1 (500 km decorrelation scale) being the starting point and most important component for the medium-range EPS. The other two patterns particularly improve the spread in seasonal forecast ensembles for which perturbations created using pattern 1 are insufficient. The decorrelation time of about 6 hours assumed in pattern 1 is loosely identified with a characteristic mesoscale time scale (e.g. for a mesoscale convective system). The longer decorrelation times used for the other two patterns in SPPT3 can be thought of as more persistent but smaller amplitude parameterization error that depends on the weather regime and thus exhibits variations on the medium-range to intra-seasonal timescales. Note that the standard deviation of the intermediate scale and large-scale patterns are much smaller (0.18 and 0.06 respectively) than that of the fastest pattern (0.52).



The three patterns underlying the SPPT3 scheme. The numbers next to the spheres indicate the horizontal spatial and temporal correlation scales in kilometres and hours. The three curves on the graph show time series of the pattern values at a point employed in the operational scheme The colour of the arrows relates the patterns to the time series.

Impact on medium-range and seasonal-range ensemble forecasts

The relative impacts of the SPPT3, SKEB and SPPT-98 schemes have been assessed using an operational configuration of the EPS (as in model cycle Cy36r4). The horizontal resolution is T639 up to day 10 and T319 thereafter. The ensemble is initialized with perturbations from the EDA combined with initial singular vector perturbations with a 50% reduced amplitude compared to the previous operational model cycle (Cy36r2). Each 15-day ensemble forecast has one control member and 50 perturbed members. The evaluation period consists of 19 equally-spaced dates in August/September 2008 and 21 equally-spaced dates in October/December 2009.

Figure 1 shows the root-mean-square (r.m.s.) error of the ensemble-mean 500 hPa geopotential and the ensemble spread (also measured as a root-mean-square) as a function of time for various representations of the model error: SPPT3+SKEB, SPPT3, SKEB and SPPT-98. Also shown are results from the EPS with only initial perturbations (labelled CONTROL) which acts as a baseline against which the effect of different model error representations can be judged. It can readily be seen that without accounting for model error, the EPS is under-spread (i.e. the ensemble spread is less than the r.ms. error of the ensemble mean). With SPPT-98 there is some small increase in spread and reduction in r.m.s. error but considerably less than both SPPT3 and SKEB. The combination of SPPT3 and SKEB gives a close match of spread to error.

The relative merits of the different model error choices are clearly seen in the Continuous Ignorance Score (CIS) with the lowest values representing the most skilful prediction (Figure 2). It is clear from Figure 1 that SKEB generates more spread than the SPPT-98 scheme and Figure 2 shows that this results in more skilful forecasts. Acting on their own, SKEB and SPPT3 give similar increases in spread over the 'no model error'

doi:10.21957/fbqmkhv7

case and yet SPPT3 seems to provide better improvements in the CIS. When combined with SPPT3, SKEB gives some modest additional reduction in CIS.

In the tropics, the EPS versions without SPPT3 are substantially under-spread and skill is low – presumably due to the inability of the forecast model to represent the interaction between parametrized convection and its local environment. Figure 3 shows that here, SKEB is less effective than the SPPT-98 scheme in generating spread in the 850 hPa temperature for the tropics. Presumably SKEB's wind forcing only generates weak temperature perturbations since the small Coriolis parameter there is unable to support balanced, horizontal temperature gradients.

Figure 3 shows, on the other hand, that SPPT3 is highly effective in generating spread and this results in substantially more skilful CIS (see Figure 4).

The Stochastic Kinetic Energy Backscatter (SKEB) scheme

B

As with SPPT3, the SKEB scheme is based on the product of a spectrally-generated pattern field and a derived model field. However, instead of using model tendencies, the backscatter scheme uses a horizontally-smoothed dissipation rate field to modulate the pattern field and defines this to be a streamfunction forcing function. The pattern uses a noise term with a different wavenumber dependence to that used in SPPT3 and one that gives a power law spectrum. This choice was determined by coarse-graining the streamfunction tendency (obtained from the u and v tendencies) in high-resolution forecasts and comparing with their counterparts in low resolution forecasts (see the section on coarse-graining).

Unlike SPPT3, the backscatter scheme allows for pattern variation with height and does this by randomly shifting the phase of each spectral mode using a first-order autoregressive process based on a Laplace probability distribution function. Again, coarse-graining results have been used to calibrate the dependence on pressure and wavenumber.

The kinetic energy dissipation rate field calculated here is not actually a true dissipation rate at all but is meant to provide an estimate of the sub-gridscale production of kinetic energy. For instance, the convective dissipation rate component is the product of the kinetic energy based on a vertically-averaged, updraught velocity multiplied by the convective mass flux detrainment rate – both terms being obtained from the convection parametrization scheme.

Parametrized mountain form drag and gravity wave drag remove energy from the forecast model yet some of this energy loss should go into sub-gridscale quasi-balanced eddies rather than turbulent energy dissipation. These eddies, although not represented explicitly, could interact with the resolved flow and cascade their energy upscale as a kind of backscatter process. Similarly, numerical dissipation of energy via explicit horizontal diffusion terms (or through the smoothing effect of interpolation in the semi-Lagrangian advection scheme) loses energy from the model without any relation to what should truly be dissipated into thermal energy. A certain fraction of this lost energy should therefore be backscattered to the resolved scales and this is what the SKEB scheme aims to do.

Thanks go to Martin Steinheimer (now at Austro Control GmbH, Vienna) for his substantial contribution to the development of SKEB.

doi:10.21957/fbqmkhv7

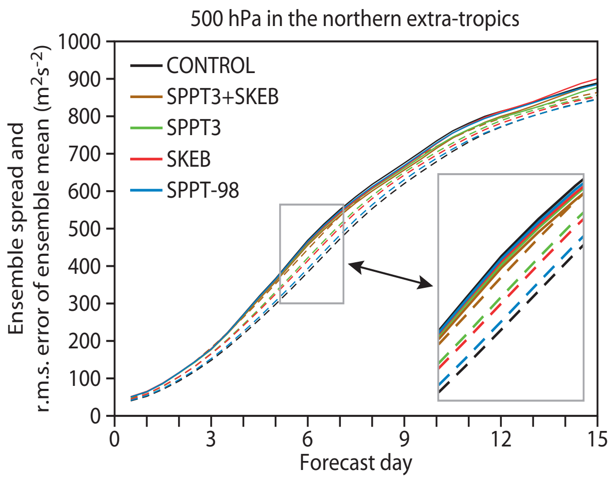


Figure 1 The r.m.s. error of the ensemble-mean (solid lines) and ensemble spread (dashed lines) versus forecast time (in days) for 500 hPa geopotential in the northern extra-tropics (20°–90°N) for various representations of model error: SPPT3+SKEB, SPPT3, SKEB and SPPT-98. Also results for the EPS with only initial perturbations are shown (CONTROL). All 40 EPS forecasts have 51 members and use initial perturbations from the Ensemble of Data Assimilations and initial singular vectors. The horizontal resolution of the forecasts is T639 and the number of model levels is 62. An expanded view of the marked rectangular region is shown for clarity.

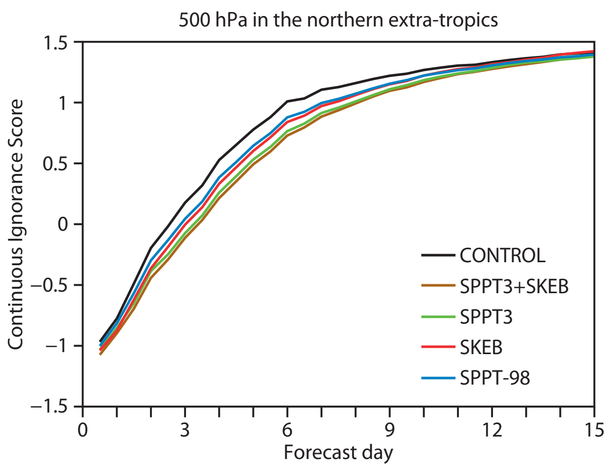


Figure 2 Continuous Ignorance Score for 500 hPa geopotential height in the northern extra-tropics for various representations of model error: SPPT3+SKEB, SPPT3, SKEB and SPPT-98 plus CONTROL. Note that the CIS is computed as the logarithmic score of the Gaussian distribution with mean and variance corresponding to the ensemble mean and ensemble variance.

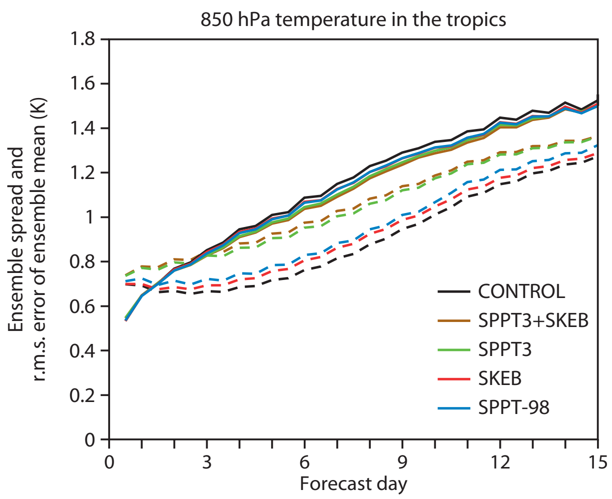


Figure 3 As Figure 1 but for temperature at 850 hPa in the tropics versus time.

doi:10.21957/fbqmkhv7

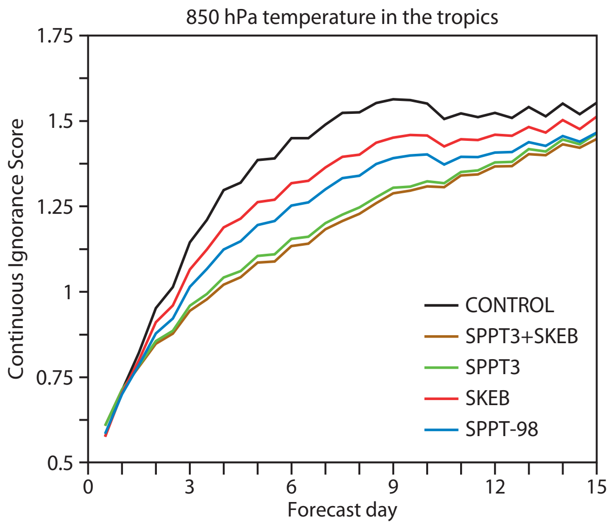


Figure 4 As Figure 2 but for temperature at 850 hPa in the tropics versus time.

Seasonal range

The impact of the new stochastic parametrization schemes has also been tested in ECMWF's seasonal-range, coupled ocean-atmosphere ensemble forecasting system. A set of retrospective ensemble forecasts with 11 ensemble members over the re-forecast period 1989–2005 has been carried out where 4-month long forecasts were initialised on 1 May and 1 November each year. The forecasts were made with a system that closely resembles the new Seasonal Forecast System 4 which is due to become the operational seasonal forecasting system at the end of 2011. The simulations were run with T255L91 resolution and used IFS Cycle 36r4, coupled to the 1°-NEMO ocean model.

The El Niño Southern Oscillation (ENSO) phenomenon is of crucial importance for seasonal forecasting and thus we focus our comparison on the performance of predicting tropical Pacific sea surface temperatures (SSTs), specifically for the Niño3 region (5°S–5°N, 150°W–90°W). Figure 5 shows the impact of the new schemes in forecasting SST anomalies in terms of the evolution of the ensemble-mean r.m.s. error (solid curves) and ensemble spread (dashed curves) over lead time. In Figure 5a, the simulations using the new model uncertainty representation (SPPT3 +SKEB) are shown along with the control simulation without any representation of model error (CONTROL). For comparison, the r.m.s. error of a simple persistence forecast is shown (PERSISTENCE). For a well-calibrated forecasting system one would expect that the ensemble-mean r.m.s. error would match the ensemble spread. This is clearly not the case for the simulations shown in Figure 5a. Here, the forecasting system is under-dispersive, or over-confident, by not generating enough ensemble spread. However, it can be seen that the stochastic tendency perturbation schemes have an overall positive effect on the problem of over-confidence by noticeably increasing the spread and slightly reducing the ensemble-mean r.m.s. error.

What are the relative contributions of SPPT3 and SKEB schemes to increasing ensemble spread and reducing the RMSE? Figure 5b shows results from ensemble forecasts where each scheme was switched on individually. As can be seen, the biggest impact in terms of spread and r.m.s. error comes from the SPPT3 scheme. The SKEB scheme also tends to increase the ensemble spread but to a much smaller extent.

Model uncertainty can be represented in different ways and stochastic physical parametrization is a newly-emerging field for long-range forecasts. The 'traditional' approach to address model uncertainty on seasonal and longer time-scales is the multi-model ensemble which relies on the assumption that individual models were developed quasi-independently. It is considered to be an 'ensemble-of-opportunity' for sampling model error. In the past, multi-model ensembles have been very successful in improving the skill of seasonal forecasts by reducing the over-confidence of the individual model ensembles. The ENSEMBLES multi-model ensemble (*Weisheimer et al.*, 2009), shown in Figure 5c, demonstrates the very good spread-skill relationship obtained by the multi-model approach.

Another method for modelling uncertainty uses ensemble forecasts with perturbed physical model parameters. For comparison, Figure 5d shows results from seasonal forecast experiments with perturbed parameters carried out in the ENSEMBLES project.

From Figure 5 it can be concluded that the stochastic physical parametrization provides a powerful alternative to other approaches for representing model uncertainty in seasonal forecasts and it is suggested that these schemes should now be developed for multi-decadal climate predictions using Earth System Models as well (*Weisheimer et al.*, 2011).

doi:10.21957/fbqmkhv7

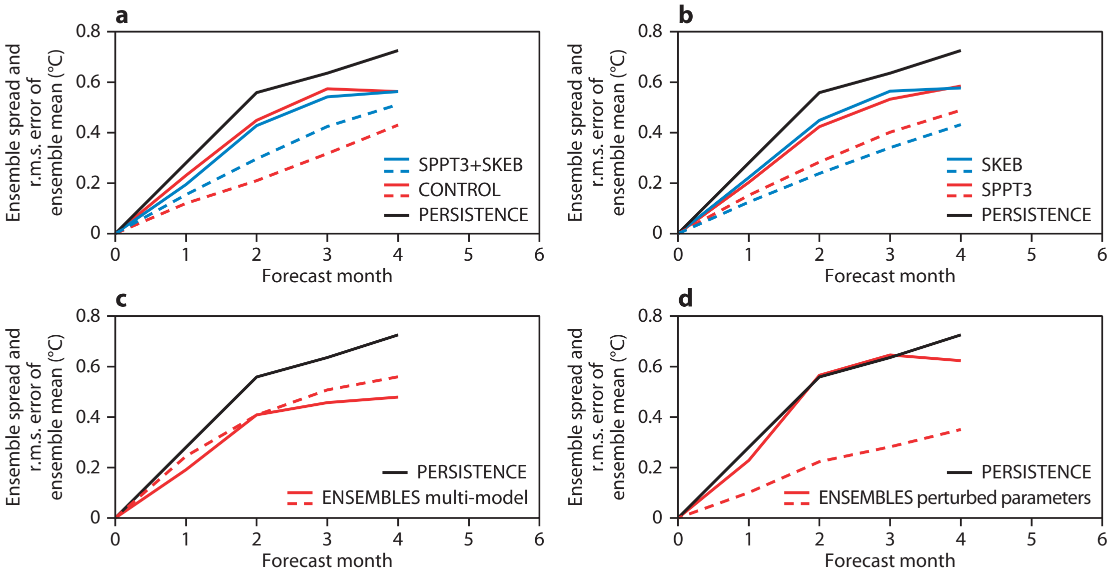


Figure 5 Quality of the SST anomaly forecast for the Niño3 region. Solid coloured lines indicate the r.m.s. error of the ensemble mean and dashed coloured lines show the ensemble standard deviation. The black lines show, as a reference, the r.m.s. error of a simple persistence forecast (PERSISTENCE). (a) Blue: SPPT3 + SKEB; red: CONTROL. (b) Blue: sensitivity simulations with SKEB only; red: sensitivity simulations with SPPT3 only. (c) Red: ENSEMBLES multi-model ensemble. (d) Red: ENSEMBLES perturbed parameters ensemble.

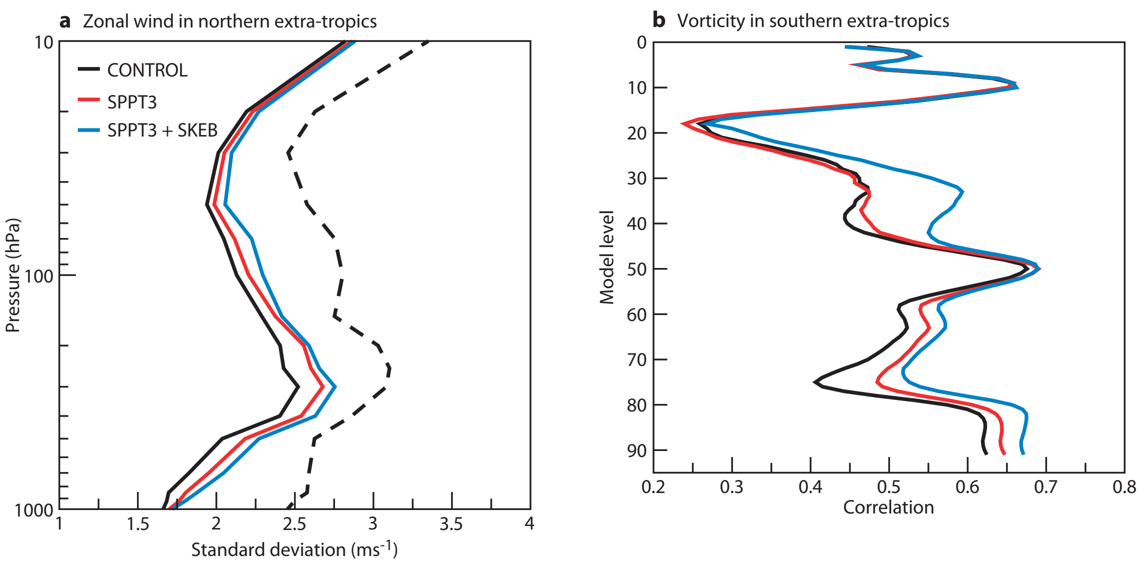


Figure 6 (a) Radiosonde innovation standard deviations for zonal wind in the northern extra-tropics (dashed line) and predicted innovation standard deviations for an EDA experiment without model error (CONTROL), with SPPT3 active (SPPT3) and with both SPPT3 and SKEB active (SPPT3+SKEB). (b) Spatial correlation of the EDA vorticity spread with the EDA mean background error vorticity field in the southern extra-tropics (model levels on the y axis).

doi:10.21957/fbqmkhv7

Model error parameterizations in the EDA

The ECMWF Ensemble of Data Assimilations (EDA) is a system of *N* (*N*=10 at the time of writing) independent, reduced-resolution, assimilation cycles which differ by using randomly-perturbed observations, sea-surface temperature fields and model physics tendencies. If the perturbations are drawn from the true distributions of observation and model error, then the spread of the EDA about the control (unperturbed) analysis will be representative of the analysis error (*Isaksen et al.*, 2010). The use of EDA perturbations has already proved to have a beneficial impact on the representation of initial uncertainties in the EPS and on the estimation of flow-dependent background errors in the deterministic 4D-Var assimilation system (*Isaksen et al.*, 2010).

The ability of the EDA to correctly capture the analysis and background errors of the reference analysis is based on an accurate representation of all the relevant sources of uncertainty in the deterministic analysis cycle, among which model error plays an important role. It is then important to evaluate how the different model error schemes affect the performance of the EDA. This is, in fact, a very stringent test of their ability to represent the true sources of model error because the effects of using a certain model error representation accumulate in time over the analysis cycles and they are confronted with the observational reality, both directly and through the EDA-sampled statistics used in the deterministic high-resolution analysis. A further distinction is that in an EPS context, one is typically concerned with the verification and use of univariate probability distributions at a given lead time and location, while one of the main uses of the EDA is to diagnose spatial and multivariate covariances. Finally, background errors (i.e. forecast errors at 12 hours lead times, in the present case) have been shown to span a much larger portion of the error space than errors at longer forecast lead times, since they have not collapsed yet on to the dominant modes of instability of the system. This obviously makes their estimation a more challenging problem.

Figure 6 shows two diagnostics of the impact of SPPT3 and SKEB on the EDA variances. Figure 6a compares the observed radiosonde innovation standard deviations for the zonal wind component in the northern extra-tropics with the expected innovations (square root of the sum of the EDA variance and observation error variance) for three different EDA systems: one with no model error representation (CONTROL), one with SPPT3 active (SPPT3), and one with both SPPT3 and SKEB active (SPPT3+SKEB). Since a statistically-consistent EDA should have matching observed and expected innovation standard deviations, it is apparent that the use of model error parametrizations improves the reliability characteristics of the EDA. This is confirmed by Figure 6 (b) which plots, as a function of model level, the spatial correlation coefficient of the EDA mean background error vorticity field with the corresponding EDA vorticity spread for the three EDA experiments.

The impact of the model error parametrizations on the EDA sample covariances is the subject of ongoing investigation. Preliminary results indicate that while the SPPT3 scheme has an overall neutral impact, the SKEB parametrization tends to slightly degrade the quality of the EDA covariances. This result, if confirmed, could be an indication that the spatially-correlated error structures introduced in the SKEB scheme in the EPS configuration are not appropriate for the estimation of background errors.

Improving the stochastic schemes by coarse-graining

Considerable effort is currently aimed at calibrating the schemes, or at the very least, providing some justification for the chosen parameters (e.g. like the standard deviation of the random pattern values about their mean value of unity). The coarse-graining method compares high- and low-resolution forecasts to infer the statistical character of tendency error in the low-resolution forecast, e.g. by coarse-graining operational T1279 and matching T159 forecasts.

Estimates made so far suggest a somewhat lower standard deviation than that currently assumed in SPPT3 although this is to be expected since a T1279 forecast parametrizes convection and much of the gravity wave spectrum. Coarse-graining using model data (cloud-resolving model and IFS forecasts) has also been used to determine the power spectrum of streamfunction forcing and the probability distribution function of vertical phase shifts in SKEB (*Palmer et al.*, 2009).

Current research is aimed at better targeting the uncertainty in physical parametrization and in backscatter. For instance it may be wrong to perturb the radiative temperature tendency in SPPT3 since the origin of radiative flux uncertainty lies in the representation of cloud principally. If we assume that the dominant uncertainty arises from spatial truncation, coarse-graining will be able to quantify the uncertainties associated with lower resolution versions of the model.

 doi:10.21957/fbqmkhv7

Future developments

SPPT3 is fairly straightforward to code in the forecast model whereas the SKEB scheme is complex and costly. In spite of its complex implementation, the approach upon which SKEB is based is quite crude and, for instance, there is no phase relationship between the streamfunction forcing and individual flow features. Work is underway to devise schemes which generate backscatter vorticity perturbations from the model's instantaneous vorticity field and in such a way that energy is more directly transferred to large-scale flow features. This type of 'negative viscosity' effect has been shown in recent studies by *Thuburn* (2011) which used the barotropic vorticity equation and has also been revealed by coarse-graining IFS forecasts.

Refinement of SPPT3 using the coarse-graining methodology is currently focused on assessing the uncertainty associated with individual processes and improved representation of the pattern generator. The assumption that the standard deviation of the perturbations is proportional to the magnitude of the tendency is currently under scrutiny and there is evidence from coarse-graining that the variance of the perturbations is proportional to the mean.

The performance of ensemble data assimilation with different formulations of random model error provides a more stringent test on their underlying physical basis than their impact on medium-range probability skill scores. Indeed it may even be possible to use EDA to determine optimal parameter settings in the stochastic algorithms. Ultimately, it would be desirable to have stochastic forcing formulations that work across all time scales from those of data assimilation to climate modelling. Only then can one be confident in the physical basis for the chosen model error representations.

Further reading

Berner, **J.**, **G.J. Shutts**, **M. Leutbecher** & **T.N. Palmer**, 2008: A spectral stochastic backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.

Bonavita, **M.**, 2011: Model Error in the ECMWF Ensemble of Data Assimilations. *In Proc. of the Workshop on 'Representing Model Uncertainty and Error in Weather and Climate Prediction'*, ECMWF, Reading, UK.

Buizza, **R.**, **M. Miller** & **T.N. Palmer**, 1999: Stochastic representation of model uncertainty in the ECMWF Ensemble Prediction System, *Q. J. R. Meteorol. Soc.*, **125**, 2887–2908.

Isaksen, **L.**, **J. Haseler**, **R. Buizza** & **M. Leutbecher**, 2010: The new Ensemble of Data Assimilations. *ECMWF Newsletter No. 123*, 17–21.

Palmer, **T.N.**, **R. Buizza**, **F. Doblas-Reyes**, **T. Jung**, **M. Leutbecher**, **G.J. Shutts**, **M. Steinheimer** & **A. Weisheimer**, 2009: Stochastic parametrization and model uncertainty. *ECMWF Tech. Memo No. 598*

Shutts, **G.J.**, 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.*, **131**, 3079–3102.

Thuburn, **J.**, 2011: Energy and enstrophy cascades in numerical models. *ECMWF Workshop on Representing Model Uncertainty and Error in Weather and Climate Prediction*, 20–24 June 2011, http://www.ecmwf.int/newsevents/meetings/workshops/2011/Model_uncertainty/presentations/Thuburn.pdf

Weisheimer, **A.**, **F.J. Doblas-Reyes**, **T.N. Palmer**, **A. Alessandri**, **A. Arribas**, **M. Deque**, **N. Keenlyside**, **M. MacVean**, **A. Navarra** & **P. Rogel**, 2009: ENSEMBLES – a new multi-model ensemble for seasonal-to-annual predictions: Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.*, **36**, L21711, doi:10.1029/2009GL040896.

Weisheimer, **A.**, **T.N. Palmer** & **F. Doblas-Reyes** (2011). Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.*, **38**, L16703, doi:10.1029/2011GL048123.

© Copyright 2016

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England

The content of this Newsletter article is available for use under a Creative Commons Attribution-Non-Commercial-No-Derivatives-4.0-Unported Licence. See the terms at https://creativecommons.org/licenses/by-nc-nd/4.0/.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.