

Relevance of climatological background error statistics for mesoscale data assimilation

By JELENA BOJAROVA*, and NILS GUSTAFSSON, *Research Department, Swedish Meteorological and Hydrological Institute, Norrköping, SE-60176, Sweden*

(Manuscript received 25 October 2018; in final form 1 May 2019)

ABSTRACT

The relevance of climatological background error statistics for mesoscale data assimilation has been investigated with regard to basic assumptions and also with regard to the ensemble generation techniques that are applied to derive the statistics. It is found that background error statistics derived by simulation through Ensemble Data Assimilation are more realistic than the corresponding statistics derived by downscaling from larger scale ensemble data. In case perturbation of observations is used to inject a spread into the ensemble, and the ensemble is integrated over a few hours only, it was found that the derived structure functions may be contaminated by the geometry of the observing network. The effects of the assumptions of stationarity, homogeneity and isotropy, that are generally applied in the generation of background error statistics, and the implications of the background error covariance model have also been illustrated. Spatial covariances derived under these assumptions were contrasted against spatial covariances obtained by ensemble averaging only, preserving the signals from forecast errors of the day. This indicates that it is likely to be favourable to apply data assimilation with ensemble background error statistics obtained from ensemble averaging, like in ensemble Kalman filters or in hybrids between variational and ensemble data assimilation techniques.

Keywords: Mesoscale data assimilation, background error covariance model, ensemble prediction system, climatological balance relationships, homogeneity and isotropy

1. Introduction

The number of observations for initialization of Numerical Weather Prediction (NWP) models is generally much smaller than the number of model state variables. For this reason, *a priori* information is needed and is generally taken as a background model state, for example a short-range forecast, and some statistical measure of the uncertainty of this background state - the background error statistics (BGE). The BGE statistics includes information on spatial scales of the background forecast errors and on balances between errors of different model variables.

It was early recognized that ideal BGE statistics needs to be flow dependent since forecast errors have a strong time variability. For most operational NWP systems, however, data assimilation schemes have been based on flow independent BGE statistics, derived by averaging in time. Besides this implicit assumption of stationarity, other simplifying assumptions like

horizontal homogeneity and isotropy are also generally applied.

A model for the BGE statistics of limited area forecast models was formulated by Berre (2000). This model is based on a spectral representation of spatial covariances and can be applied to spectral as well as grid point forecast models. A time series of forecast differences is taken as a proxy for background forecast errors and parameters of a background error model are estimated. The forecast differences may be taken as differences between forecasts of different forecast lengths but valid at the same time and this is referred to as the NMC method (Parrish and Derber, 1992). Here we will apply Ensemble Data Assimilation to obtain the time series of forecast differences between ensemble members. The Berre (2000) BGE statistics model has been successful for synoptic scale data assimilation, while significant efforts have been needed to show positive impact of data assimilation for models with a grid resolution of a few km (Gustafsson et al., 2018).

In this paper, we will investigate the application of the Berre (2000) BGE statistics model to the HARMONIE

*Corresponding author. e-mail: Jelena.Bojarova@smhi.se

(HIRLAM ALADIN Research on Mesoscale Operational NWP In Europe) AROME convective scale forecasting system. The HARMONIE AROME model (Bengtsson et al., 2017) is essentially the same as the AROME model applied by Météo-France (Seity et al., 2011).

Two sets of BGE statistics based on two different time series of short-range ensemble forecasts will be compared in the present study. One time series of ensemble forecasts is obtained by downscaling of a global ECMWF ensemble that was derived through an Ensemble of Data Assimilations (EDA) with different perturbations of observations for each ensemble member. The second time series of ensemble forecasts is a high resolution ensemble derived by EDA for the HARMONIE model. The characteristics of the two sets of ensembles and their suitability for BGE statistics will first be investigated in sections 2–3 and then the derived statistics will be described and compared in sections 4–5. The effects of the assumptions of stationarity, homogeneity and isotropy will be demonstrated and discussed in section 6 and the relevance of the statistical balance constraints for mesoscale data assimilation will be explored in section 7. A discussion will follow in section 8 and a summary and conclusions are provided in section 9.

2. Atmospheric spectra and grid point structure functions

The development of high resolution models with a grid separation of a few km has made it possible to study model spectra ranging from the global scales down to scales resolved by these models (Skamarock et al., 2014). Atmospheric energy spectra can also be studied by observations, in particular aircraft observation of winds at flight levels (8–12 km above the ground) have proven to be useful for this purpose (Nastrom and Gage, 1985). With regard to kinetic energy, model simulation spectra and spectra based on observed data agree reasonably well. They indicate (1) a spectral energy slope proportional to k^{-3} , where k is the horizontal wave number, for horizontal scales ranging from the largest global scales down to approximately a few hundred km; (2) a more flat spectral energy slope proportional to $k^{-5/3}$ for smaller scales down to the smallest horizontal scales resolved by the models. Below the smallest scales resolved by the models, generally for wave-lengths shorter than $6\text{--}10 \Delta s$ where Δs is the grid separation of the model, the spectral energy densities fall off quickly.

In case a high-resolution model is initialized with data from a coarser resolution model or data assimilation (“down-scaling experiment”), the model integration will start with no energy in the smallest resolved scales, and during the initial model integration the energy content in these scales will adjust to the larger scale input energies

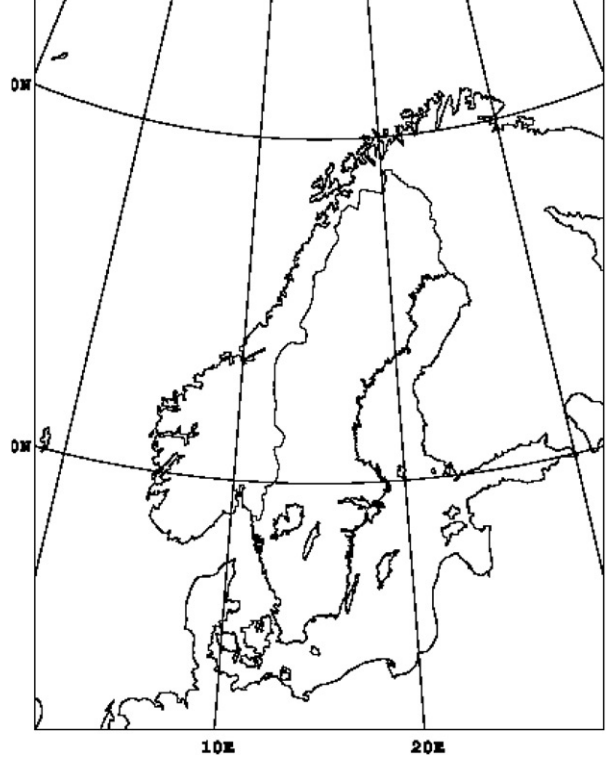


Fig. 1. Model domain for the HARMONIE ensemble experiments (739×949 horizontal grid points with a 2.5 km grid resolution).

such that the simulated model energy spectra will approach the theoretically predicted spectra. This initial spectral adjustment or spectral spinup process may last for as long as 6–24 hours (Skamarock et al., 2014).

The HARMONIE forecasting model is a limited area spectral model (Bubnova et al., 1995) utilizing an area extension (Haugen and Machenhauer, 1993) for its spectral transforms. The application of the extension zone for the calculation of energy spectra is associated with horizontal aliasing problems producing distorted spectra (Blazika et al., 2015). An alternative physical space calculation of structure functions (Lindborg, 1999; Frehlich and Sharman, 2004) will be applied here in order to obtain a result not influenced by aliasing effects.

We will thus calculate the second order structure function $f(s)$ for the longitudinal wind u component in the east-west direction:

$$f(s) = \overline{(u(x+s, y, t) - u(x, y, t))^2}^{xyt} \quad (1)$$

where s is distance separation in the x -direction and $\overline{(\)}^{xyt}$ denotes an average in the horizontal and in time. We are using the concept *structure function* in two slightly different meanings in this paper. To see that they are related one can re-write equation (1) slightly:

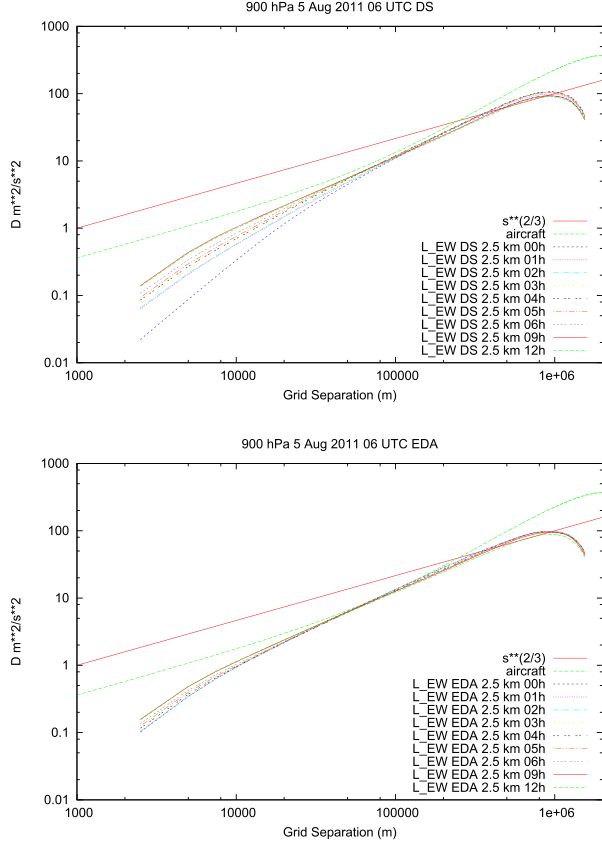


Fig. 2. Structure functions for the longitudinal component of the wind in the east-west direction. Horizontal domain average at model level 47 (≈ 900 hPa) for 5 August 2011 06 UTC. Separate curves for forecast lengths +0 h, +1 h, +2 h, +3 h, +4 h, +5 h, +6 h, +9 h and +12 h. Downscaling (top) and EDA (bottom).

$$f(s) = \frac{\overline{u(x+s, y, t)^2}^{xyt} + \overline{u(x, y, t)^2}^{xyt}}{-2\overline{u(x+s, y, t) \times u(x, y, t)}^{xyt}} \quad (2)$$

The structure function $f(s)$ is thus simply proportional to the difference between the sum of the variances and the covariance multiplied by 2. A covariance in physical space corresponds to a variance density spectrum in spectral space. It can be proven that a kinetic energy spectrum $\sim k^{-5/3}$ corresponds to a structure function $f(s) \sim s^{2/3}$ (Frehlich and Sharman, 2004).

3. Characteristics of HARMONIE ensembles based on downscaling and ensemble data assimilation

The HARMONIE model was applied for the month of August 2011 at a horizontal resolution of 2.5 km over the domain illustrated in Fig. 1 and with 65 vertical levels. The lateral boundary conditions for two ensemble

experiments were taken from operational ECMWF EDA (Ensemble of Data Assimilations) ensemble forecasts at 06 UTC and 18 UTC with forecasts available up to +12 h. This had the effect that HARMONIE ensemble forecasts could be generated up to +12 h at 06 UTC and 18 UTC, but only up to +6 h at 00 UTC and 12 UTC. The ECMWF EDA used a T399 horizontal resolution and 91 vertical levels. The ECMWF 4D-Var assimilation windows were from 00 UTC until 12 UTC, and from 12 UTC until 00 UTC, for the assimilation cycles at 06 UTC and 18 UTC, respectively.

The first set of HARMONIE ensemble forecasts was generated by a direct downscaling of the ECMWF EDA ensemble forecasts at 06 UTC and 18 UTC. The HARMONIE forecasts were integrated up to +12 h. The second set of HARMONIE ensemble forecasts were generated from initial data obtained by a 3D-Var data assimilation scheme (Fischer et al., 2005) applied with a 6 h data assimilation cycle. Similar to what is done in the ECMWF EDA, observations were perturbed with Gaussian random numbers, applying different observation perturbations for different ensemble members. Only conventional observations (TEMP, SYNOP, SHIP and aircraft data) were assimilated. The standard deviations of the observation perturbations were taken to be the same as the assumed observation error standard deviations for the data assimilation. The two ensemble experiments were both run with 6 ensemble members. The lateral boundary conditions for any member k ($k = 1, 2, \dots, 6$) were coming from the same ECMWF EDA member in the HARMONIE downscaling and in the HARMONIE EDA ensemble runs. No initialization scheme was applied.

As a diagnosis of the spectral characteristics of the two sets of HARMONIE ensemble forecasts, we will show structure functions for the longitudinal component of the wind in the east-west direction calculated in accordance with equation (1). Fig. 2 shows horizontally averaged structure functions at model level 47 (≈ 900 hPa) downscaling (top) and EDA (bottom) forecasts for 5 August 2011 06 UTC with forecast length ranging from +0 h to +12 h. The curve for the theoretical slope $s^{2/3}$ of the structure function, corresponding to the slope $k^{-5/3}$ of the kinetic energy spectrum, has also been included in Fig. 2, as well as a structure function curve derived from aircraft observations (Frehlich and Sharman, 2004).

We may first notice that the downscaling structure function curves are associated with a strong spinup during the initial 6–12 h of the forecast. This is mainly because the HARMONIE forecast model, in the downscaling experiment, was initialized with ECMWF coarse resolution data. A similar spinup effect was reported by Skamarock et al. (2014). The HARMONIE EDA

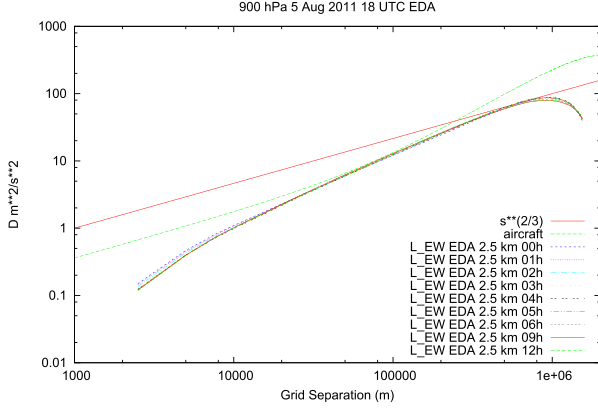


Fig. 3. Structure functions for the longitudinal component of the wind in the east-west direction. Horizontal domain average at model level 47 (≈ 900 hPa) for 5 August 2011 18 UTC. Separate curves for forecast lengths +0h, +1h, +2h, +3h, +4h, +5h, +6h, +9h and +12h. The EDA experiment.

structure function curves seem to be affected by a weaker spinup, but one needs to consider also the summer day time spinup of planetary boundary layer structures (turbulence and convection, for example), since the forecasts were initialized in the local morning time at 06 UTC. The presence of daytime spinup processes reflected in the structure functions for the EDA forecasts initialized at 06 UTC in Fig. 2 (bottom) is confirmed by the corresponding structure functions for forecasts initialized at 18 UTC in Fig. 3. We may notice a significant spindown of the structure functions with forecast length during the night for the shortest horizontal scales, most likely corresponding to the dying out of turbulence in the planetary boundary layer. The night time spindown in Fig. 3 is not as strong as the day time spinup in Fig. 2, probably indicating an additional model spinup due to deficiencies in the data assimilation process, both at 06 UTC and at 18 UTC.

Secondly, we can notice in Fig. 2 that the slopes of the model structure function curves follow the slope of the observation structure function reasonably well for grid separations between approximately 15–20 km and 500–1000 km. The upper limit is simply a sign of a rather small model domain, while the lower limit is an indication that the **real resolution** of the HARMONIE model with a grid distance $\Delta s = 2.5$ km is in the range $6\Delta s - 8\Delta s$ (Frehlich and Sharman, 2004). This is expected, due to limitations in the numerical treatment of highly non-linear processes in the model physics and the artificial actions of horizontal diffusion, for example, to counteract aliasing effects.

Comparing the two sets of structure functions for downscaling and EDA, we may notice that EDA curves have slightly larger values for a mesoscale range of grid

separations 50–200 km, see Fig. 2. This may indicate that there are spinup processes that are slower than what is possible to describe with the differences between the structure function curves for forecast of length 0–12 h, another explanation is that the data assimilation process in the EDA experiment may add and retain further energy in the mesoscale range.

4. Brief description of the model for the background error statistics

The model for the BGE statistics in the HARMONIE data assimilation was formulated by Berre (2000). Since the background error covariance matrix in physical model space (the **B**-matrix) is too large to be stored explicitly, the forecast background error vector δX is transformed to a control vector χ , the elements of which are assumed to be statistically independent.

$$\chi = VDF\delta X \quad (3)$$

where F is a horizontal 2-dimensional Fourier transform from physical grid point space to spectral space. D is a balancing operator and V is a vertical transform utilizing the eigen-vectors of vertical covariance matrices. The application of 2-dimensional Fourier transforms (F) together with the assumption that spectral components for different wave-lengths are statistically independent corresponds to the assumption that the BGE statistics in physical space are horizontally homogeneous.

The assumption that spectral components for different wave numbers are statistically independent has the convenient consequence that horizontal covariances can simply be expressed as the variances for different wave components. If we furthermore assume horizontally isotropy with respect to the covariances, the horizontal covariances can be represented by 1-dimensional variance density spectra for the control variables. We will study such variance density spectra below.

The balance operator (or statistical de-correlation operator) D is derived in spectral space with a step-wise multiple regression technique for each wave number component separately, having the effect that the balancing becomes scale dependent. The multiple regression proceeds stepwise, starting from the background forecast errors of model variables vorticity (ζ), divergence (η), temperature (T), surface pressure (P_s) and specific humidity (q) in spectral space. The first step is the derivation of a horizontal balance operator (H) with a regression between vorticity and linearised geopotential (P_b , a linear combination of surface pressure and temperature). No horizontal variations of the Coriolis parameter are considered in this statistical (f -plane) balance. The second step in the multiple regression is the derivation of a

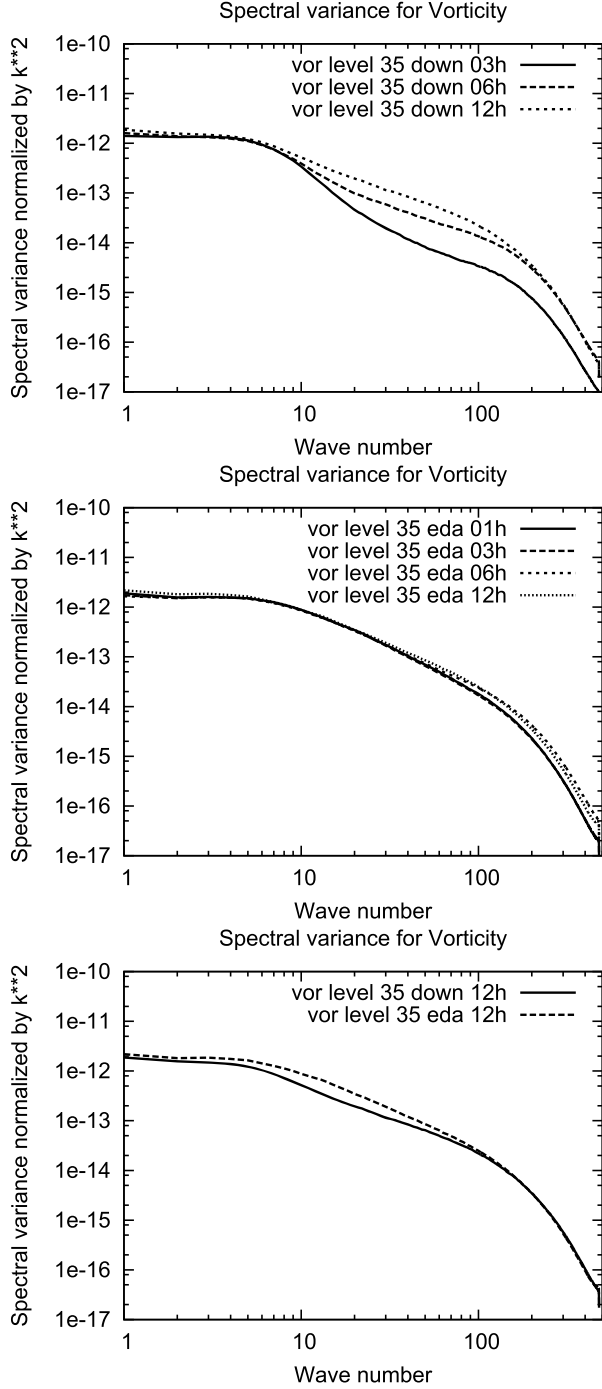


Fig. 4. Horizontal spectral variance densities of +1h, +3h, +6h and +12h vorticity background error as estimated by the ensemble downscaling (top) and the Ensemble Data Assimilation (EDA, middle) techniques. Model level 35. Spectral variance densities of +12h vorticity background error as estimated by the downscaling and EDA techniques (bottom). The spectral variance densities have been normalized with k^2 , where k^* is the one-dimensional wave number.

balance operator (M) for the relation between divergence and vorticity, via the linearised geopotential. For the

difference between total divergence (η) and the balanced part ($MH\zeta$) we will use the term unbalanced divergence (η_u). The stepwise multiple regression then proceeds with regression equations for temperature, surface pressure and specific humidity. The regression equations can be summarized as follows:

$$\begin{aligned}\zeta &= \zeta \\ \eta &= MH\zeta + \eta_u \\ (T, P_s) &= NH\zeta + P\eta_u + (T, P_s)_u \\ q &= QH\zeta + R\eta_u + S(T, P_s)_u + q_u\end{aligned}\quad (4)$$

where the matrices H , M , N , P , Q , R and S include the regression coefficients. Note that the regression is for vertical profiles of the spectral coefficients of the model variables. With the exception for the horizontal balance operator H , the balance operators are vertical operators.

In summary, the model for the background error statistics include: (1) The balance operators H , M , N , P , Q , R , and S ; (2) Horizontal variance density spectra for the control variables ($\zeta, \eta_u, (T, P_s)_u, q_u$) and (3) Vertical correlations for the control variables.

5. Comparisons of background error statistics based on downscaling and EDA

HARMONIE BGE statistics have been derived for a Nordic domain (Fig. 1) with a 2.5 km horizontal grid resolution and with 65 vertical levels. The HARMONIE model is a spectral model based on semi-Lagrangian and semi-implicit time integration. Due to the semi-Lagrangian representation of advection, it is possible to apply a “linear” grid where the shortest wave in the spectral model representation corresponds to 2 grid lengths (=5 km).

Here we will discuss the estimation of BGE statistics for 06 UTC initial data only, forecasts were run up to +12h for the whole month of August 2011 and a 6h data assimilation cycle was applied in the case of the HARMONIE EDA run. Both the downscaling forecast and the HARMONIE EDA run included 6 ensemble members. The following ensemble member model state differences were taken as proxies for background forecast errors in the estimation of the BGE statistics: member 1 - member 2, member 2 - member 3, member 3 - member 4, member 4 - member 5, member 5 - member 6 and member 6 - member 1. This combination of difference fields is linearly dependent, but it is kept in order to apply the same efficient weight for each ensemble member. The only effect of the linear dependency will be a scaling of the resulting statistics (a factor 2.0 for variances).

5.1. Spectral densities

Figure 4 (top and middle) shows horizontal variance density spectra for vorticity background errors, at model level 35 (≈ 500 hPa), corresponding to different forecast lengths for downscaling and EDA, respectively. We have normalized the vorticity spectra in the figure with $k^*{}^2$, where k^* is the one-dimensional wave number. This makes the amplitude of the spectral densities proportional to rotational kinetic energy. Fig. 4 (bottom) compares spectra of +12h vorticity background errors for downscaling and EDA. From Fig. 4, it is clear that the downscaling BGE statistics are affected by severe spinup processes for wavelengths shorter than a few hundred km, at least at +3h and +6h. The reason, as discussed above, is that the HARMONIE downscaling forecasts were initialized from much coarser resolution initial ECMWF EDA data. From the comparison between downscaling and EDA spectra at +12h (Fig. 4, bottom), in particular with regards to the shapes of the spectra, it seems that spectral spinup process is still ongoing at +12h in the down-scaling case. It should be noted, however, that not only the model spectral spinup, but also the applied data assimilation calculations and the perturbation of observation errors can influence the EDA spectra.

With regard to the variations in the EDA spectra depending on background error forecast length, influences from at least two processes can be identified in Fig. 4 (middle). For most scales there is a general increase in the amplitude of the spectral variance with forecast length, in particular for horizontal scales smaller than a few hundred km. This corresponds to the growth of forecast errors with forecast length. On top of this there are also daily cycle effects at the smallest scales, with larger amplitude of the spectral variance at +6h (valid time 12 UTC) than at +12h (valid time 18 UTC). It may seem that the growth of the forecast error with forecast length, as manifested in the amplitude of the spectral variances, is rather weak. This will be further discussed below.

Horizontal spectra for unbalanced divergence and unbalanced temperature (not shown), unbalanced surface pressure (Figure 5) and unbalanced humidity (Figure 6) background errors show similar features for downscaling and EDA as seen and discussed for vorticity spectra. For the unbalanced surface pressure EDA spectra, we may notice that for the longest waves there is a significant decrease in spectral variance with forecast length from +1h to +3h and then a modest increase in spectral variance from +3h until +12h. The decrease of unbalanced surface pressure variance at the start of the model integrations may be interpreted as a sign of the lack of initialization in the HARMONIE 3D-Var based data

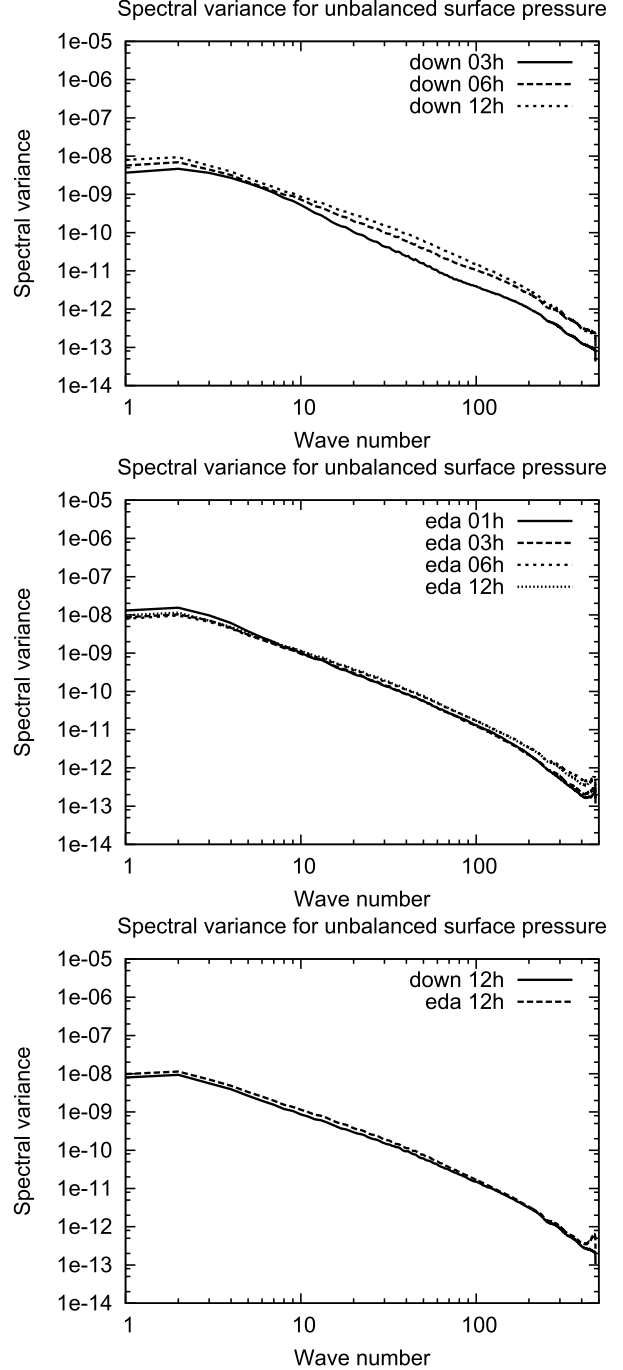


Fig. 5. Spectral variance densities of +1h, +3h, +6h and +12h unbalanced surface pressure background error as estimated by the ensemble downscaling (top) and the Ensemble Data Assimilation (EDA, middle) techniques. Model level 35. Spectral variance densities of +12h unbalanced surface pressure background error as estimated by the downscaling and EDA techniques (bottom).

assimilation, leaving for the model to carry out geostrophic adjustment during the model integration as a

compensation for this lack of initialization. Due to this geostrophic adjustment, observed information for surface pressure on the largest horizontal scales may be rejected and lost. A less optimistic interpretation is that there are systematic errors in the large-scale physics and/or dynamics of the model resulting in large scale geostrophically balanced model errors. One may say that the model cannot describe the balanced climate of the real atmosphere inherent in the observations. Forecast verification scores seem to support the latter hypothesis.

From the unbalanced humidity spectral variance densities (Figure 6), we may observe that EDA spectral densities are generally larger than the spectral densities of the down-scaling spectral densities. Since also the balanced part of the humidity variances are larger for the EDA case than for the down-scaling case (see below), we may conclude that the HARMONIE EDA generated model “atmosphere” has a more vivid moisture climate than the downscaling model “atmosphere”, and that it takes more time than 12 hours for the HARMONIE model to spin up to this more vivid moisture model climate.

The spectral variance densities for vorticity and unbalanced divergence at model level 35 are compared in Figure 7. We may conclude that the rotational part of the flow dominates the estimated background errors for scales corresponding to wave numbers 1–10.

5.2. Standard deviations and balance relations

Background error variances for the different variables can be obtained by summing the spectral variance densities over horizontal wave numbers. This can be carried out for the unbalanced component as well as for the different balanced components of the different variables. Figure 8 shows standard deviations for vorticity, for +3 h and for +12 h as well as for statistics generated by down-scaling and by EDA. Note the strong effect of spinup from +3 h to +12 h in the downscaling case, and the far smaller increase due to vorticity error growth from +3 h to +12 h in the case of the EDA.

Standard deviations for various components of the temperature background error are shown in Figure 9 for the EDA case. Note the relative weak balance between temperature and, for example, vorticity (geostrophic balance) and also note the relatively small temperature background error amplification from +3 h to +12 h.

Balance relations with the temperature involved are strongly dependent on horizontal scale. This can be seen in Fig. 10 which shows percentages of temperature variance explained by vorticity and unbalanced divergence as a function of horizontal wave number. For the largest horizontal scales (wave length around 1000 km) around 50% of the temperature variance is explained by a linear

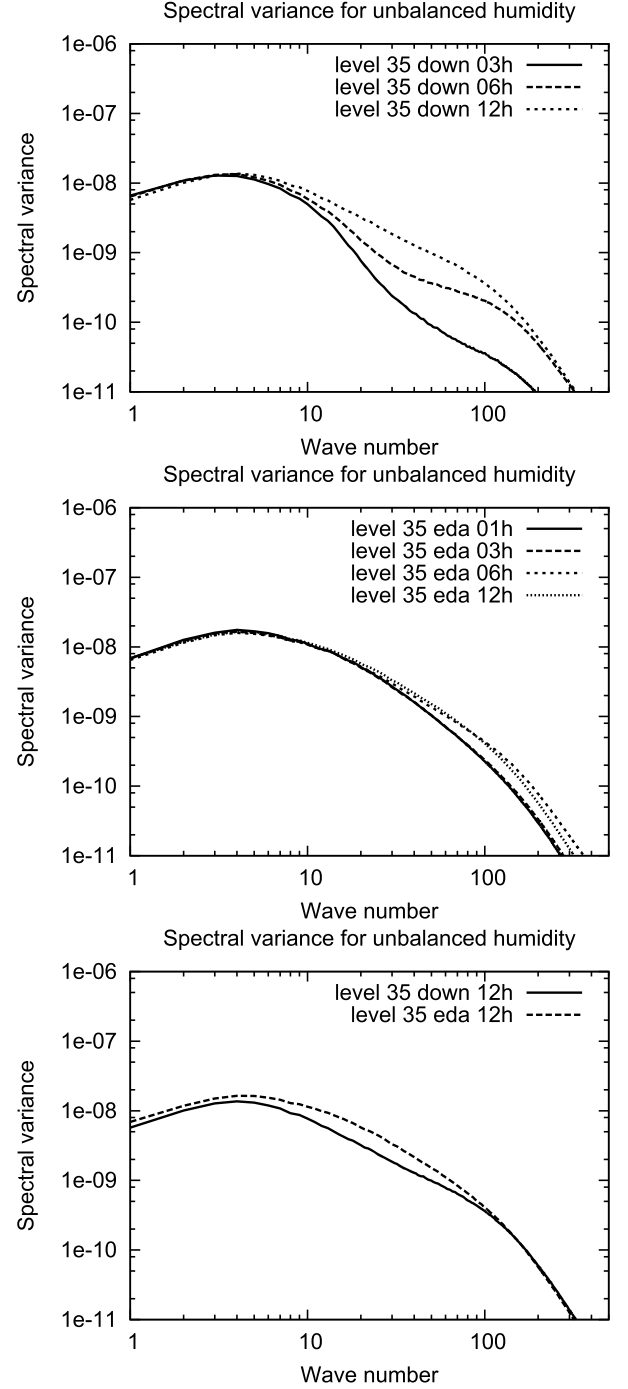


Fig. 6. Spectral variance densities of +1 h, +3 h, +6 h and +12 h unbalanced humidity background error as estimated by the ensemble downscaling (top) and the Ensemble Data Assimilation (EDA, middle) techniques. Model level 35. Spectral variance densities of +12 h unbalanced humidity background error as estimated by the downscaling and EDA techniques (bottom).

balance with vorticity (thus a linear geostrophic balance), while this balance more or less disappears for horizontal

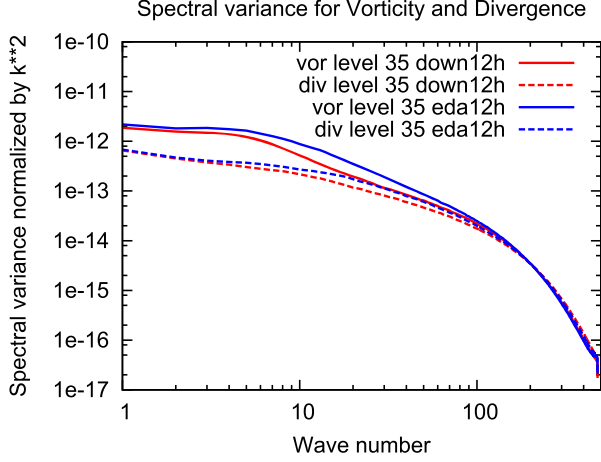


Fig. 7. Spectral variance densities of +12h vorticity and unbalanced divergence background errors as estimated by the ensemble downscaling and the Ensemble Data Assimilation (EDA) techniques. Model level 35. The spectral variance densities have been normalized with k^{*2} , where k^{*} is the one-dimensional wave number.

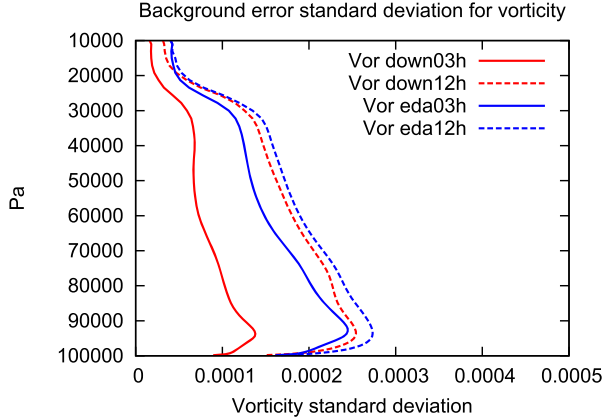


Fig. 8. Background error standard deviations for vorticity, +3h and +12h, downscaling and EDA.

scales shorter than 100 km. The balance between temperature and unbalanced divergence is very weak, being strongest (around 15%) at small horizontal scales. The maxima that appear at the shortest waves (2 grid lengths, 5 km) most likely represent noise in the form of grid scale interaction between wind, temperature (and surface pressure, see below).

The horizontal wave length dependencies of the balance relations with the (logarithm of) surface pressure involved are illustrated in Figure 11. As is the case for temperature, there is a significant degree of balance between surface pressure and vorticity for the largest horizontal scales. Furthermore, this degree of balance increases with forecast length during the model

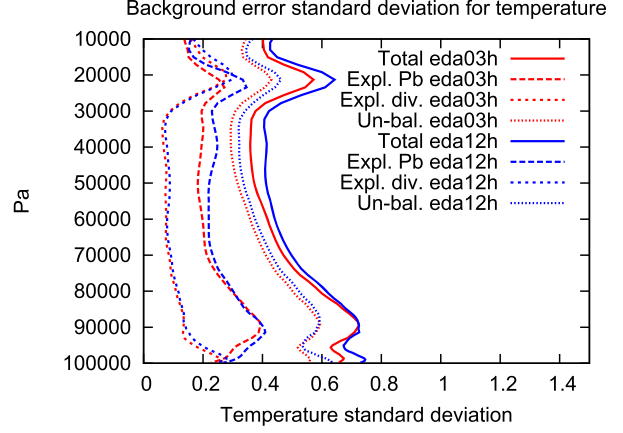


Fig. 9. Background error standard deviations for temperature, +3h and +12h, EDA; The component balanced with vorticity (via linearized geopotential P_b), the component balanced with unbalanced divergence, the unbalanced component and total standard deviation.

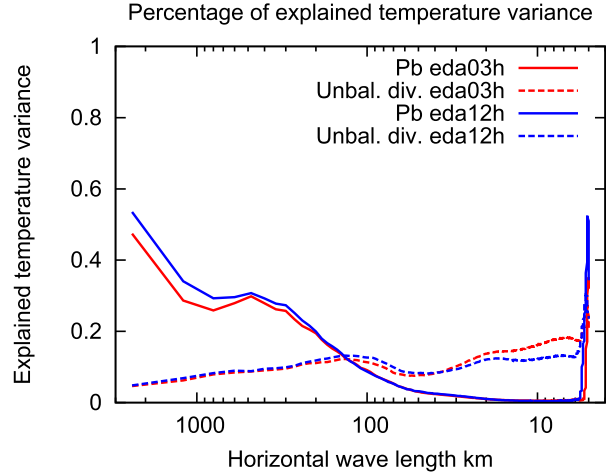


Fig. 10. Percentages of temperature background error variances explained by vorticity and unbalanced divergence as a function of horizontal wave number. EDA +3h and EDA +12h.

integration, from around 60% in explained surface pressure variance at +3h to almost 80% at +12h. At the same time, forecast verification scores (not shown) indicates that there is a significant drift of surface pressure away from the observed state during model integrations on the same time scale. One interpretation is that the model, due to limitations of some physical processes, establishes a large-scale geostrophic balance that is different from the balance in the real atmosphere realized via the observations. If this is a correct interpretation, it is of course very serious since it makes utilization of surface pressure observations from the real atmosphere very difficult, if not even meaningless. The local maximum in the explained surface pressure variance at ≈ 300 –400 km

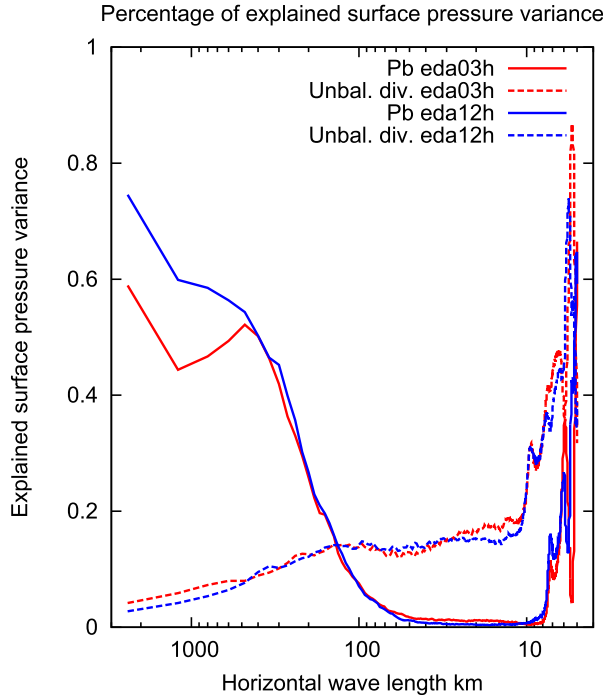


Fig. 11. Percentages of surface pressure background error variances explained by vorticity and unbalanced divergence as a function of horizontal wave number. EDA +3h and EDA +12h.

wave length, seen in particular at +3h forecast range and also for temperature, is further investigated and discussed in subsection 7.3 below.

A second likely model deficiency is also indicated in Figure 11 for horizontal scales shorter than approximately 10 km, thus for scales corresponding to 2–4 grid lengths. There seems to be a strong interaction between surface pressure, vorticity and divergence variations in the model at these small scales. In fact, this strong variability on the smallest scales has been introduced on purpose by a reduction of the horizontal diffusion coefficient in the HARMONIE model with AROME physics. The AROME physics does not include a deep convection parameterization scheme and relies on the model dynamics to establish convective cells that carries out the needed vertical transports associated with convection. With the normal setting of horizontal diffusion, the HARMONIE AROME model experienced a tendency to create very unrealistic “super-cells” of convection, that could only be cured in the operational setting by a reduction of horizontal diffusion.

This problem of small-scale noise raises many questions:

- The HARMONIE AROME model uses a linear grid (the shortest wave in the spectral model representation is 2 grid lengths). This was introduced together with the semi-Lagrangian advection, since aliasing due to the quadratic Eulerian advection no

longer is present. But the model, and in particular the physics, include many other non-linearities of even higher order than the quadratic advection. The small-scale noise that we observe in Figure 11 could be due to aliasing effects. One possibility is to re-introduce a quadratic or even a cubic grid to avoid aliasing of non-linear model terms.

- It seems over-optimistic to explicitly describe the effects of convective cells on the kilometre-scale with a model using a 2.5 km grid resolution. Taking our consideration that the model can accurately describe phenomena at a horizontal scale of 5–6 grid lengths, a model grid resolution of the order of 500 m would be needed.

Balance relations with moisture involved are illustrated in Figure 12. It can be seen that moisture includes a large-scale component balanced with vorticity. This may be interpreted as moistening (drying) associated with frictional inflow (outflow) in synoptic scale disturbances. For smaller horizontal scales, balance relations with moisture involved are more associated with temperature variations, probably related to convection and condensation processes (i.e. latent heat release). The increased importance of divergence-humidity relations at the 10–30 km scale may be interpreted as the effect of inertia-gravity waves associated with adjustment processes.

6. On the basic assumptions of the HARMONIE data assimilation background error statistics

The model for the HARMONIE BGE statistics is based on the three assumptions of stationarity, homogeneity and isotropy with respect to horizontal covariances. In this section, we will investigate the effects of these assumptions. We will take temperature at a few model levels as examples, and we will use the HARMONIE EDA ensemble data from 2–31 August 2011, 06 UTC + 12h only, for a demonstration. Figure 13 shows time-averaged horizontal correlations for temperature at model level 35 with reference to a grid point at (66° N, 08° E). The time-averaged “raw” correlations, thus without any further assumptions applied, are shown in Figure 13 (upper, right), correlations based on horizontal homogeneity are shown in Figure 13 (lower, left) and correlations based on horizontal homogeneity and isotropy are shown in Figure 13 (lower, right). A map illustrating orography has been added in Figure 13 (upper left).

Figure 13 clearly shows that time-averaging over 25 days and 6 ensemble members only is not enough to establish a stable model for the background error statistics. The effects of sampling errors are significant, although the general shapes of the correlations in Figure 13 (upper, right) have similarities to the correlations where horizontal homogeneity and isotropy have been

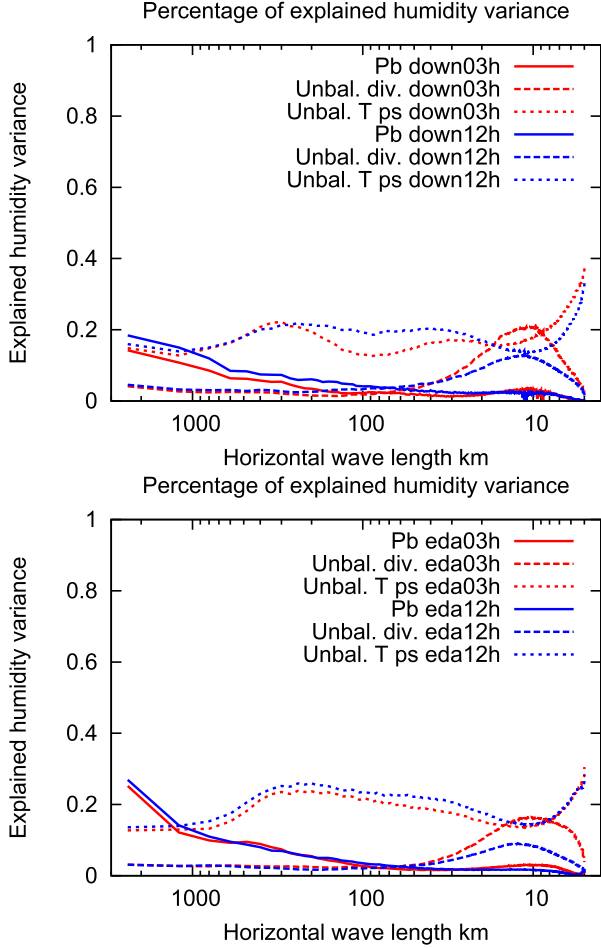


Fig. 12. Percentages of humidity background error variances explained by vorticity, unbalanced divergence and unbalanced temperature and surface pressure as a function of horizontal wave number. Downscaling +3h, +12h (top), EDA +3h, +12h (bottom).

applied (Figure 13, lower left and right). We may also notice that horizontal averaging (through the assumption of horizontal homogeneity) is a very efficient remedy to obtain stable and smooth background error statistics. The assumption of horizontal isotropy adds some further minor horizontal smoothing.

Although the smooth statistics obtained through the assumptions of homogeneity and isotropy may be attractive from a computational stability point of view, one may ask how representative the derived stationary, homogeneous and isotropic horizontal correlations are for real forecast error correlations with a strong case-to-case variability. To illustrate this, we have selected a single case, 12 August 2011 06 UTC + 12h, and we show in Figure 14 the same correlations maps as shown in 13 but estimated from the single case only. We have also introduced the background state from one of the ensemble members (member 1) in Figure

14. The reference point for the horizontal correlations is the same as in Figure 13 and was selected to be in the area of an upper air frontal zone (model level 35 $\approx$ 500 hPa). It may appear too naive to try to estimate horizontal background forecast error correlations from 6 ensemble members only, but the message of the “raw” horizontal correlations of Figure 14 (upper, right) is clear and physically relevant. The correlation patterns in the vicinity of the reference point are large scale, elongated along the frontal zone and are indicating an uncertainty in the horizontal position of the upper air frontal zone.

The effects of the assumption of horizontal homogeneity are very strong in this particular case, compare Figure 14 (upper, right) with Figure 14 (lower, left). Due to the averaging over all horizontal grid points, as inherent in the effects of the assumption about horizontal homogeneity, the message about error structures due to uncertainties in frontal positions is completely wiped out and replaced with the same small scale, almost isotropic, structures seen in the time averaged statistics of Figure 13. Some remaining large-scale irregularities are smoothed out by the assumption of horizontal isotropy, see Figure 14 (lower, right).

For atmospheric variables close to surface of the earth, one would expect horizontal correlations that are non-isotropic and inhomogeneous, depending on orography and on land/sea differences, for example. This has been demonstrated with the use of observational data for standalone mesoscale analysis (Häggmark et al., 2000) as well as from time-averages of model simulations (Sattler and Huang, 2002). One possibility would be to estimate parameters for anisotropic and inhomogeneous correlation models from time-series of ensemble information. Figure 15 shows one example for temperatures at model level 62 (a few hundred meters above the ground) in order to illustrate these possibilities. Data from a sub-domain over the Western Baltic Sea, including the Islands of Gotland (57.5°N, 18.5°E) and Öland (56.5°N, 16.5°E) and the east coast of Mainland Sweden were extracted. The time average horizontal correlations with reference to a position over the central part over the Island of Gotland are shown in Figure 15 (upper right), and horizontal correlations for two individual cases, averaged over ensemble members only, are shown for 12 August 2011 06UTC + 12h (lower, left) and 21 August 2011 06 UTC + 12h (lower, right).

For individual cases, the horizontal correlations in Figure 15 indicate a strong dependency on the underlying surface conditions such that, for example, the temperature forecast differences in the boundary layer over the Island of Gotland are more correlated with the boundary layer temperatures over the Island of Öland and mainland Sweden than with temperatures over the Baltic Sea water surfaces in between. Such correlation dependencies are of course also related to atmospheric stability and

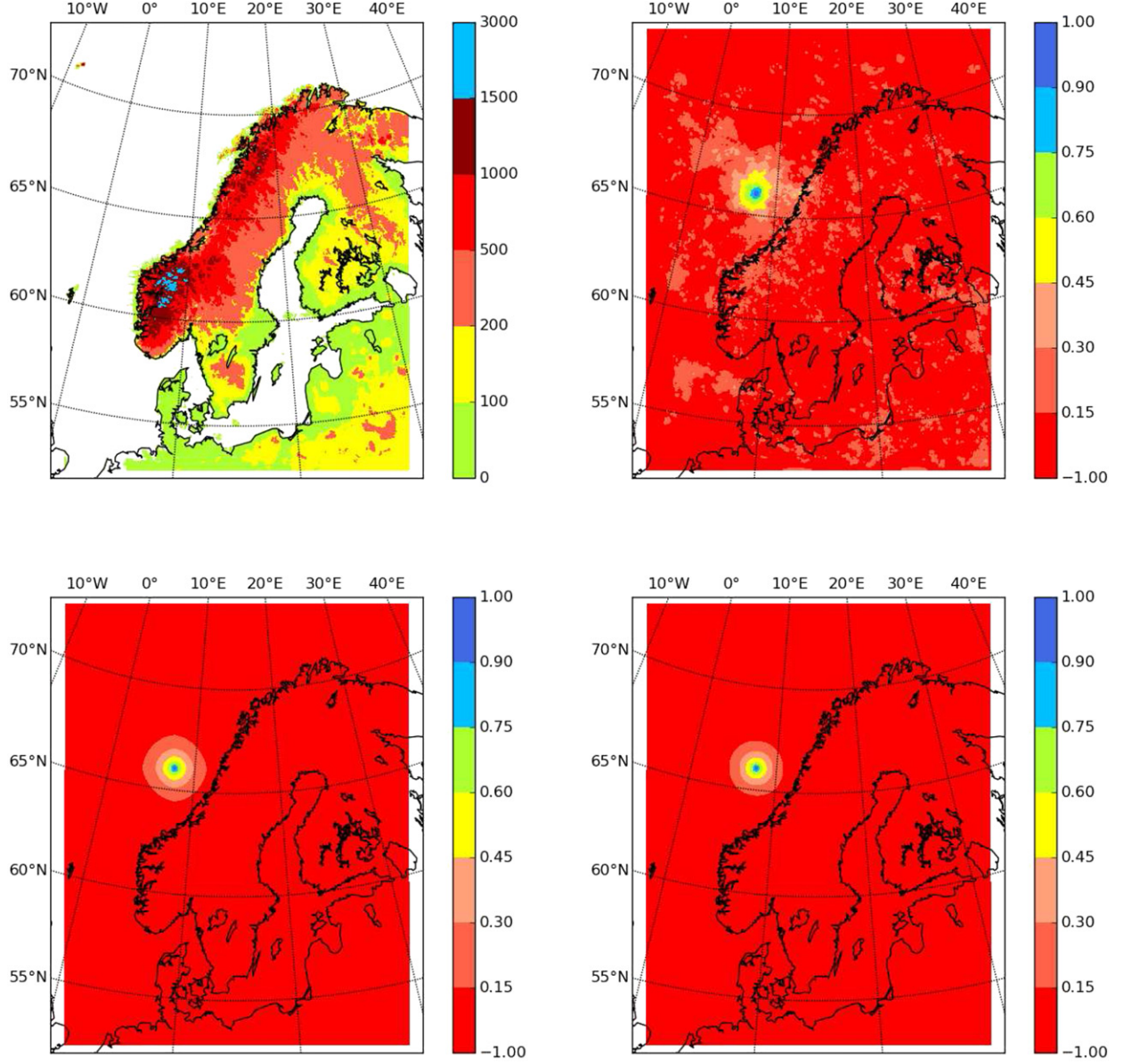


Fig. 13. Orography (upper left). Horizontal correlations for temperature at model level 35 (≈ 500 hPa) with reference to the grid point $(x,y)=(235,595)$, averages over ensemble members and time: no further assumptions (upper right), homogeneity assumption (lower left), homogeneity and isotropy assumptions (lower right).

other air mass conditions. In the time averages over the whole month of August 2011, the horizontal correlations become more homogeneous and isotropic due to averaging over different stability and air mass conditions.

7. Relevance of the statistical balances and horizontal spectra for mesoscale data assimilation

We will now evaluate the relevance of the background error covariance model through its impact on the

forecast quality. To measure the forecast quality in a statistical sense, we have conducted a series of one-month HARMONIE 3D-Var experiments for the period from the 13th of June 2016 until the 13th of July 2016. We use standard verification scores, Bias and Root Mean Square Error (RMSE), averaged over the last 24 days. All experiments run in a data assimilation cycling mode with a 3 h update frequency. The domain is centred over Denmark, it covers an area of $1620 \times 1620 \text{ km}^2$ and employs a linear grid with a 2.5 km grid resolution.

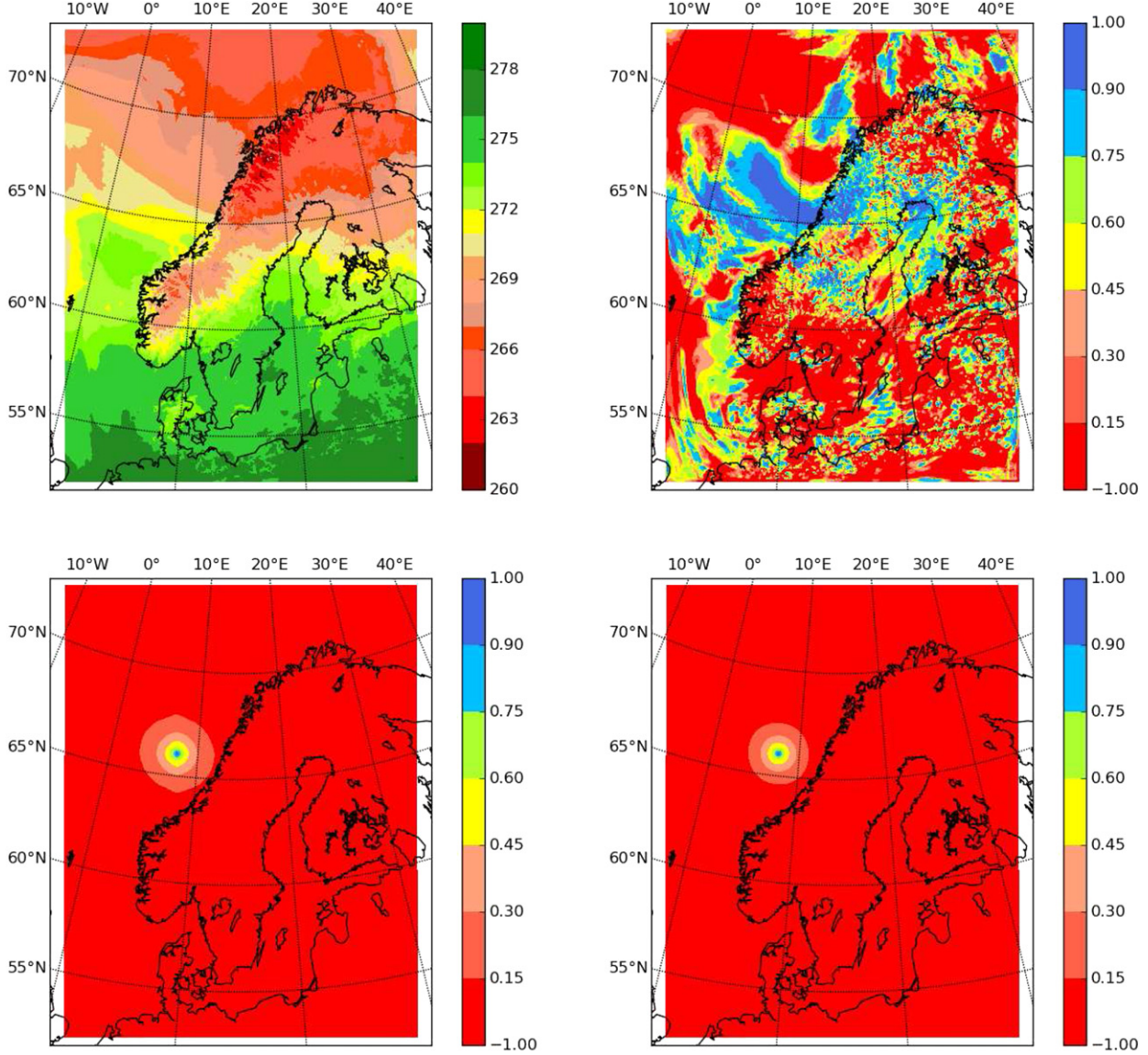


Fig. 14. Temperature at model level 35 (≈ 500 hPa) 12 August 2011 +12h, ensemble member 1. Horizontal correlations for temperature with reference in the grid point ($66^\circ\text{N}, 08^\circ\text{E}$), averages over ensemble members: no further assumptions (upper right), homogeneity assumption (lower left), homogeneity and isotropy assumptions (lower right).

7.1. Minimization on restricted horizontal scale

In the first series of experiments, we use identical climatological structure functions but we restrict analysis increments to a coarse horizontal resolution. Such sensitivity experiments are easy to carry out with the HARMONIE data assimilation, since the assimilation control variables are formulated in spectral space. For a grid-point representation of horizontal covariances through horizontal diffusion operators, this would, for example, require a reformulation of these diffusion operators. Thus, the observations filtered through the background error covariance influence the initial model state for a specified

range of scales only. At the same time, the impact from observations is redistributed to the restricted range of spatial scales. The experiments are summarized in Table 1.

The domain geometry with a linear grid allows 324 one-dimensional isotropic waves in the horizontal. An elliptic truncation is used to relate a one-dimensional isotropic wave to 2D harmonics. For example, in `cntr50` (see Table 1) the analysis increment is restricted to the wave components corresponding to the longest 50 1D waves and the grid-distance of the analysis increment becomes 16.2 km instead of 2.5 km of the forecast model. The

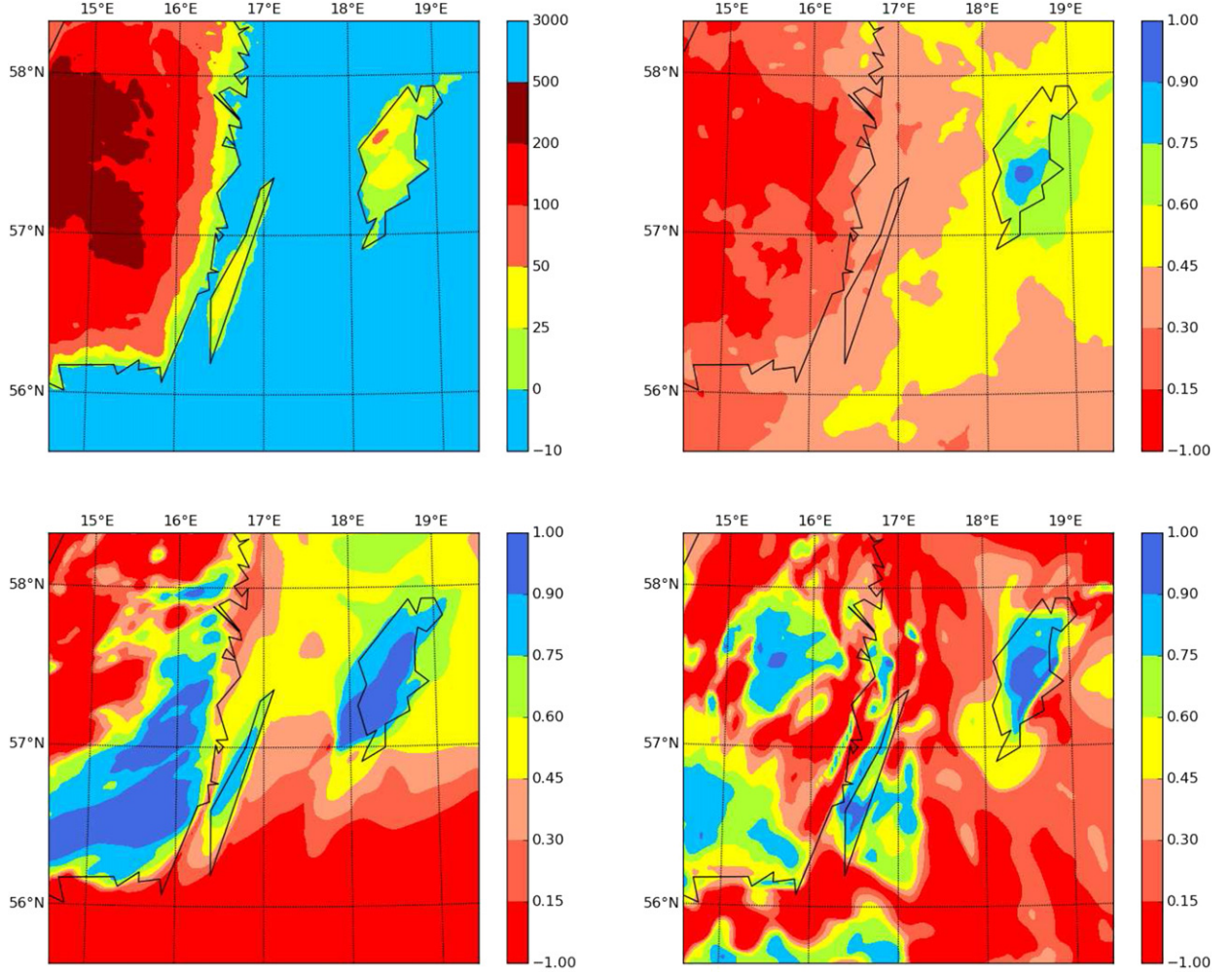


Fig. 15. Orography of a sub-domain over the Western Baltic Sea including the east coast of mainland Sweden and the Islands of Öland and Gotland. Horizontal correlations for temperature at model level 62 (a few hundred meters above the ground) with a reference point over central Gotland. Time averaged correlations for August 2011 (upper right) and averages over ensemble members only for 12 August 2011 06UTC +12 h (lower left) and 21 August 2011 06UTC +12 h (lower right).

shortest wave that can be described on a linear grid with 16.2 km resolution has a wave length of 32.4 km. Constructing the initial state for cntr_{50} , shorter scales come directly from the high resolution +3 h HARMONIE forecast.

Figure 16 shows an example of the analysis increment for temperature at model level 47 from the original 3-D Variational data assimilation (upper left plot) and those from restricted ones. The analysis increment corresponding to cntr_{100} is shown in the upper right plot, the analysis increment corresponding to cntr_{50} is shown in the lower left plot and the one corresponding to cntr_{10} is shown in the lower right plot. What exactly happens with the analysis increment when restricting it to larger scales depends on several factors: the dominating signal from “observations - minus- forecast” differences, the energy

Table 1. Summary of the conducted experiments and their abbreviations. M denotes number of 1D wave-components in the analysis increment, Δs denotes grid-size of the analysis increment and L_s denotes the length of the shortest resolved wave component of the analysis increment.

abbr	cntr	cntr ₁₀₀	cntr ₅₀	cntr ₁₀
M	324	100	50	10
Δs	2.5	8.1	16.2	81
L_s	5	16.2	32.4	162

spectra of the control variables and the balance relationships assumed in the BGE covariance model. In general, the amplitude of the analysis increment increases when the impact from observations is restricted to a narrower range of spatial scales for the control variables. See for

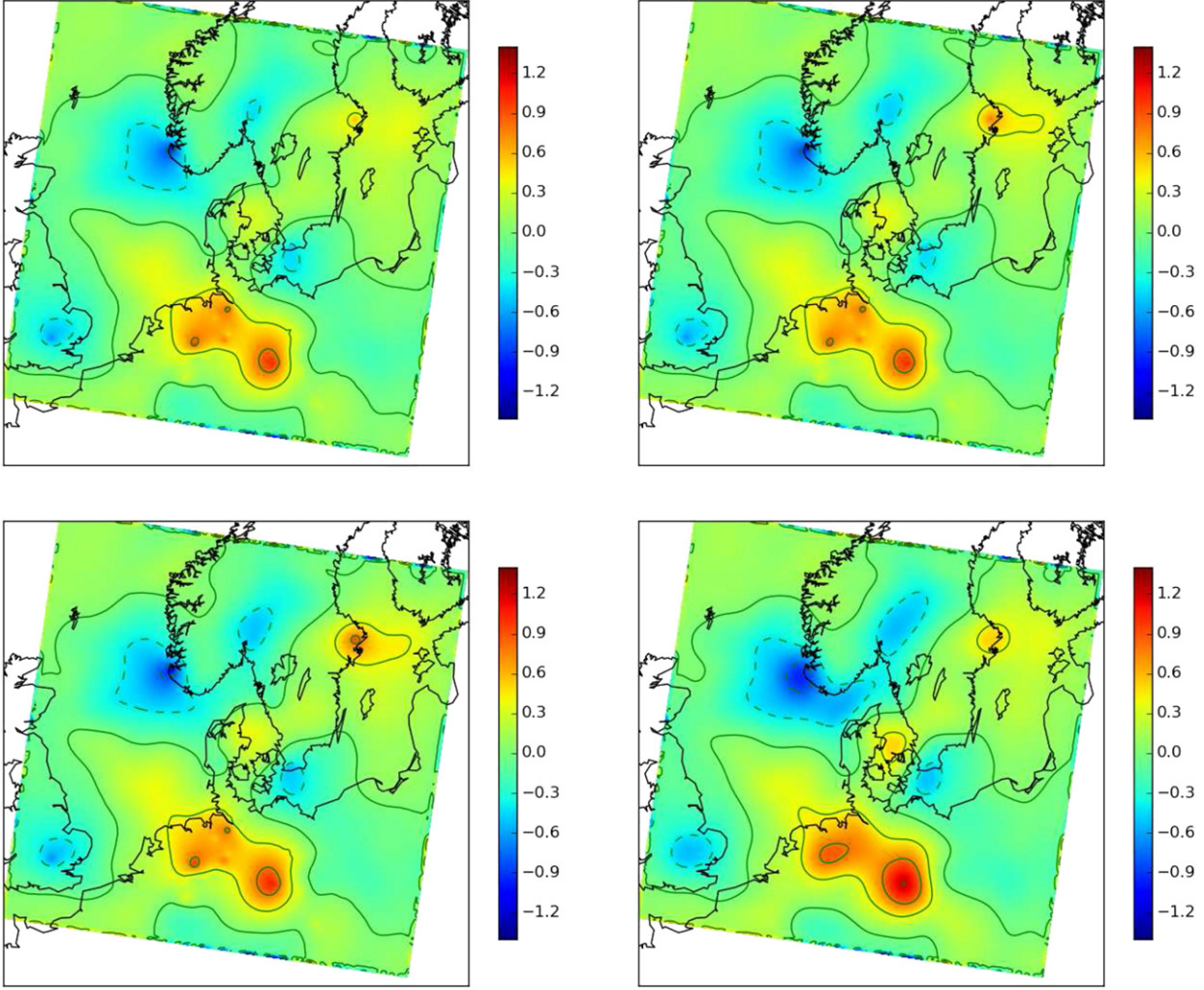


Fig. 16. Temperature analysis increment at model level 47 obtained from the original (upper left) and from the restricted data assimilation schemes: **cntr₁₀₀** (upper right), **cntr₅₀** (lower left) and **cntr₁₀** (lower right). The analysis increment is valid on the 2 July 2016 06 UTC.

example, temperature increment for subsequent experiments in the vicinity of Berlin. Balances, however, play an important role here. If geostrophic balance is important, as it is in the relation between mass and vorticity (see Figure 10), a larger scale temperature increment will get more energy. If on contrary, an observed signal comes from a humidity related observation, the meso-scale temperature humidity balance will play an important role instead (see Figure 12). In this case, the increment restricted to large scales may even loose energy, as it is for the temperature increment in the vicinity of Stockholm. Table 2 shows bias and RMSE of surface pressure forecasts as a function of the forecast lead time. The differences between the verification scores of the experiments are quite small and hardly significant, but our main emphasis is to show that no degradation of

forecast quality occur due to the restriction of the scale of the assimilation increment. It is remarkable, however, that the 10 longest waves of the analysis increment seem to be enough to maintain surface pressure forecasts up to 24 hours with a marginal loss of forecast quality. Restricting analysis increment to an 81 km resolution leads to less than 2% relative loss in the RMSE score at the initial time and 0.3% for the +18h forecast. The RMSE score for surface pressure was in fact improved already for the +6h forecast, although marginally, when the analysis increment was restricted to an 8.1 km resolution. Red colour in Table 2 indicates scores which were improved by restricting minimization to a lower resolution.

The results are even more pronounced for other model variables than surface pressure with smaller scale spatial

Table 2. Three weeks average of Bias and RMSE scores for surface pressure forecasts from **cntr** and **cntr₁₀₀** (top), **cntr₅₀** and **cntr₁₀** (bottom) experiments as a function of forecast lead time. Red colour denotes scores that are improved by restricting the analysis increment to a low horizontal resolution.

	cntr		cntr ₁₀₀	
	Bias	RMSE	Bias	RMSE
+hh				
+00	-0.025	0.408	-0.025	0.409
+06	-0.094	0.566	-0.092	0.565
+12	-0.100	0.621	-0.098	0.620
+18	-0.187	0.694	-0.181	0.689
+24	-0.172	0.746	-0.167	0.745
	cntr ₅₀		cntr ₁₀	
	Bias	RMSE	Bias	RMSE
+hh				
+00	-0.025	0.410	-0.027	0.415
+06	-0.095	0.565	-0.097	0.569
+12	-0.101	0.622	-0.105	0.624
+18	-0.183	0.693	-0.190	0.696
+24	-0.170	0.748	-0.175	0.748

variability. Figure 17 shows analysis of specific humidity at the model level 25 valid on the 2 July 2016 06UTC. Here, the analysis obtained from the original 3D-Variational scheme is shown to the left and the one obtained with the analysis increments restricted to an 81 km resolution (**cntr₁₀**) is shown on the plot to the right. The weather situation is dominated by a synoptic scale front with dry and cold air to the west and warm and moist air to the east. On the warm side of the front, a meso-scale activity is going on, note for example a meso-scale system with a lifting of warm air at the coast over the Gulf of Gdansk in the vicinity of Gdynia (54.5°N, 18.7°E). In the **cntr₁₀** experiment, when the analysis increment is restricted to the largest scale and meso-scales phenomena are allowed to evolve freely by the forecast model, the meso-scale system is stronger. The shape of system is anisotropic and is able to better follow the orography.

In Figure 18, we show time and domain averaged RMSE and Bias for 12h accumulated precipitation as a function of forecast lead time for the different experiments. The 12h accumulated precipitation scores are chosen in order to compensate for a double penalty error which is quite pronounced for such fields as precipitation. Accumulated scores allow to smooth the signal to certain extent and provide a more reliable verification when a RMSE is used as a score. Restricting minimization to lower resolution improves the quality of the accumulated precipitation forecasts. As we have already discussed in Section 6, the small-scale structures of the background error statistics derived assuming homogeneity, isotropy and stationarity are neither representing the current

weather situation nor are they a response to orographic forcing. They are simply a result of mathematical construction. Projecting observed information by such structure functions onto small scales brings in unrealistic signals. In case of the restricted minimization small scales structures of the initial state are not affected by the isotropic and homogeneous structures of the background error covariance. Instead, they come from a high-resolution first guess and they represent orographic forcing well. A well represented orography forcing is important for quality of the precipitation forecasts. The challenge that appears in the restricted minimization approach is how to constrain the small-scale structures that comes from the first guess such that they would be consistent with the large-scale features of the analysis field corrected from observations. Inconsistent large and small scales are likely to provoke spin-up processes that may appear as spurious precipitation. As we can see in Figure 18, the precipitation forecast from experiment **cntr₁₀₀**, with an analysis increment filtered to 8.1 km resolution, outperforms both the **cntr**, with analysis increment of 2.5 km resolution, and **cntr₅₀** and **cntr₁₀**, with analysis increment filtered to respectively 16.5 and 81 km horizontal resolution.

Small scale structures of the analysis increment of course help to get a closer fit of the initial state to observations. Figure 19 present time and domain averaged Bias and RMSE of relative humidity at 850 hPa as a function of forecast lead time for all four experiments. Experiment **cntr** (red) provides the best fit to observations at initial time. As it can be expected a close fit to observations at the analysis time does not necessarily guarantee a high forecast quality. The restricted minimization experiments **cntr₁₀₀** (green) and **cntr₅₀** (blue) outperform **cntr** at 12 hours, and **cntr₅₀**, with the analysis increment filtered to 16.2 km, provides a clearly superior +24 h forecasts. Experiment **cntr₁₀**, with analysis increments filtered to 81 km, has lost information relevant for moisture, and has the worst scores.

Figure 20 shows time, domain and forecast lead time averaged Bias and RMSE for wind speed profiles. In fact, the wind speed is the only field that is degraded when a restricted minimization is applied. The best verification scores belong to the **cntr** experiment (red curve), where the resolution of the analysis increment is the same as the resolution of the forecast model. According to geostrophic adjustment theory, winds adjust to mass on the large scales while mass adjusts to winds on the short scales. This is the reason why wind observations are so important on mesoscales.

7.2. The statistical background error balance

To get a better understanding of what impact the background error balance model (eqn. 4) has on the quality of

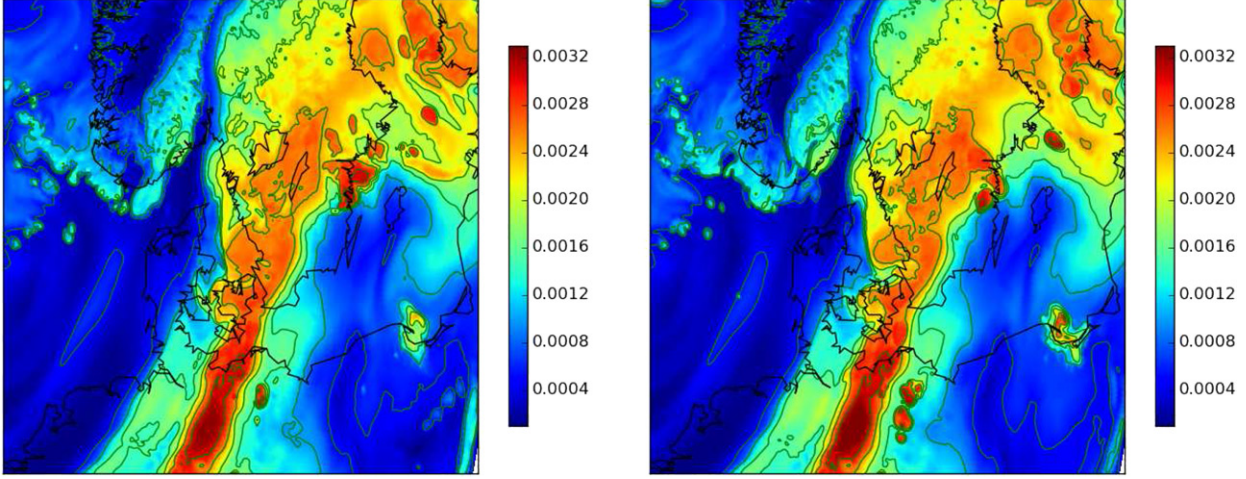


Fig. 17. The analysis of the specific humidity at model level 25 obtained from original (left plot) and from the restricted **cntr₁₀₀** (right plot) variational minimization schemes. The analysis is valid on the 2 July 2016 06UTC.

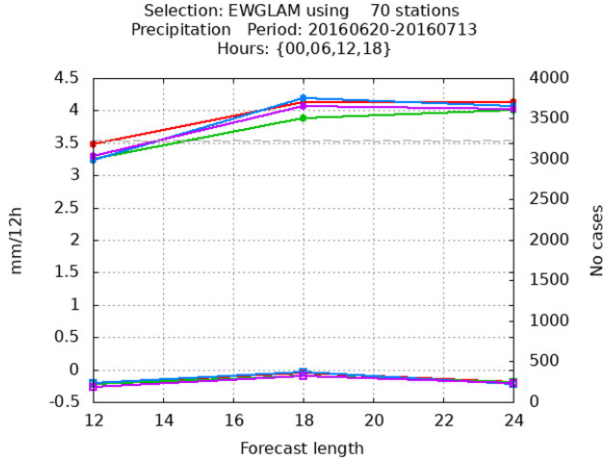


Fig. 18. Time and domain averaged Bias and RMSE scores of 12h accumulated precipitation for **cntr** (red), **cntr₁₀₀** (green), **cntr₅₀** (blue) and **cntr₁₀** (magenta) experiments as a function of forecast lead time. The dashed line shows the number of cases used in verification.

the forecasts we conduct another series of experiments. Now we run a restricted minimization (**cntr₁₀₀**, with the analysis increment filtered to 8.1 km) and we progressively switch off different balance operators in the background error covariance model. Table 3 gives a summary of the experiments and what balance operators are used in them. The experiment **noNPQRS** retain the balance operators *H* and *M* for the wind increments and treat temperature and humidity increments univariately. This means that the response of the mass and humidity to wind observations will be created during the model integration instead of being constructed during assimilation. In the same way the observations of mass and humidity

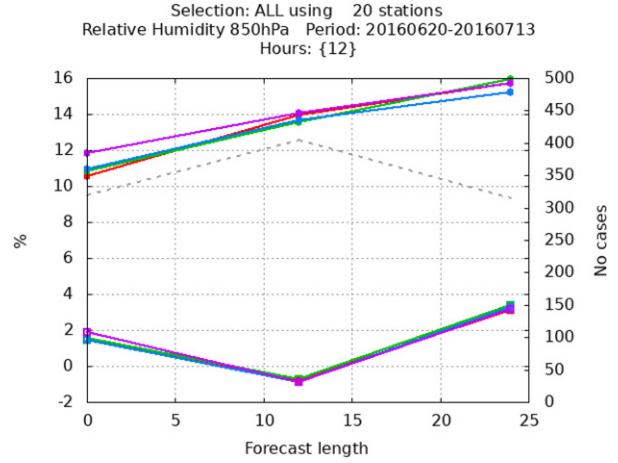


Fig. 19. Time and domain averaged Bias and RMSE scores of RH at 850hPa level for **cntr** (red), **cntr₁₀₀** (green), **cntr₅₀** (blue) and **cntr₁₀** (magenta) experiments as a function of forecast lead time. The dashed line shows the number of cases used in verification.

will not have a direct impact on the wind field, but this will similarly be created through adjustment processes during the forecast model integration. The experiment **noPRS** adds the geostrophic balance through the linear balance operators, *N* (a linear balance between vorticity and mass) and *Q* (a linear balance between vorticity and humidity). The contribution to the increments due to *N* and *Q* is small at short scales but increases for the larger scales. The experiment **noS** adds two more balance operators when constructing temperature and humidity increments, *P* (linear relationship between mass and divergent wind) and *R* (linear relationship between humidity and divergent wind). The contribution from divergent wind is

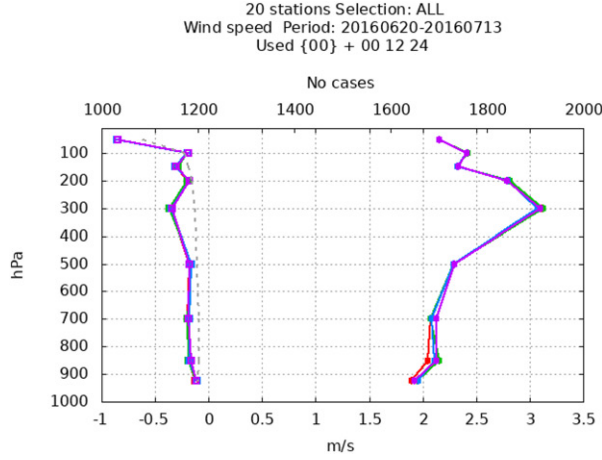


Fig. 20. Wind field verification scores (valid at 00 UTC): Bias and RMSE scores of wind speed; **cntr** (red) **cntr₁₀₀** (green), **cntr₅₀** (blue) and **cntr₁₀** (magenta) experiments; scores are averaged over time, domain and forecast lead time. The dashed line shows the number of cases used in verification.

strongest on the short scales. Finally, **cntr₁₀₀** in addition to balance operators H , M , N , Q , P and R includes S , a linear dependency between humidity and mass. The contribution to the humidity increments due to S is dominating on meso-scales.

Figure 21 shows the domain and time averaged Bias and RMSE scores of wind speed at 850 hPa valid at 00 UTC for all five experiments. Recall that restricted minimization has degraded the quality of wind speed forecasts in the lower troposphere. It is interesting to note that by omitting the linear balance relation S between humidity and unbalanced mass (**noS**, **noPRS** and **noNPQRS** experiments), the quality of the wind speed forecasts is improved. Being derived based on the assumptions of homogeneity, isotropy and stationarity the linear balance relationship between temperature and humidity is likely to be less representative, in particular in cases of strong convective developments. Applied during the data assimilation homogeneously and isotropically such a balance will amplify noise even more through feedback mechanisms.

It is interesting to see the impact of the background error covariance model on the development of the meso-scale and synoptic scale systems. Figure 22 shows the analysis of specific humidity at model level 25 valid on the 2 July 2016 06UTC, the same case as on Figure 17. The analysis of the specific humidity obtained from the restricted minimization **cntr₁₀₀** is shown on the left upper plot. The other three plots are the analyses from experiments with progressively switched off balances, **noS** (upper right), **noPRS** (lower left) and **noNPQRS** (lower right). The progressively switched off balances weaken

Table 3. Summary of the conducted statistical balance experiments and their abbreviations. “+” sign denotes what balance operators are included in the background error covariance model for each of experiments. See eqn. (4) for an explanation of the different balance operators (H , M , N , P , Q , R , S).

Operator	cntr ₁₀₀	noS	noPRS	noNPQRS
H	+	+	+	+
M	+	+	+	+
N	+	+	+	–
P	+	+	–	–
Q	+	+	+	–
R	+	+	–	–
S	+	–	–	–

meso-scale activity (for example, the meso-scale system on the coastline of Gulf of Gdansk or the meso-scale system north of Berlin) but strengthen synoptic scale activity in the north over Gulf of Bothnia. When the meso-scale mass-humidity balance S is omitted (experiment **noS**, upper right plot) the meso-scale activity remains relatively strong and the system is able to better follow orography. Even if the mass and humidity are obviously correlated, the homogeneity and isotropy in the construction of the analysis increment have a negative impact on the evolution of the meso-scale system. When the relations between the mass and humidity and divergent wind, linear operators P and R , are omitted in addition (**noPRS**, lower left plot), an important mechanism of convergent-divergent flow is destroyed and the meso-scale systems weaken considerably. When temperature and humidity are treated univariately and temperature and humidity observations do not influence the wind field (**noNPQRS**, lower right plot) the synoptic development becomes stronger and the meso-scale systems weaken even more. Deeper investigations are needed in order to understand the mechanisms behind, however we have noticed that jet-level winds become somewhat weaker and slightly displaced in case of **noNPQRS**. This may be due to the importance of geostrophic balance, i.e. the strong correlation between the mass and wind fields, on large scales (see Figure 10). When geostrophic balance is omitted the larger scale environmental conditions for development of meso-scale systems are changed.

7.3. Impact of ensemble generation methodology on background error statistics

From investigations performed in section 5 we see that the structure functions are sensitive to the quality of the ensemble even if the homogeneity and isotropy assumptions impose a strong smoothing effect. The spectral

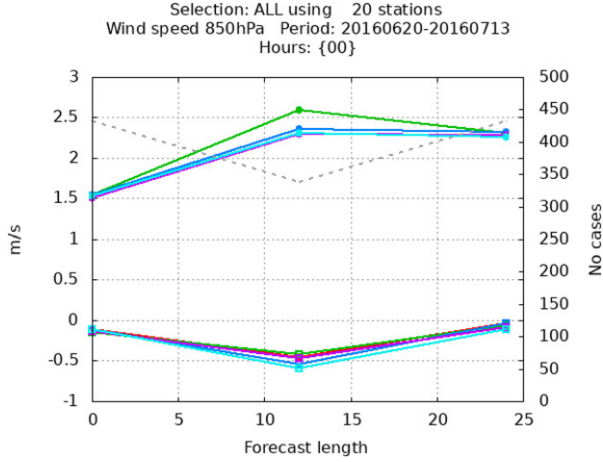


Fig. 21. Time and domain averaged Bias and RMSE scores of wind speed at the 850 hPa level for the **cntr** (red), **cntr₁₀₀** (green), **noS** (blue) and **noPRS** (magenta) and **noNPQRS** (cyan) experiments as a function of forecast lead time. The dashed line shows the number of cases used in verification.

density for unbalanced humidity (see Figure 6) looks different when it is computed from the ensemble based on downscaling or the ensemble based on HARMONIE EDA. Even if longer forecast lead times make structure functions computed from both ensembles quite similar, one can see that certain adjustment processes on meso-scales continue even after 12 h of forward integration with the HARMONIE model. Both ensemble sets, the ensemble of downscaling and of the HARMONIE EDA, are generated using the same methodology, i.e. Ensemble of Data Assimilation (EDA) runs. Following this methodology, the variability of the ensembles originates from variability imposed in observation space. Namely, observations are perturbed independently in different locations according to a Gaussian PDF with a standard-deviation of the assumed observation error. Different realisations of such artificial observations are assimilated to create different ensemble members. These ensemble members will serve as initial conditions for downscaling or for a HARMONIE EDA ensemble. The difference is that for HARMONIE EDA a HARMONIE first-guess will be used to assimilate the observations, while in the downscaling ensemble the data assimilation is performed in a low-resolution IFS ECMWF model. One may ask to what extent the design of the observing network, for example what scales and what quantities are observed, impacts the derived structure functions that are scale-dependent. To get an insight into this problem, we derived structure functions from two further ensemble sets and compared them with the structure functions described in section 5. The first set of structure functions, abbreviated **3h 3DVAR EDA MetCoOp**, uses EDA

methodology with a 3 h update frequency and a high-resolution observing network, that includes radar reflectivities, ground-based GNSS (Global Navigation Satellite System) zenith total delay measurements and ATOVS micro-wave satellite radiances in addition to the conventional observations. The second set of structure functions, abbreviated **BRAND**, is derived from an ensemble, the generation of which does not include perturbation of observations. **BRAND** (B-matrix RANdom) perturbations are generated as Gaussian ($N(0,1)$) random numbers in the entire control vector space and are projected to the physical space of the model state applying the square-root of background error covariance (see eqn. 3). The perturbations are added to the high-resolution control analysis. Both the **BRAND** ensemble and the **3h 3DVAR EDA MetCoOp** ensemble are generated every 3 h.

Figure 23 shows some diagnostics derived from ensembles generated with and without perturbing observations, namely the percentage of the surface pressure variance explained by vorticity and unbalanced divergence computed from **BRAND** and **EDA** ensembles. Structure functions derived from the **BRAND** based ensemble (red curves) are plotted against those derived from **6h 3DVAR EDA CONV** (blue curves, to the left) and those derived from **3h 3DVAR EDA MetCoOp** (blue curves, to the right). The forecast lead time is 3 h for all three ensemble sets. First of all, we can notice a gravity wave signature, a significant correlation between surface pressure (integrated mass-field) and unbalanced divergence at scales below 25 km, that are much stronger for 3 h update cycle (blue curves, right plot) than for 6 h update cycle (blue curves, left plot). Note that the **BRAND** ensemble is generated around the control that performs ordinary, unperturbed data assimilation every three hours. As one could expect the gravity wave signature is slightly stronger for the **BRAND** ensemble than for **EDA** because **BRAND** perturbations are added throughout the entire model space. In addition, one may notice a strong correlation between surface pressure and unbalanced divergence at the $2\Delta s$ scale (80% of explained surface pressure variance), due to numerical truncation and aliasing on the linear grid. However, the largest differences between **BRAND** and **EDA** based structures functions concern decomposition of the flow on the divergent and rotational ones, in particular what role the geostrophic balance plays on meso-scales. Note a peak in the P_b curve (the explained variance of surface pressure through vorticity via linearized geopotential P_b) at scales around 200-600 km for statistics computed from **EDA** ensembles. The "peak" is high and narrow in case of **6h 3DVAR EDA CONV** (blue curve, left plot) and smooth and low in case of **3h 3DVAR EDA MetCoOp** (blue curve, right plot). A similar peak does not appear for statistics

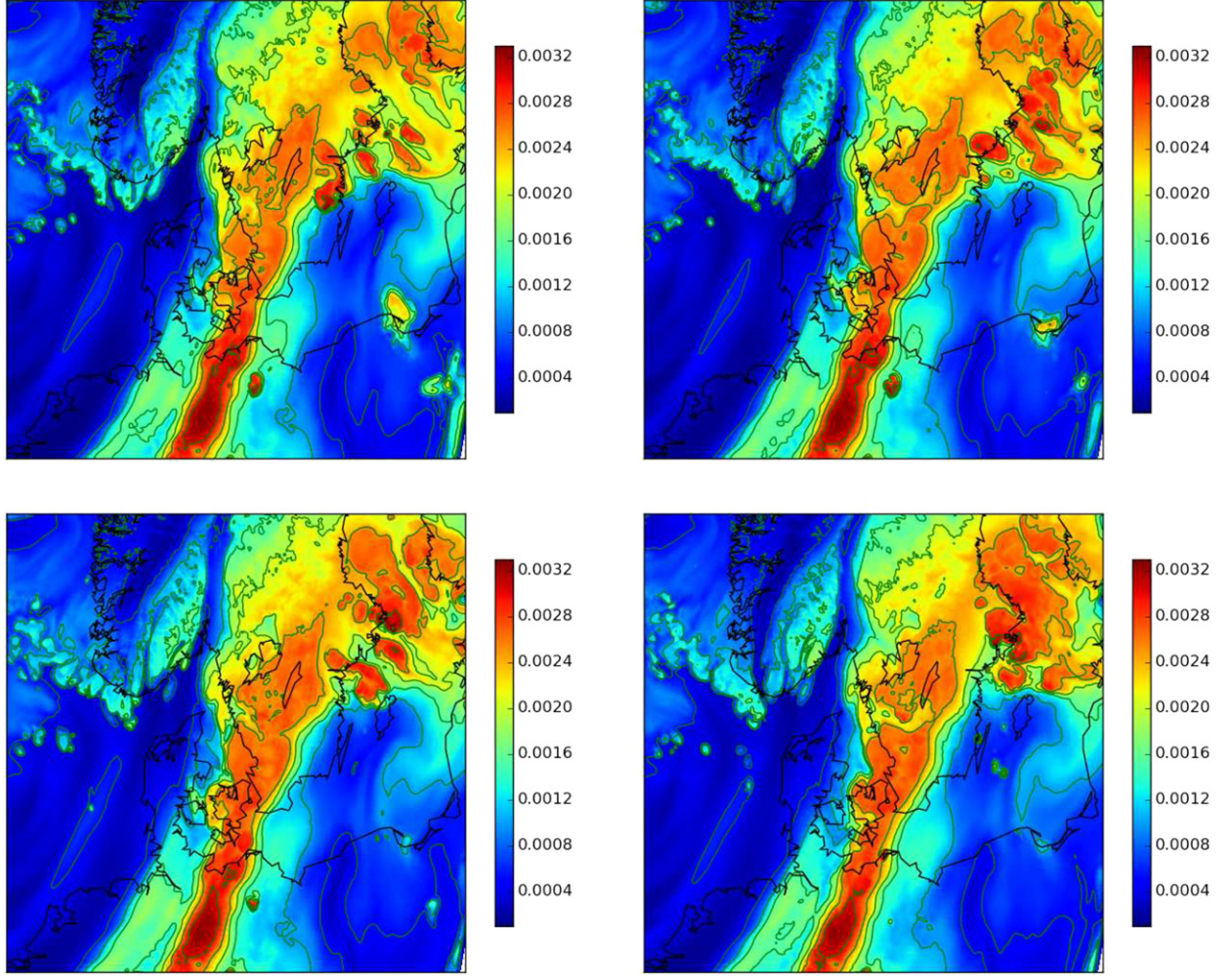


Fig. 22. Analysis of specific humidity on the model level 25 obtained from **cntr₁₀₀** (upper left) and from the experiments with progressively switched balance operators **noS** (upper right), **noPRS** (lower left) and **noNPQRS** (lower right). The analysis is valid on the 2 July 2016 06 UTC.

computed from the **BRAND** ensemble (red curve) where the signature of geostrophic balance is very clean as an increase in correlation between surface pressure and vorticity at synoptic scales, above 800 km. We may suggest that such a “peak” is a result of the observation perturbation methodology for ensemble generation. Properties of short-range forecasts reflect the way in which an ensemble was generated. In **6h 3DVAR EDA CONV** uncorrelated perturbations were injected at the scales corresponding to the scales of the conventional observing system that is dominated by mass field measurements (surface pressure) over land only. The background error covariance model is used to transform this unstructured noise into the entire model space. At large scales, wind adjust to mass following the geostrophic adjustment theory that underlies the background error covariance model. The result manifests itself in the peak of P_b curve

located at scales of the conventional observing network, that reflects for example the distribution of land and sea in model domain. The observing network perturbed in the **3h 3DVAR EDA MetCoOP** is much more diversified and includes high-resolution observing networks of radars and GNSS. The same background covariance model is applied to transform the injected noise into the entire model space in this case also, but the spread is injected on wider range including smaller scales. As a result, the “peak” is lower and smoother.

Probably a long-time integration (see Figure 11) could allow to spread the energy injected at specific spatial scales to the entire model space due to non-linear model dynamics, but a few hours of model integration of the regional model is obviously not enough. On the contrary, **BRAND** perturbations inject energy into the entire model space and are able to maintain the same relationship

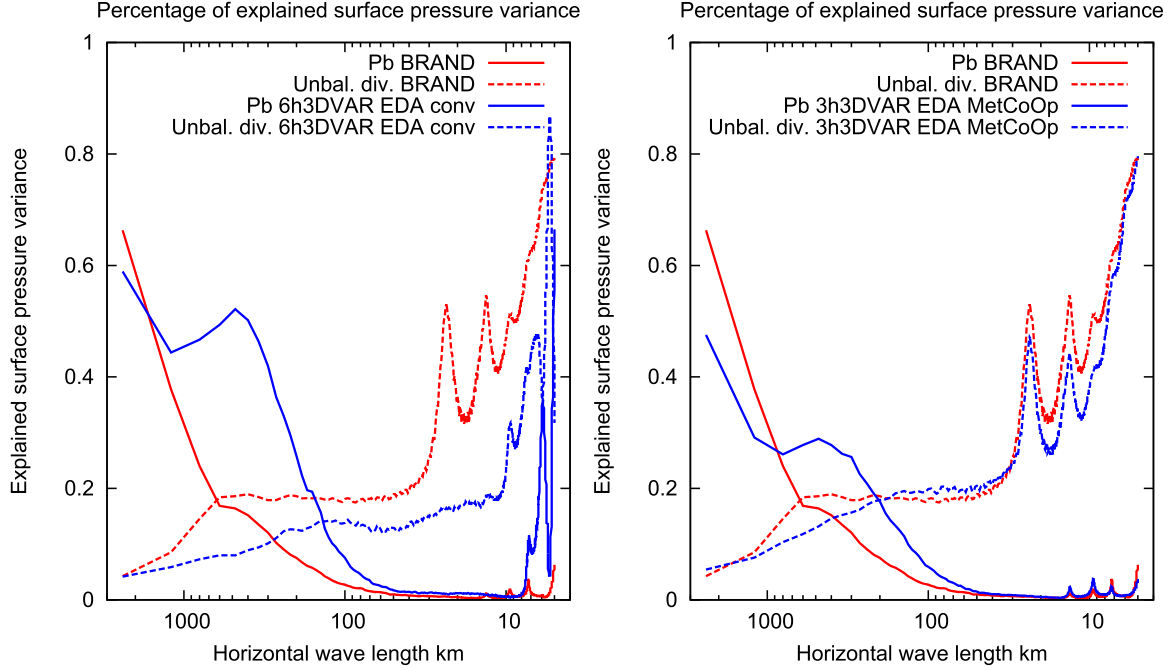


Fig. 23. Percentage of surface pressure variance explained by vorticity and unbalanced divergence derived from the **EDA** (blue curves) and **BRAND** (red curves) ensemble sets. **BRAND** based structure functions are plotted against **6h 3DVAR EDA CONV** (left) and against **3h 3DVAR EDA MetCoOP** (right).

between the mass-field and the unbalanced divergence throughout the mesoscales. As we have seen in Section 7.2 the mass- and wind-field relationship and its decomposition into divergent and rotational flow increments impacts the development of meso-scale systems. On average 20% of the variability of the integrated mass field can be explained by the variability of unbalanced divergence throughout the mesoscales. This relatively low value is probably due to smoothing effects related to the assumptions of homogeneity and isotropy for deriving the statistics from fields with more narrow elongated anisotropic structures.

8. Discussion

The results presented in this paper are primarily relevant for the **HARMONIE** forecasting system, although some conclusions might be quite general. The presented results are furthermore obtained from experiments that are subject to certain simplifications, assumptions and known deficiencies that might influence the results:

1. We have noted a rather weak growth of the **HARMONIE** ensemble spread between +3h (or even +1h) and +12h of the model integration. It may be speculated whether this problem is associated with the lateral boundary conditions (LBCs), that are of strong importance over the small **HARMONIE**

model domain. The LBCs are taken from 3-12h forecasts of the ECMWF 4D-Var based EDA runs at 06 UTC and 18 UTC. These 3-12h forecasts are efficiently within the 12h 4D-Var assimilation windows, and could be considered as analyses influenced by observations. This may explain the weak growth of the ensemble spread over the small **HARMONIE** domain.

2. The horizontal 2-dimensional Fourier transforms used in the formulation of the background error covariance model require an extension zone in both horizontal directions when they are applied over the limited area model domain in order to obtain a bi-periodicity. In the current version the extension zone is eleven grid distances wide (27.5 km). Such a narrow extension zone may lead to strong artificial gradients close to the boundaries. We cannot exclude that a wrap-around of information in the analysis increment, due to the periodicity assumption, and artificial gradients may impact the results. The sensitivity of the forecast quality to the width of the extension zone is a subject of on-going work.
3. We have noticed that **HARMONIE** forecasts are contaminated by a high-frequency and small scale ($2\Delta s$ – $5\Delta s$) noise. The current configuration of the **HARMONIE** model employs a linear grid. Numerical truncation and aliasing from non-linear processes

might be one of the reasons for this noise. The lack of initialisation might contribute as well. A statistical balance constraint is included in the formulation of the background error covariance model. One should be aware that this statistical balance formulation was derived with a perspective of a synoptic scale system and forecast lead time between 12 and 48 hours, exceeding the time-scales of adjustment processes. In mesoscale systems used to forecast convective scale phenomena, the interest is in forecast of shorter lead time. In assimilation cycling mode the frequent update of initial conditions is applied and 3h forecasts are used to bring in background information even if geostrophic adjustment processes are still ongoing. The development of methodologies to take adjustment processes into account during assimilation is outside the scope of this paper. Furthermore, a digital filter initialization is available in the HARMONIE forecasting system, but it is less appropriate since it is based on cut-off frequencies of the order of $(3\text{-}6\text{h})^{-1}$.

4. There are several other pragmatic decisions that have been taken in order to make the work affordable. For example, only conventional observations are assimilated in the experiments for which we have presented verification scores in [section 7](#). We wanted to avoid assimilation of high-resolution observations because of known deficiencies in the treatment of correlated observation errors in our system.
5. In the experiments described in [Table 3](#) we decided not to re-adjust the error variance for the unbalanced control vector part even if the contribution from a balanced part was switched off. This imposes that the amplitude of the analysis increment for the mass component in **noPRS** is smaller than the amplitude of the analysis increment in **cntr₁₀₀** even if the same observations are used. We were reluctant to add more noise to the system that would happen if we would increase the variance of the unbalanced control vector components. Instead we have decided to see if and how the missed part of the analysis increment will be re-established through the model integration. Furthermore, we have switched off balance operators through the entire model spectrum including large scales. We could note that experiment **noNPQRS** resulted in somewhat weaker jet-stream and somewhat stronger mesoscale activity than those from experiment **cntr₁₀₀**. It might be that this is an impact of a smaller amplitude of the analysis increment rather than the formulation of the statistical background error covariance model.
6. The generation of structure functions in [section 5](#) and the computation of verification scores in [section 7](#)

were carried out over different model domains and different time periods. Due to homogeneity and isotropy assumptions, the exact geographical position of the domain plays a marginal role as soon as the same weather regime is dominating.

7. The generation of ensemble variability through observation perturbations in the Ensemble Data Assimilation (EDA) approach and the direct generation of model space perturbations in the BRAND (B-matrix RANDom) approach are both dependent on the applied BGE statistics (the B matrix). Longer time integrations may help to establish ensemble perturbations more representative of “real” forecast errors (as in the NMC method). An iterative approach to determine new BGE statistics could also be investigated, although convergence is uncertain.

9. Conclusions and outlook

The relevance of climatological BGE statistics for meso-scale data assimilation has been investigated. We have found limitations in the usefulness of such climatological BGE statistics, both with regard to basic covariance modelling assumptions and also with regard to the simulation techniques used to estimate parameters of BGE covariance models. From the results presented and discussed in this study, we can conclude that the BGE statistics derived by simulation through Ensemble Data Assimilation are more realistic than the corresponding statistics derived by downscaling of ECMWF EDA data, since the downscaling data are associated with a model spinup, that is not completed even after 12 hours of model integration.

We have also seen that structure functions derived from an ensemble of model state differences are very sensitive to the way the ensemble is generated. A few hours of model integration with a high-resolution model are not enough to spread the variability injected on a particular scale to the entire model space. For example, when a perturbation of observations is used to inject a spread into the ensemble, the structure functions are contaminated by the geometry of the observing network. In case of spatially incomplete observing networks this might project arbitrary background error structures into the ensemble instead of sampling forecast errors developed by the model.

Furthermore, we have illustrated that the generation of BGE statistics, in particular statistical balance operators, with model simulation techniques is very sensitive to small scale noise generated during the model simulations. For this reason, one needs to restrict the use of such BGE statistics to spatial scales that are resolved properly

by the model, i.e. scales larger than 3–5 Δs , with Δs being the model grid resolution.

Finally, we have illustrated the effects of the basic assumptions of stationarity, homogeneity and isotropy that are generally applied in the generation of BGE statistics. Spatial covariances derived under these assumptions were contrasted against simplified examples of spatial covariances obtained by ensemble averaging only, preserving the signals from forecast errors of the day. This indicates that it is likely to be favourable to apply data assimilation with ensemble background error statistics obtained from ensemble averaging, like in ensemble Kalman filters or in hybrids between variational and ensemble data assimilation techniques. At the same time, one should be aware that the quality and the affordable number of members of the ensemble systems for operational applications, in particular for short forecast ranges, is premature and not sufficient, despite a number of different proposed measures (for example lagging, inflation and localisation). Climatological background error statistics will furthermore continue to play an important role in the operational data assimilation in order to assure enough search directions during the minimization process (Gustafsson and Bojarova, 2014, Lorenc and Jardak, 2018).

Acknowledgements

Many thanks to Magnus Lindskog and Martin Ridal for help with generation and access to ensemble information and many thanks to Nedjeljka Zagar for inspiration to start this study.

Disclosure statement

Abdelwaheb Hannachi was the responsible Subject Editor for the handling of this paper.

References

- Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J. co-authors. 2017. The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. *Mon. Wea. Rev.* **145**, 1919–1935. doi:10.1175/MWR-D-16-0417.1
- Berre, L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.* **128**, 644–667. doi:10.1175/1520-0493(2000)128<0644:EOSAMF>2.0.CO;2
- Blazika, V., Gustafsson, N. and Zagar, N. 2015. The impact of periodization methods on the kinetic energy spectra for limited-area numerical weather prediction models. *Geosci. Model Dev.* **8**, 87–97. doi:10.5194/gmd-8-87-2015
- Bubnova, R., Hello, G., Benard, P. and Geleyn, J.-F. 1995. Integration of the fully elastic equations cast in the hydrostatic pressure terrain following coordinate in the framework of the ARPEGE/Aladin NWP system. *Mon. Wea. Rev.* **123**, 515–535. doi:10.1175/1520-0493(1995)123<0515:IOTFEE>2.0.CO;2
- Fischer, C., Montmerle, T., Berre, L., Auger, L. and Ștefănescu, S. E. 2005. An overview of the variational assimilation in the Aladin/France NWP system. *Q. J. R. Meteorol. Soc.* **131**, 3477–3492. doi:10.1256/qj.05.115
- Gustafsson, N. and Bojarova, J. 2014. Four-dimensional ensemble variational (4D-En-Var) data assimilation for the High Resolution Limited Area Model (HIRLAM). *Nonlin. Processes Geophys.* **21**, 745–762. doi:10.5194/npg-21-745-2014
- Gustafsson, N., Janjic, T., Schraff, C., Leuenberger, D., Weissman, M. co-authors. 2018. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Q. J. R. Meteorol. Soc.* **144**, 1218–1256. doi:10.1002/qj.3179
- Frehlich, R. and Sharman, R. 2004. Estimates of turbulence from numerical weather prediction model output with applications to turbulence diagnosis and data assimilation. *Mon. Wea. Rev.* **132**, 2308–2324. doi:10.1175/1520-0493(2004)132<2308:EOTFNW>2.0.CO;2
- Haugen, J.-E. and Machenhauer, B. 1993. A spectral limited-area model formulation with time-dependent boundary conditions applied to the shallow-water equations. *Mon. Wea. Rev.* **121**, 2618–2630. doi:10.1175/1520-0493(1993)121<2618:ASLAMP>2.0.CO;2
- Hägmark, L., Ivarsson, K.-I., Gollvik, S. and Olofsson, P.-O. 2000. Mesan, an operational mesoscale analysis system. *Tellus* **52A**, 2–20.
- Lindborg, E. 1999. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.* **388**, 259–288. doi:10.1017/S0022112099004851
- Lorenc, A. C. and Jardak, M. 2018. A comparison of hybrid variational data assimilation methods for global NWP. *Q. J. R. Meteorol. Soc.* **144**, 2748–2760. doi:10.1002/qj.3401
- Nastrom, G. D. and Gage, K. S. 1985. A climatology of atmospheric wave number spectra observed by commercial aircraft. *J. Atmos. Sci.* **42**, 950–960. doi:10.1175/1520-0469(1985)042<0950:ACOWS>2.0.CO;2
- Parrish, D. and Derber, J. 1992. The National meteorological center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763. doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2
- Sattler, K. and Huang, X.-Y. 2002. Structure function for 2 meter temperature and relative humidity in different horizontal resolution. *Tellus* **54A**, 14–33.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Benard, P. co-authors. 2011. The AROME-France convective scale operational model. *Mon. Wea. Rev.* **139**, 976–991. doi:10.1175/2010MWR3425.1
- Skamarock, W. C., Park, S.-H., Klemp, J. B. and Snyder, C. 2014. Atmospheric kinetic energy spectra from global high-resolution nonhydrostatic simulations. *J. Atmos. Sci.* **71**, 4369–4381. doi:10.1175/JAS-D-14-0114.1