ECMWF Feature article

from Newsletter Number 123 – Spring 2010

METEOROLOGY

Combined use of EDA- and SV-based perturbations in the EPS

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

doi:10.21957/codf7bpp

This article appeared in the Meteorology section of ECMWF Newsletter No. 123 – Spring 2010, pp. 22–28.

Combined use of EDA- and SV-based perturbations in the EPS

Roberto Buizza, Martin Leutbecher, Lars Isaksen, Jan Haseler

The simulation of initial uncertainties is one of the key aspects in ensemble prediction. At ECMWF, since the implementation of the first version of the Ensemble Prediction System (EPS) in 1992, these uncertainties have been simulated with singular vectors (SVs), perturbations characterized by the fastest growth, measured using a total energy norm, over a finite time interval.

With the forthcoming implementation in cycle 36r2 of an Ensemble of Data Assimilations (EDA, see the companion article in this edition of the *ECMWF Newsletter*, pages 17 to 21), the methodology used to generate the EPS initial perturbations will be changed. EDA-based perturbations will replace evolved singular vectors in the generation of the EPS initial conditions. Following this change, the EPS initial perturbations will have a better geographical and vertical coverage than in the earlier system. This results in a better spread-skill relationship in the early forecast range over the extra-tropics, and for the whole forecast range over the tropics. Limited-area ensemble prediction systems (e.g. COSMO-LEPS) that use EPS initial and boundary conditions will benefit from this improvement. Over the tropics the substantial increase of the EPS spread leads to much smaller spread under-dispersion. In terms of skill, the EDA-SVINI configuration of the EPS has a higher skill than the earlier SV-based system everywhere.

This article briefly describes the new EDA-SVINI implementation and discusses some results.

The old SV-based EPS

In the old SV-based system (Figure 1a), the EPS initial perturbations are generated using SVs growing over two different time periods:

- Evolved SVs (EVO) growing optimally during the 48 hours leading to the analysis time represent uncertainties that are likely to contribute most to analysis errors.
- Initial-time SVs optimally growing during the first 48 hours of the forecast (SVINI) sample directions in phase space likely to contribute most to forecast uncertainty.

Practically, the SVs are computed separately over the northern and the southern hemisphere extra-tropics, and for up to six local regions in the tropics to improve the geographical sampling of the initial uncertainties. The initial-time and evolved SVs for the different areas are re-scaled to have initial amplitude comparable to the analysis error estimate given by the high-resolution data-assimilation system. The background research that lead to ECMWF's SV-approach to simulating initial uncertainty is briefly summarized in Box A, and more detailed information about the configuration used to compute the SV component is given in Box B.

ECMWF's SV-approach to simulating initial uncertainty — A

The ECMWF SV-approach to simulating initial uncertainty using SVs was inspired by earlier work by, among others, *Lorenz* (1965) and *Farrell* (1990), who showed that these type of perturbations dominates the system dynamics over a finite time-interval. It is worth also quoting the work of *Ehrendorfer & Tribbia* (1997), who showed that if the objective of an ensemble system is the optimal prediction of the forecast error covariance matrix (optimal in the sense of maximum possible fraction of forecast error variance), then the singular vectors constructed using covariance information at the initial time constitute the most efficient means for predicting the forecast error covariance matrix. *Palmer et al.* (1998) discussed the issue of the impact of the norm definition on the SVs, and argued that the total energy norm used in the ECMWF system is a good approximation of the covariance matrix mentioned by Ehrendorfer & Tribbia.

 doi:10.21957/codf7bpp

The old EVO-SVINI configurations of the EPS

B

In the old EVO-SVINI configuration, the initial conditions are defined by adding to the unperturbed analysis a linear combination of initial-time and evolved SVs. To optimize the use of computational resources, SVs for any specific day/time are computed along a forecast trajectory, defined by the 6-to-54 hour forecast starting from a 6-hour earlier analysis (see *Leutbecher*, 2005 for more information).

For the EPS starting at time t, the initial-time SVs are the initial-time (t=0) SVs computed along a 6-to-54 hour forecast starting at t= –6 h. The evolved SVs are the final-time (t+48 h) SVs computed along a 6-to-54 h forecast starting at t= –54 h. The initial perturbations are symmetric, with the EPS even members having the opposite sign perturbation of the odd members. The coefficients that determined the linear combination and the amplitude of the SVs are computed using a Gaussian sampling method.

Over the extra-tropics (northern and southern hemispheres) the 50 leading initial-time and extra-tropical SVs are used, and over the tropics only the 5 leading initial-time SVs for each tropical target area are used (see *Leutbecher & Palmer*, 2008 for details). These perturbations are added only to the temperature and wind component of the model state vector, and to the surface pressure (no perturbations are added to the specific humidity or to any surface field). In the new EDA-SVINI configuration of the EPS only the initial-time SVs are used.

Replacement of the evolved SVs with EDA-based perturbations in the EPS

Buizza et al. (2008), who discuss in details the rationale behind the proposed change, have shown that replacing the evolved SVs with EDA-based initial perturbations leads to a better ensemble system. In the new EDA-SVINI configuration (Figure 1b):

- EDA-based perturbations are used instead of the evolved SVs to represent uncertainties that have been growing during the data assimilation cycles.
- Initial-time SVs optimally growing during the first 48 hours of the forecast (SVINI) were used to sample directions likely in phase space to contribute most to forecast uncertainty.

The EDA perturbed members are generated by (a) perturbing all observations and the sea-surface temperature field and (b) using the stochastically perturbed parametrization tendency (SPPT) scheme that perturbs the total parametrized tendency of physical processes to simulate random model error. More details on the EDA methodology can be found in the companion article by *Isaksen et al.* published in this edition of the *ECMWF Newsletter*.

In this new EDA-SVINI configuration, the EPS initial conditions are defined by adding to the unperturbed analysis an EDA-based perturbation and a linear combination of initial-time SVs. The initial-time SV component is identical to the one that was used in the old EVO-SVINI configuration except for a reduction of the amplitude by 10%. This reduction of the SVINI component is needed to achieve a better spread-skill relationship since the EDA-based perturbations have larger amplitude than the EVO component.

Each EDA-based initial perturbation is defined by the difference between one perturbed and the unperturbed 6-hour forecasts (the first-guess) started from the previous EDA cycle (i.e. from the EDA 4D-Var analyses run during the 12-hour period preceding the most recent analysis used to generate the unperturbed analysis). The EDA-based perturbations are symmetric, thus the EPS initial perturbations are still symmetric in the new configuration.

The reason why differences between 6-hour first guesses from the preceding EDA assimilation cycle are used instead of differences between analyses, is that the EDA suite runs with a 12-hour delayed mode to achieve a timely dissemination of the EPS products. Earlier experimentation compared ensemble forecasts using EDA-based perturbations defined by analyses and perturbations defined by 6-hour forecasts from the preceding cycle. Published (*Buizza et al.*, 2008) and recent results indicated that the use of 6-hour forecast perturbations instead of analyses perturbations does not degrade the probabilistic skill of the EPS. It should be pointed out that the EDA-based perturbations are also added to the upper-level specific humidity component of the unperturbed initial conditions (this variable was not perturbed in the old EVO-SVINI configuration). The reader is referred to *Buizza et al.* (2008) for a more detailed discussion of the similarity and differences between the old SV-based and the new EDA-based configurations.

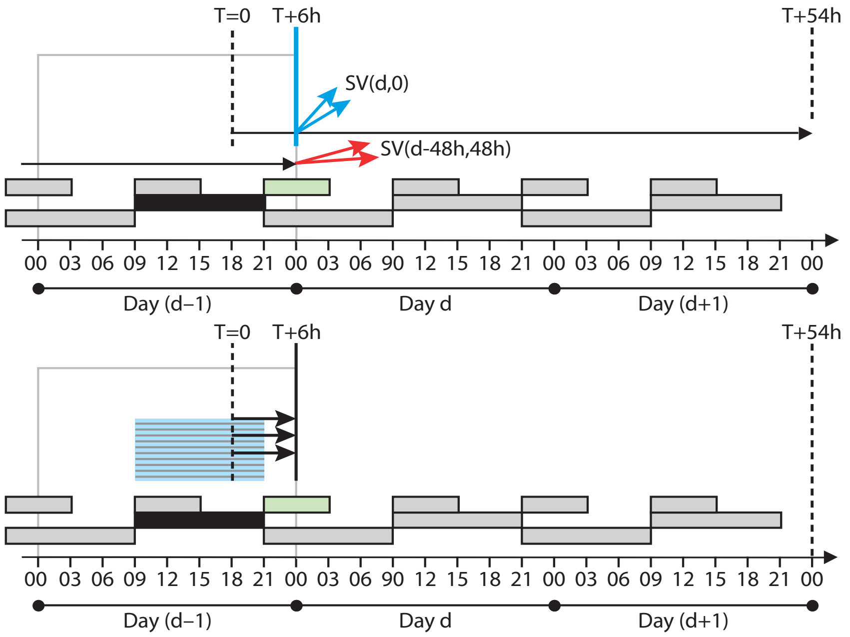

Figure 1 (from *Buizza et al.*, 2008). Schematic of the configuration used to generate the EPS initial conditions at 00 UTC. (a) The 12-hour-long black and grey boxes mark the time window of the 12-hour 4D-Var; while the 6-hour-long dark-grey and grey boxes for 21–03 UTC and 09–15 UTC mark the time window of the early-delivery 6-hour 4D-Var. The EPS unperturbed analysis at 00 UTC is defined by the 6-hour 4D-Var analysis generated by the early-delivery suite (green box). The evolved *SV(d-48h,48h)* (red vectors) are computed from 00 UTC of day *(d-2)*, and the initial-time *SV(d,0)* (blue vectors) are computed from 00 UTC of day d. The trajectory along which the *SV(d,0)* grow starts from the +6-hour forecast initiated at 18 UTC of day *(d-1)*, and the trajectory along which *SV(d-48h,48h)* grows starts from the +6-hour forecast initiated at 18 UTC of day *(d-3)*. (b) The EDA members used at day *d* (blue box with black lines) are generated by 12-hour 4D-Var cycles running between 09 UTC and 21 UTC of day *(d-1)*.

Replacement of the evolved SVs with EDA-based perturbations in the EPS re-forecast suite

Since March 2008, when the 15-day variable resolution ensemble was merged with the monthly prediction system, a key component of the EPS system has been the EPS re-forecast suite (*Hagedorn*, 2008; *Hagedorn et al.*, 2010). The EPS re-forecast suite is based on a 5-member ensemble starting from ERA-Interim analyses, and run for the same calendar day of the past 18 years. These 90 forecasts are used operationally to generate monthly anomaly products and to compute EPS products such as the Extreme Forecast Index (to be more precise, the EFI uses 450 forecasts, i.e. the 90 re-forecasts starting on the 5 weeks centred on the current day).

The new EDA-SVINI configuration will be used also for the re-forecast suite, but since the EDA has not been run for the past years, the re-forecast suite has to use the EDA-based perturbations computed for the current year. More precisely, since the re-forecasts are run up to 2-weeks in advance, the re-forecasts for any specific day for the past 18 years use the EDA-based perturbations from the current year minus 14 days. Extensive experimentation has indicated that the EDA-SVINI re-forecast EPS, despite using EDA perturbations from another year, has spread and skill characteristics closer to the real-time EPS than the old EVO-SVINI re-forecast EPS (not shown).

Comparison of the old EVO-SVINI and the new EDA-SVINI EPS configurations

It is worth noting that since 26 January 2010, the EPS has been running with a T639L62 (spectral triangular truncation at wave-number 639 with a linear grid and 62 vertical levels) resolution between day 0 and 10, and with T319L62 resolution from day 10 to day 15 (day 32 at 00 on Thursdays) – this is referred to as the 639v319 EPS. Table 1 summarizes the resolution of the key components of the EDA-SVINI configuration of the EPS that will become operational with model cycle 36r2 (the re-forecast suite has the same characteristics, but it includes only 4 instead of 50 perturbed members).

doi:10.21957/codf7bpp

Initial conditions			EPS forecasts day 0–15/32	
Unperturbed analysis	SV-based initial perturbations	EDA-based initial perturbations	Leg A (day 0–10)	Leg B (day 10–15/32)
T639L62 (interpolated from T1279L91 operational analysis)	T42L62	T399L62 analyses interpolated from the T399L91 analyses	T639L62	T319L62

Table 1 Resolution of the components used to generate the EPS initial conditions and produce the EPS forecasts since 26 January 2010.

EPS initial perturbations

The replacement of the evolved SVs with EDA-based perturbations has a large impact on the EPS initial perturbations As an example, Figure 2 shows the initial perturbations in terms of temperature and the zonal wind component at 700 hPa for one randomly-chosen member, member number 5, of the EVO-SVINI (left panels) and the EDA-SVINI (right panels) ensembles started on 1 December 2009. Also Figure 3 shows the corresponding results for a vertical cross section. The EDA-SVINI perturbations are less localized geographically and in the vertical and provide a better coverage of the globe. This is true especially over the tropics which were sampled (by design) only in a limited fashion by the initial perturbations of the old EVO-SVINI system.

The EDA-based perturbations, computed at T399 resolution, have smaller scales than T42 SV-based perturbations and a less evident vertical tilt with height. In addition, the EDA-based perturbations grow slower than SV-based perturbations, which explains why an ensemble with only EDA-based initial perturbations would have too little spread and a poorer performance than an EDA-SVINI ensemble (see *Buizza et al.*, 2008). By contrast, the blend of EDA-based perturbations and the initial-time SVs combines the benefits of both sets of diverse perturbations, and provides a superior performance to EVO-SVINI.

The effect of the modified initial perturbations is detectable over the extra-tropics during the first 48 hours and over the tropics for the first week. The EDA-SVINI perturbations start with a larger initial amplitude than the EVO-SVINI perturbations but after 24 hours their amplitude is close to the amplitude of the EVO-SVINI perturbations, especially over the extra-tropical regions where the SVINI component starts dominating the perturbation growth (Figure 4). This is the reason why from this forecast time on, in the regions sampled by initial-time SVs, the perturbations from the two ensembles have similar structures (e.g. south-east of Greenland (20°W, 60°N) or south-east of Cuba (60°W, 10°N)). But in the regions not sampled by the initial-time SVs (e.g. over most of the tropics) the EDA-SVINI initial perturbations provide a better geographical coverage. After 48 hours, over the extra-tropics the difference becomes, on average, smaller and gradually disappears in the medium-range (say around forecast day 7).

Member States' users of EPS initial and boundary conditions for limited area ensemble prediction systems over Europe will benefit from the increased spread over the extra-tropics in the early forecast range. This can be seen in Figure 5; this example shows the ensemble spread at initial time and at T+12 and T+24 hours for the EPS started at 12 UTC on 11 December 2009. The EDA-SVINI has a larger initial spread at 20°W where the EDA perturbed analyses differ slightly in the positioning and intensification of a low-pressure system, which propagates in time and leads to larger spread at T+24 hours west of Spain.

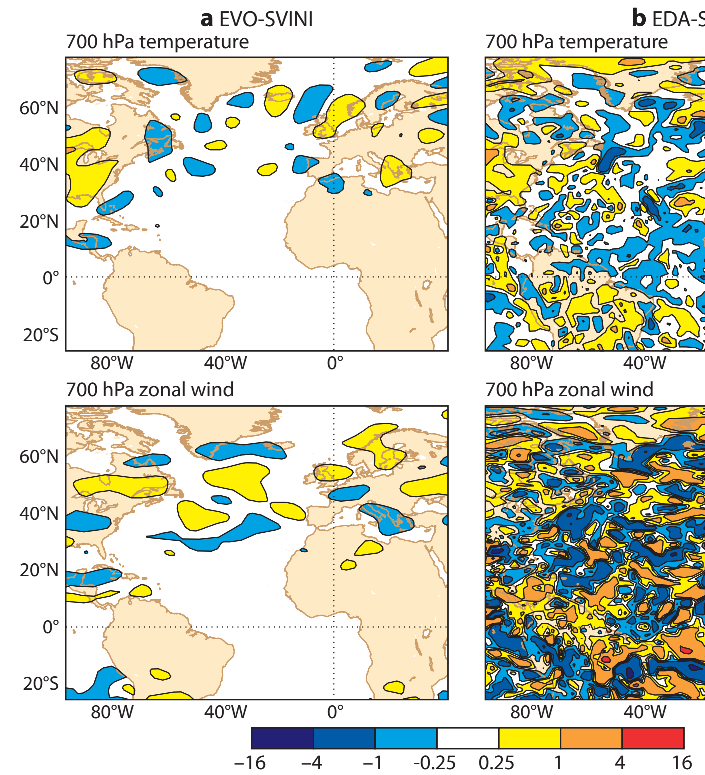

Figure 2 Initial-time perturbation of EPS member number 5 of (a) EVO-SVINI and (b) EDA-SVINI started at 12 UTC on 1 December 2009, in terms of the 700 hPa temperature (top panels) and the 700 hPa zonal wind (bottom panels).

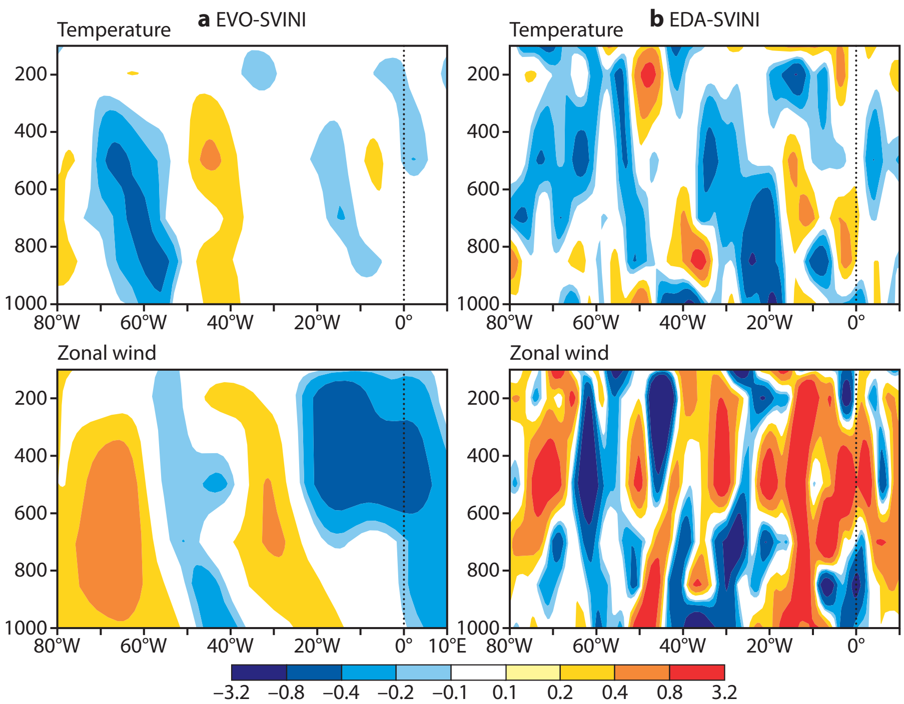

Figure 3 As Figure 2 but for a vertical cross section at latitude 50°N between 80°W and 10°E of the perturbation of EPS member number 5 of (a) EVO-SVINI and (b) EDA-SVINI at 12 UTC on 1 December 2009, in terms of temperature (top panels) and zonal wind (bottom panels).

doi:10.21957/codf7bpp

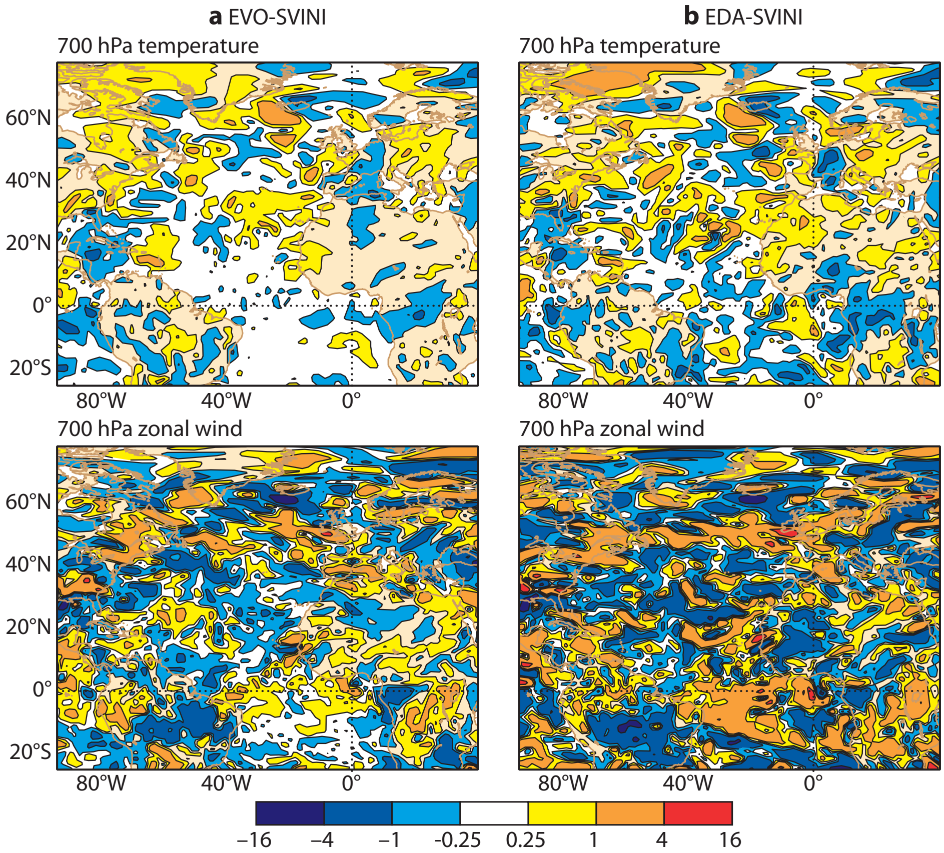

Figure 4 As Figure 2 but for the t+24 hour perturbation of EPS member number 5 of (a) EVO-SVINI and (b) EDA-SVINI started at 12 UTC on 1 December 2009, in terms of the 700 hPa temperature (top panels) and the 700 hPa zonal wind (bottom panels).

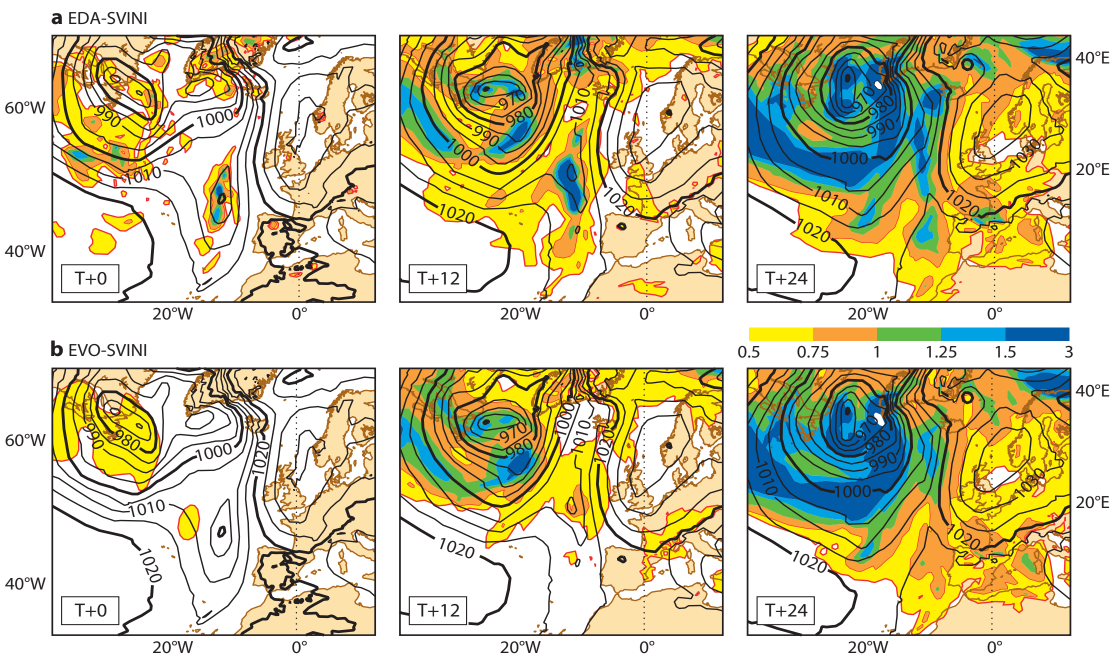

Figure 5 Ensemble-mean (black contours) and standard deviation (coloured shading) in terms of mean-sea-level-pressure (MSLP) of (a, top panels) EDA-SVINI and (b, bottom panels) EVO-SVINI ensembles at initial time (left panels), at T+12 hour (middle panels) and at T+24 hour (right panels) for EPS forecasts started on 11 December 2009. The contour interval for the ensemble-mean fields is 5 hPa; the shading for the standard deviation is for 0.5, 0.75, 1.0, 1.25, 1.50 and 3 hPa.

EPS spread and skill characteristics

The difference in the characteristics (amplitude, scale, growth rate, coverage) of the initial-time perturbations affects the average ensemble spread, as can be seen in Figure 6. This shows the 10-day average (forecasts with initial date from 1 to 19 December 2009, every other day) spread of the EVO-SVINI and the EDA-SVINI ensembles in terms of the 850 hPa temperature and the 700 hPa kinetic energy. The EVO-SVINI initial perturbations (left panels) are much more localized and have a smaller amplitude than the EDA-SVINI initial perturbations (right panels).

Figure 7 shows the average impact on 639v319 ensemble forecasts of the 850 hPa temperature, based on the comparison of EVO-SVINI and EDA-SVINI initial perturbations for 88 cases (from 5 October to 31 December 2009). As already mentioned in the introduction, in the EDA-SVINI ensemble, the amplitude of the SVINI perturbations has been decreased by 10% to improve the spread and skill relationship.

The EDA-SVINI ensemble has, on average, a better-tuned ensemble spread and a higher skill. Over the extra-tropics, there is a clear increase in ensemble spread during the first 48 hours: this reduces the spread under-estimation of the old ensemble system by about 50%. There is also a small positive impact on the error of the ensemble-mean and on the skill of probabilistic scores measured by the continuous rank probability skill score (CRPSS): although small, differences are statistically significant at the 5% level up to forecast day 6 over the northern hemisphere, and up to forecast day 10 over the southern hemisphere.

The positive impact on the ensemble spread and skill is more evident over the tropics, where the use of the EDA-based perturbations has a large impact on the ensemble spread. Over this region, the EDA-SVINI ensemble-mean has a smaller root-mean-square-error that is statistically significant (at the 5% level) and the probabilistic forecast has a higher continuous ranked probability skill score up to forecast day 9.

Similar conclusions can be drawn from other probabilistic accuracy measures, such as the area under the relative operating characteristic curve or the Brier skill score (not shown).

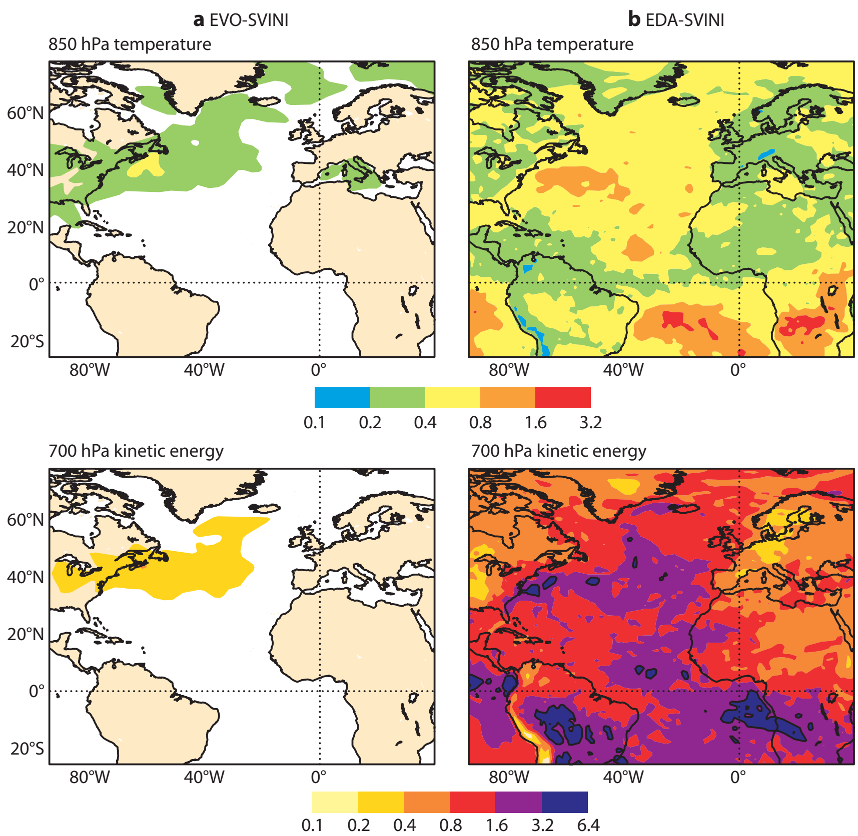

Figure 6 Ensemble spread: 10-case average (1 to 19 December 2009, every 2 days) standard deviation at initial time of an (a) EVO-SVINI ensemble and (b) EDA-SVINI ensemble, measured in terms of the 850 hPa temperature (top panels, K) and the 700 hPa kinetic energy (lower panels, ms^{-1}).

doi:10.21957/codf7bpp

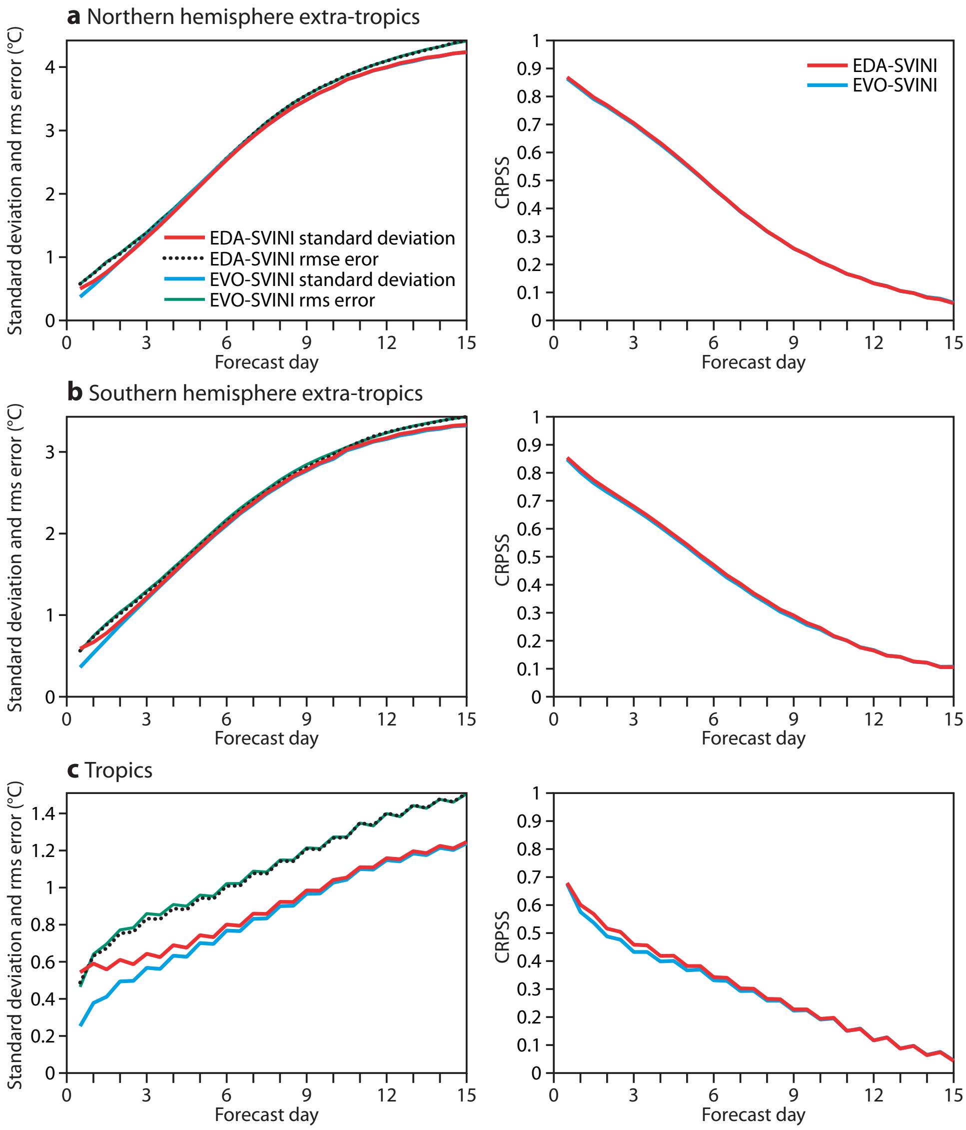

Figure 7 Average (88 cases from 5 October to 31 December 2009) statistics for the 850 hPa temperature over (a) northern hemisphere extra-tropics, (b) southern hemisphere extra-tropics, and (c) tropics. Left: root-mean-square error of the ensemble-mean forecast of the EDA-SVINI and the EVO-SVINI ensembles, and standard deviation of the EDA-SVINI and the EVO-SVINI line) ensembles. Right: continuous rank probability skill score (CRPSS) of the EDA-SVINI and the EVO-SVINI ensembles.

Future developments of the EPS

Work will continue in four key areas to further improve the skill of the EPS:

- *EDA membership:* the potential benefit of using a larger ensemble of perturbed analyses (25 or 50 instead of 10) will be assessed.
- *EDA-based land-surface perturbations:* spread of the EPS in the boundary layer and for surface variables will be assessed more thoroughly, and the potential use of EDA-based perturbations to perturb surface variables (e.g. soil moisture and soil temperature) will be investigated.
- *Combination of EDA- and SV-based perturbations:* possible ways to combine the EDA- and the SV-based perturbations different from the one implemented in the EPS will be explored, with the final aim to provide a better tuned and more skilful ensemble system for the entire forecast range and for the whole vertical structure of the atmosphere.
- *Stochastic model error:* revised and new stochastic schemes are under final development and will be tested in the EPS and the EDA to improve the simulation of model uncertainties.

Progress in these areas will be reported in due course.

Futher Reading

Buizza, **R.**, **M. Leutbecher** & **L. Isaksen**, 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **134**, 2051–2066.

Ehrendorfer, **M.** & **J.J. Tribbia**, 1997: Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.*, **54**, 286–313.

Farrell, **B.F.**, 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, **47**, 2409–2416.

Hagedorn, **R.**, 2008: Using the ECMWF re-forecast dataset to calibrate EPS forecasts. *ECMWF Newsletter No. 117*, 8–13.

Hagedorn, **R.**, **R. Buizza**, **M.T. Hamill**, **M. Leutbecher** & **T.N. Palmer**, 2010: Comparing TIGGE multi-model forecasts with re-forecast calibrated ECMWF ensemble forecasts. *Mon. Wea. Rev.*, submitted.

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

Leutbecher, **M.**, 2005: On ensemble prediction using singular vectors started from forecasts. *Mon. Wea. Rev.*, **133**, 3038–3046.

Leutbecher, **M.** & **T.N. Palmer**, 2008: Ensemble forecasting. *J. Comp. Phys.*, **227**, 3515–3539.

Lorenz, **E.**, 1964: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333.

Palmer, **T.N.**, **R. Gelaro**, **J. Barkmeijer** & **R. Buizza**, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633–653.

 doi:10.21957/codf7bpp