



# Four-dimensional ensemble variational (4D-En-Var) data assimilation for the High Resolution Limited Area Model (HIRLAM)

N. Gustafsson<sup>1</sup> and J. Bojarova<sup>2</sup>

<sup>1</sup>Swedish Meteorological and Hydrological Institute, 60176 Norrköping, Sweden

<sup>2</sup>Norwegian Meteorological Institute, P.O. Box 43 Blindern, 0313 Oslo, Norway

Correspondence to: N. Gustafsson (nils.gustafsson@smhi.se)

Received: 22 November 2013 – Revised: 1 April 2014 – Accepted: 19 May 2014 – Published: 14 July 2014

**Abstract.** A four-dimensional ensemble variational (4D-En-Var) data assimilation has been developed for a limited area model. The integration of tangent linear and adjoint models, as applied in standard 4D-Var, is replaced with the use of an ensemble of non-linear model states to estimate four-dimensional background error covariances over the assimilation time window. The computational costs for 4D-En-Var are therefore significantly reduced in comparison with standard 4D-Var and the scalability of the algorithm is improved.

The flow dependency of 4D-En-Var assimilation increments is demonstrated in single simulated observation experiments and compared with corresponding increments from standard 4D-Var and Hybrid 4D-Var ensemble assimilation experiments. Real observation data assimilation experiments carried out over a 6-week period show that 4D-En-Var outperforms standard 4D-Var as well as Hybrid 4D-Var ensemble data assimilation with regard to forecast quality measured by forecast verification scores.

lation and, in addition, statistical information about the uncertainty of this background field may be utilized. This was realized already by Berghorsson and Döös (1955), who applied a successive correction spatial interpolation to a quasi-geostrophic barotropic forecast model. The successive correction interpolation was essentially based on weights given to the observation-minus-background deviations being proportional to the distances between the observation positions and the grid points. During the 1960s and 1970s more advanced spatial interpolation techniques, for example the statistical interpolation or OI (optimum interpolation, Gandin, 1963), were introduced and brought into operational use in three-dimensional versions, including also the idea of balancing between mass and wind field information (Rutherford, 1976; Gustafsson, 1981; Lorenc, 1981).

Variational data assimilation (Sasaki, 1958; Le Dimet and Talagrand, 1986; Lewis and Derber, 1985) provides an efficient framework for data assimilation with NWP models, since the possibility to apply dynamical and physical constraints for the data assimilation is introduced. In three-dimensional variational data assimilation (3D-Var) a cost function measuring the distance to the observations and the distance to the background field is minimized. In its simplest form, 3D-Var is equivalent to OI, with the exception that there are no requirements for data selection, while OI generally has to rely on local data selection schemes to find the observed values to influence each grid point of the model. One of the most important reasons of 3D-Var replacing OI is that 3D-Var can easily use non-linear observation operators such that satellite radiance data can be directly assimilated. 3D-Var is applied to global NWP models (Parrish and Derber, 1992; Courtier et al., 1998) and also to regional NWP

## 1 Introduction

Data assimilation for numerical weather prediction (NWP) is the process of using observations to create initial conditions for NWP models. The number of observed values at any particular moment of time is generally much smaller than the number of model state variables of the NWP model. For this reason, a priori information has to be taken into account. In most data assimilation schemes, a short-range forecast from the NWP model, valid at the time of the observations, is used as a *background* field for the data assimilation

models (Gustafsson et al., 2001; Lindskog et al., 2001). In four-dimensional variational data assimilation (4D-Var) the model initial state is found by minimizing a cost function measuring the distance to observations from an assimilation time window (e.g. 6 h) including also the forecast model as a strong constraint for the model development over the assimilation time window. The concept of incremental 4D-Var (Courtier et al., 1994) made it possible to practically solve the computational problem of 4D-Var by introducing simplification and linearization of the forecast model and the observation operators and by introducing a reduced spatial resolution for the 4D-Var minimization. Incremental 4D-Var schemes have been successfully introduced operationally for global NWP models (Rabier et al., 2000) as well as for regional high-resolution models (Kawabata et al., 2007; Gustafsson et al., 2012).

The application of 4D-Var for operational NWP has been very successful, but there are a few problems associated with its further development. Firstly, the application of a static background error constraint, similar to those applied in 3D-Var, at the start of the assimilation window limits the possibilities to describe flow-dependent background error structures, for example those associated with baroclinicity. This can be alleviated by introduction of longer 4D-Var assimilation time windows, allowing the implicit assimilation structure functions to develop these flow dependencies. With longer time windows it will, however, be necessary to treat model errors more efficiently, for example by introducing the weak constraint 4D-Var (Trémolet, 2006). Secondly, the incremental 4D-Var necessitates the development of the tangent linear (TL) and adjoint (AD) versions of the NWP forecast models. This requires substantial development and maintenance efforts and, furthermore, the computational cost of incremental 4D-Var is dominated by the costs for the TL and AD model integrations.

The Kalman filter (KF, Kalman, 1960) provides a framework for estimating the uncertainty of forecast model states for linear forecast models, and the extended Kalman filter (EKF, Kalman and Bucy, 1961) extends the Kalman filter theory to weakly non-linear models and observation operators. The KF and the EKF cannot be applied directly to large-scale NWP models, due to the huge size of the corresponding model state covariance matrix to be handled. The ensemble Kalman filter (EnKF, Evensen, 1994) is essentially a Monte Carlo approximation to the Kalman filter. The model uncertainty is represented through an ensemble of model states, from which the required forecast error covariances are estimated. The implementation of EnKF can roughly be subdivided into three different approaches: the EnKF with perturbed observations, which samples different sources of uncertainty (Evensen, 1994; Houtekamer and Mitchell, 2001), the square-root EnKF, which directly handles covariances involved (Tippett et al., 2003) and the reduced-rank KF (Heemink et al., 2001; Cohn and Todling, 1996), which focuses on a low-dimensional approximation

of the model state covariances. One main advantage of the EnKF is the possibility to describe flow-dependent uncertainties of the background model state. Another advantage of EnKF, as compared with 4D-Var, is that it is no longer necessary to develop and maintain TL and AD versions of the NWP forecast models. These advantages have the price of a low rank of the background error covariance for full-scale NWP data assimilation, due to the necessity to apply a limited number ( $\sim 10$ – $100$ ) of ensemble members. The low rank of the background error covariance matrix can be handled to some extent by, for example, application of localization to avoid spurious correlations at large distance separations, caused by sampling errors. Additive and multiplicative inflation techniques can also be used to compensate for the effects of the low rank.

The advantages and disadvantages of EnKF, as compared to 4D-Var, were discussed by Lorenc (2003), Kalnay et al. (2007), and Gustafsson (2007). A synthesis of this discussion is the recognition of the need to develop combinations of the robust and full rank 4D-Var algorithms and the ensemble-based methods that are able to describe flow dependencies. A first step in this direction is the development of hybrid variational ensemble data assimilation methods (Hamil and Snyder, 2000; Lorenc, 2003; Wang et al., 2008a, b; Clayton et al., 2013; Gustafsson et al., 2014) that incorporate, for example, ensemble-based error covariances within the framework of variational data assimilation. A further step is to replace the TL and AD model integrations in 4D-Var with the application of ensembles of non-linear model trajectories over the data assimilation window (4D-En-Var, Liu et al., 2008, 2009; Liu and Xiao, 2013; Buehner et al., 2010a, b; Desroziers et al., 2014). The subject of the present paper is to present the 4D-En-Var developed for the High Resolution Limited Area Model (HIRLAM, Undén et al., 2002). A more detailed review of 4D-En-Var developments is provided in Sect. 2. Sections 3 and 4 focus on the HIRLAM implementation of 4D-En-Var. Section 5 presents results of experiments, both with single simulated observations to demonstrate the abilities of the algorithm, and with real observations to show the performance of 4D-En-Var. Section 6 discusses the computational performance and challenges with the scheme and the concluding remarks are presented in Sect. 7.

## 2 Review of 4D-En-Var developments

The following cost function is minimized with respect to the assimilation increment  $\delta\mathbf{x}$  in incremental 4D-Var:

$$J = J_b + J_o = \frac{1}{2} (\delta\mathbf{x})^T \mathbf{B}^{-1} \delta\mathbf{x} + \frac{1}{2} \sum_{t_k=t_0}^{t_K} (\mathbf{H}_k \mathbf{M}_k \delta\mathbf{x} - \mathbf{d}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{M}_k \delta\mathbf{x} - \mathbf{d}_k), \quad (1)$$

where  $\mathbf{B}$  is the background error covariance,  $t_k = t_0, \dots, t_K$  the data assimilation time window,  $\mathbf{d}_k = \mathbf{y}_k - \mathbf{H}_k(\mathbf{M}_k(\mathbf{x}_b))$

the innovations with  $\mathbf{y}_k$  being the vector of observations at time  $t_k$ ,  $\mathbf{x}_b$  the model background state valid at time  $t_0$ ,  $M_k(\cdot)$  denotes integration of the non-linear model from time  $t_0$  until time  $t_k$  and  $\mathbf{M}_k$  the corresponding tangent linear model integration.  $H_k(\cdot)$  is the non-linear observation operator,  $\mathbf{R}_k$  is the observation error covariance and  $\mathbf{H}_k$  the linearized observation operator, all valid at time  $t_k$ .

The background error covariance matrix  $\mathbf{B}$  has such a large dimension that the inverse  $\mathbf{B}^{-1}$  cannot be obtained directly by matrix inversion techniques. Therefore we introduce a pre-conditioning matrix  $\mathbf{U}$  such that  $\mathbf{B} = \mathbf{U}\mathbf{U}^T$ , and  $\delta\mathbf{x} = \mathbf{U}\boldsymbol{\chi}$  with  $\boldsymbol{\chi}$  the assimilation control variable. The transformation matrix  $\mathbf{U}$  is not formulated explicitly but given, for example, as a series of simpler transform operators. The cost function to be minimized and its gradient with respect to the assimilation control variable  $\boldsymbol{\chi}$  are given by

$$J = J_b + J_o = \frac{1}{2}\boldsymbol{\chi}^T\boldsymbol{\chi} + \frac{1}{2}\sum_{t_k=t_0}^{t_K} (\mathbf{H}_k\mathbf{M}_k\mathbf{U}\boldsymbol{\chi} - \mathbf{d}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k\mathbf{M}_k\mathbf{U}\boldsymbol{\chi} - \mathbf{d}_k) \quad (2)$$

and

$$\nabla_{\boldsymbol{\chi}} J = \boldsymbol{\chi} + \sum_{t_k=t_0}^{t_K} \mathbf{U}^T \mathbf{M}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{H}_k\mathbf{M}_k\mathbf{U}\boldsymbol{\chi} - \mathbf{d}_k). \quad (3)$$

The gradient calculation (3) is applied for every iteration during the 4D-Var minimization. A forward integration of the tangent linear model  $\mathbf{M}_k$  and a backward integration of the adjoint model  $\mathbf{M}_k^T$  are required for each such iteration.

Following Liu et al. (2008) we replace the static error covariance  $\mathbf{B}$  with a flow-dependent error covariance  $\mathbf{B} \approx \mathbf{X}'_b (\mathbf{X}'_b)^T$  estimated from an ensemble of background model states.  $\mathbf{X}'_b$  is a matrix whose columns are the normalized deviations of the ensemble background states from their mean:

$$\mathbf{X}'_b = \frac{1}{\sqrt{N-1}} (\mathbf{x}_{b1} - \bar{\mathbf{x}}_b, \dots, \mathbf{x}_{bN} - \bar{\mathbf{x}}_b), \quad (4)$$

where  $N$  is the number of ensemble members.

Also following Liu et al. (2008) we may apply  $\mathbf{X}'_b$  for the pre-conditioning  $\delta\mathbf{x} = \mathbf{X}'_b\boldsymbol{\chi}$ . Here the control vector will have the dimension of the number of ensemble members and we may notice that the assimilation increment is just a linear combination of the ensemble perturbations. The cost function and its gradient will have a similar form as above ( $\mathbf{X}'_b$  will replace  $\mathbf{U}$ ). We have achieved one of our aims, to utilize a flow-dependent background error covariance at the start of the 4D-Var assimilation window, but we still need to integrate the tangent linear model forward in time and the adjoint model backward in time over the data assimilation window. To avoid this, Liu et al. (2008) used an ensemble of non-linear model integrations over the data assimilation window to replace the forward integration of the tangent linear

model with pre-calculated values derived from these non-linear model trajectories in observation space.

The background error covariance of this 4D-En-Var formulation has the maximum rank equal to the number of ensemble members minus 1 ( $N - 1$ ) and cannot directly be applied to a full-scale NWP data assimilation problem. For demonstration purposes Liu et al. (2008) applied the method to a low-dimensional shallow-water model, with the degrees of freedom of the same order as the number of ensemble members. They showed that the performance of this 4D-En-Var for the low-order problem was very similar to the performance of 4D-Var and other assimilation algorithms like EnKF and 3D-Var. An even more important message is that the computational cost of 4D-En-Var was much lower than the corresponding cost of 4D-Var, since the integrations of the TL and AD models over the assimilation window were avoided.

In a second study Liu et al. (2009) applied 4D-En-Var to observing system simulation experiments (OSSE) with the WRF (weather research and forecasting) model (Skamarock et al., 2008). The number of ensemble members is in this case several orders of magnitude smaller than the model dimension, leading to a severe rank deficiency of the ensemble-based background error covariance matrix. The standard techniques in EnKF to counteract this rank deficiency are to apply covariance localization and/or to apply the EnKF algorithm locally, for example by solving separate filtering problems for each grid point of the model domain. The application of a local EnKF brings us back to the need for local data selection and will not be further discussed here. The most common algorithm for covariance localization is to apply an element-by-element multiplication (Schur product) of the original low-rank ensemble-based covariance matrix  $\mathbf{B}_{\text{ens}} = \mathbf{X}'_b (\mathbf{X}'_b)^T$  with a localization correlation matrix  $\mathbf{C}$ , constructed such that the final covariances  $\mathbf{B} = \mathbf{C} \circ \mathbf{B}_{\text{ens}}$  applied in the EnKF will be eliminated at long distances (Houtekamer and Mitchell, 2001). Liu et al. (2009) applied the correspondence to this covariance localization within the 4D-En-Var formulation with a technique similar to the one described by Buehner (2005). The pre-conditioning was done with the matrix  $\mathbf{P}'_b$  given by

$$\mathbf{P}'_b = (\mathbf{C}' \circ \mathbf{X}'_{b1}, \mathbf{C}' \circ \mathbf{X}'_{b2}, \dots, \mathbf{C}' \circ \mathbf{X}'_{bN}), \quad (5)$$

where  $\mathbf{C}'\mathbf{C}'^T = \mathbf{C}$ ,  $\mathbf{C}'$  is an  $n \times r$  matrix with  $r \leq n$  and  $\mathbf{X}'_{bl}$  is an  $r$  column matrix with every column being equal to the  $l$ th column in  $\mathbf{X}'_b$ .  $N$  is the number of ensemble members and  $n$  is the model space dimension. The columns of the transformation matrix  $\mathbf{C}'$  consist of eigenvectors of the correlation matrix  $\mathbf{C}$  multiplied by the square root of the eigenvalues. In order to make the eigenvector decomposition possible, the eigenvectors are evaluated on a coarser resolution grid, are possibly truncated and are interpolated to the assimilation grid. It is straightforward to show that  $\mathbf{P}'_b (\mathbf{P}'_b)^T = \mathbf{C} \circ \mathbf{B}_{\text{ens}}$ .

Liu and Xiao (2013) applied 4D-En-Var including covariance localization developed for WRF to a full-scale data assimilation experiment over Antarctica. They managed to demonstrate overall superiority and robust performance in comparison with 3D-Var and 3D-Var FGAT (first guess at appropriate time).

Buehner et al. (2010a, b) implemented a 4D-En-Var similar to the 4D-En-Var of Liu et al. (2009) into the Canadian operational global NWP system. The covariance localization was the same as in Buehner (2005). 4D-En-Var was compared with 4D-Var and 3D-Var FGAT, both applied with static background error covariances as well as with ensemble-based background error covariances, and was also compared with EnKF (Houtekamer and Mitchell, 2001). The ensemble-based error covariances were all based on the ensemble of short-range forecasts produced by the EnKF. Comparisons were made for single simulated observation impact experiments (Buehner et al., 2010a) and for full-scale data assimilation and forecasts over a 1 month period (Buehner et al., 2010b). The 4D-En-Var performed as well as 4D-Var with a static background error covariance, but slightly worse than 4D-Var Hybrid with an ensemble-based background error covariance in addition.

Desroziers et al. (2014) compared different 4D-En-Var formulations, including pre-conditioning based on the  $\mathbf{B}$  matrix, the full background error covariance, rather than  $\sqrt{\mathbf{B}}$  matrix as suggested by Liu et al. (2009) and applied in the present work with the HIRLAM 4D-En-Var. It is argued by Desroziers et al. (2014) that this alternative pre-conditioning might provide larger flexibility to implement the localization needed in 4D-En-Var. The choice of pre-conditioning method can also have an influence on the dimension of the assimilation control variable and on the characteristics of the minimization problem.

### 3 Formulation of the HIRLAM 4D-En-Var Hybrid

The 4D-En-Var for the High Resolution Limited Area Model (HIRLAM, Undén et al., 2002) is an extension of the HIRLAM Hybrid variational ensemble data assimilation scheme (Gustafsson et al., 2014). It builds on the idea of an augmentation of the control vector by Lorenc (2003) and use of the  $\mathbf{B}^{\frac{1}{2}}$  for pre-conditioning. Furthermore, the HIRLAM 4D-En-Var is a hybrid between 4D-En-Var and 3D-Var FGAT (first guess at appropriate time), such that the assimilation increment  $\delta\mathbf{x}(t_k)$  at time  $t_k$  within the assimilation time window  $t_0 \leq t_k \leq t_K$  is formed as a linear combination of a 3D-Var FGAT increment  $\delta\mathbf{x}_{3\text{dvar}}(\frac{t_0+t_K}{2})$  and a 4D-En-Var increment  $\delta\mathbf{x}_{\text{ens}}(t_k)$ :

$$\delta\mathbf{x}(t_k) = \delta\mathbf{x}_{3\text{dvar}}\left(\frac{t_0+t_K}{2}\right) + \delta\mathbf{x}_{\text{ens}}(t_k). \quad (6)$$

According to HIRLAM 3D-Var FGAT experiences, in order to utilize time-averaging effects, the assimilation increment

related to the static part of the background error covariance should preferably be applied for non-linear model initialization at time  $\frac{t_0+t_K}{2}$  in the middle of the data assimilation window.

Pre-conditioning for the two components of the assimilation increment is done separately. The pre-conditioning for the 3D-Var FGAT component  $\delta\mathbf{x}_{3\text{dvar}} = \mathbf{U}\boldsymbol{\chi}_{3\text{dvar}}$  with  $\mathbf{B}_{3\text{dvar}} = \mathbf{U}\mathbf{U}^T$  is done as in the HIRLAM 3D-Var (Berre, 2000; Gustafsson et al., 2001). The transform operator  $\mathbf{U}$  includes vertical transforms of the spectral control vector  $\boldsymbol{\chi}_{3\text{dvar}}$  from vertical eigenvector space to model levels, inversion of vertical balance operators and projection by Fourier transforms from spectral space to grid point space. An area extension is applied to the regional model domain in order to make application of fast Fourier transform possible (Haugen and Machenhauer, 1993). For the 4D-En-Var component  $\delta\mathbf{x}_{\text{ens}}$  we follow Liu et al. (2009), see also Eq. (5):

$$\begin{aligned} \delta\mathbf{x}_{\text{ens}} &= \mathbf{P}'_b \boldsymbol{\chi}_{\text{ens}} = (\mathbf{C}' \circ \mathbf{X}'_{b1}, \mathbf{C}' \circ \mathbf{X}'_{b2}, \dots, \mathbf{C}' \circ \mathbf{X}'_{bN}) \boldsymbol{\chi}_{\text{ens}} \\ &= \sum_{l=1}^N \mathbf{C}' \circ \mathbf{X}'_{bl} (\boldsymbol{\chi}_{\text{ens}})_l, \end{aligned}$$

where  $\boldsymbol{\chi}_{\text{ens}}$  is the control vector of dimension  $N \cdot r$  and  $(\boldsymbol{\chi}_{\text{ens}})_l$  the  $l$ th component of the control vector, corresponding to ensemble member  $l$ . If we introduce a vector  $\boldsymbol{\alpha}_l = \mathbf{C}' (\boldsymbol{\chi}_{\text{ens}})_l$  we can show by substitution into component form that

$$\delta\mathbf{x}_{\text{ens}} = \sum_{l=1}^N \boldsymbol{\alpha}_l \circ (\mathbf{X}'_b)_l \quad (7)$$

and this is 4D-En-Var in the form of the control vector augmentation method (the  $\alpha$ -method) suggested by Lorenc (2003). The  $\boldsymbol{\alpha}_l$  can be considered as a localized weight field for the ensemble perturbations of ensemble member  $l$ . In the most general case, and without any truncation of the localization correlation function,  $\boldsymbol{\alpha}_l$  will be a vector of dimension  $K \times n$ . In the first version of HIRLAM 4D-En-Var Hybrid, we will assume the same localized ensemble weights for each observation time window, corresponding to a strong constraint 4D-Var, and also for each model level, corresponding to application of horizontal localization only. The impact of an additional vertical localization is studied separately, see Sect. 5.2.3. In the HIRLAM 4D-En-Var Hybrid the transform matrix  $\mathbf{C}'$  is derived from the localization correlation matrix  $\mathbf{C}$  in a similar way as for the static background error covariance matrix. The localization correlation matrix  $\mathbf{C}$  is thus generated in the extended domain with bi-periodic variations, without any loss of generality we may assume homogeneity and isotropy with regard to the horizontal localization correlation function and then we may obtain a correlation spectrum representing the localization correlation in spectral space. Its square root, the transform matrix  $\mathbf{C}'$ , consists of horizontal inverse fast Fourier transforms to grid point space. The transforms are carried out to obtain the

weights  $\alpha_l$  for the different ensemble members. By construction these weights include the implicit localization of ensemble perturbations  $(X'_b)_l$ . The same procedure is also applied in the HIRLAM hybrid variational ensemble assimilation algorithm and is discussed in Gustafsson et al. (2014).

For the localization in the horizontal direction we apply the power law correlation function with a length scale  $c$  of 500 km:

$$\rho(c, r) = \frac{1}{1 + \frac{1}{2}(\frac{r}{c})^2}, \quad (8)$$

where  $r$  is the distance. In order to limit the negative localization effects on model balances, we perform localization on vorticity, divergence, temperature, specific humidity and surface pressure perturbations. A similar localization of stream function and velocity potential was discussed and applied by Keperth (2009).

The cost function expression for the HIRLAM 4D-En-Var Hybrid is given by

$$J(\chi_{3dvar}, \chi_{ens}) = \beta_{3dvar} J_{b,3dvar}(\chi_{3dvar}) + \beta_{ens} J_{b,ens}(\chi_{ens}) + J_o, \quad (9)$$

where the weights  $\beta_{3dvar}$  and  $\beta_{ens}$  for the two different components of the background error constraint should fulfil

$$\frac{1}{\beta_{3dvar}} + \frac{1}{\beta_{ens}} = 1. \quad (10)$$

Slightly different from the 4D-En-Var implementation by Liu et al. (2009) we apply explicitly the linearized observation operator  $\mathbf{H}_k$  in the observation constraint part  $J_o$  of the cost function to be minimized:

$$J_o = \frac{1}{2} \sum_{t_k=t_0}^{t_K} (\mathbf{H}_k \delta \mathbf{x}(t_k) - \mathbf{d}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \delta \mathbf{x}(t_k) - \mathbf{d}_k), \quad (11)$$

where  $\delta \mathbf{x}(t_k)$  given by Eq. (6) can be expressed in the assimilation control variables as

$$\begin{aligned} \delta \mathbf{x}(t_k) &= \mathbf{U} \chi_{3dvar} + \kappa \sum_{l=1}^N \alpha_l \circ (\mathbf{X}'_b)_l(t_k) \\ &= \mathbf{U} \chi_{3dvar} + \kappa \sum_{l=1}^N (C'(\chi_{ens})_l) \circ (\mathbf{X}'_b)_l(t_k), \end{aligned} \quad (12)$$

where  $\kappa$  is a tuning factor.

There are several parameters of the HIRLAM 4D-En-Var Hybrid that can be considered as tuning parameters – for example, the form of the horizontal localization correlation function and its length scale, the weights  $\beta_{3dvar}$ ,  $\beta_{ens}$  associated with the two different components of the background error constraint and the tuning factor  $\kappa$  to scale the ensemble perturbations for estimation of the flow-dependent forecast error covariance. Here we will make experiments with  $\beta_{3dvar} = \beta_{ens} = 2$ , thus with equal weights given to the two

components of the background error constraint, and with almost full weight given to the ensemble-based background error constraint. We will also compare experiments without any scaling of the ensemble perturbations ( $\kappa = 1$ ) with experiments with such a scaling ( $\kappa = 4$ ). For sensitivity of the 4D-Var Hybrid scheme to these tuning parameters, see Gustafsson et al. (2014).

#### 4 ETKF re-scaling

The 4D-En-Var Hybrid as implemented in HIRLAM provides an initial state for the ensemble control only. For the generation of the perturbed ensemble members we apply an ensemble transform KF (ETKF) based re-scaling scheme (Bishop et al., 2001). Essentially this is a low-rank estimation of the analysis error variances, preserving also dynamical structures. Further details on the HIRLAM ETKF re-scaling are described and discussed in Bojarova et al. (2010) and Gustafsson et al. (2014).

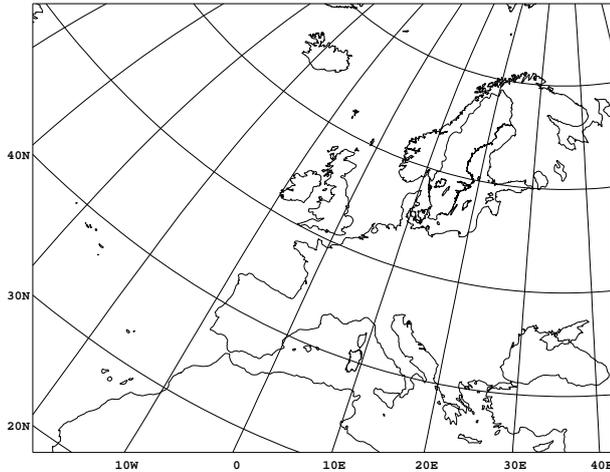
#### 5 Results

Data assimilation experiments to validate the performance of the HIRLAM 4D-En-Var Hybrid and to compare it with HIRLAM 4D-Var, as well as with the HIRLAM 4D-Var Hybrid data assimilation, were done over the period 17 January–28 February 2008. Single simulated observation impact experiments were carried out in order to illustrate the flow-dependent assimilation increments due to the use of ensemble-based background error covariances. The average effects on forecast quality are illustrated through a data assimilation and forecast experiment, using real observations, over the whole data period (43 days). The model domain applied during the experiments is shown in Fig. 1; the horizontal grid resolution was 11 km and the number of vertical levels of the forecast model was 40.

The forecast model used in the experiments was the HIRLAM grid point forecast model. It is hydrostatic and it utilizes a semi-implicit, semi-Lagrangian two time level integration scheme (Undén et al., 2002). The physical parameterizations were the CBR turbulence scheme (Cuxart et al., 2000), the Kain–Fritsch convection scheme (Kain, 2004), the Rasch–Kristjánsson cloud water scheme (Rasch and Kristjánsson, 1998), the Savijärvi (1990) radiation and the ISBA surface scheme (Noilhan and Mahfouf, 1996).

##### 5.1 Single observation experiments with 4D-Var, 4D-Var Hybrid, and 4D-En-Var Hybrid

To illustrate the influence of ensemble based error covariances on flow dependency, we selected a case with a strong baroclinic development over the Northern Atlantic south of Iceland on 7 February 2008. The 6 h data assimilation window selected was from 6 February 2008, 21:00 UTC,



**Figure 1.** Model domain for the data assimilation and forecast experiments.

**Table 1.** Description of experiments carried out with the HIRLAM 4D-Var, the HIRLAM Hybrid 4D-Var ensemble data assimilation and with the HIRLAM 4D-En-Var for the winter period 17 January–28 February 2008.

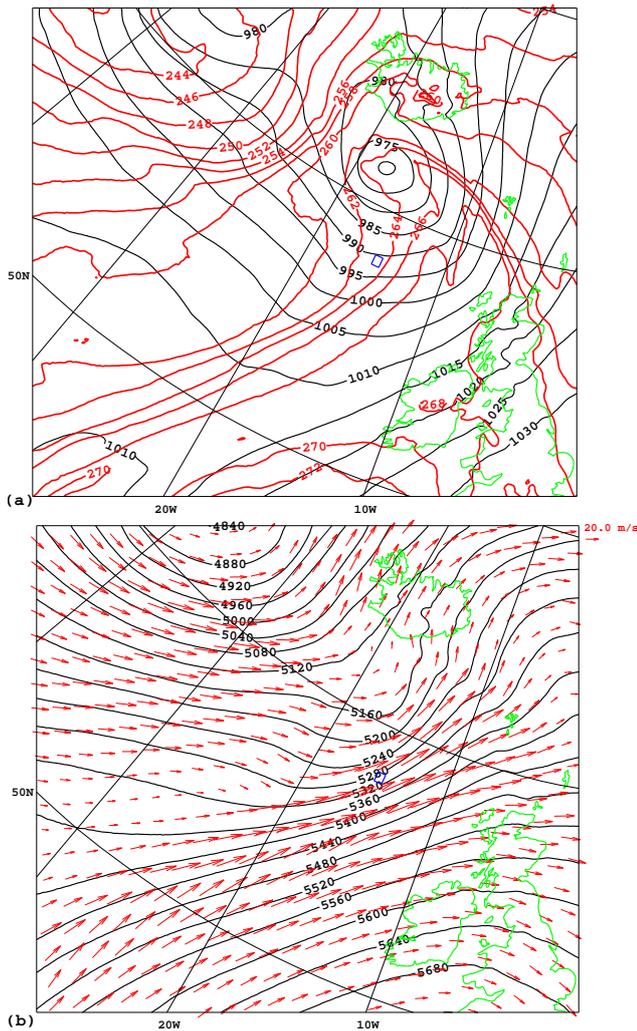
Experiment name	Description
4DVar	4D-Var, no ensemble constraint, two outer loops, 66 and 44 km increment resolution.
4DVar-Hyb25-Sc	4D-Var Hybrid, $\beta_{\text{var}} = 4$ , $\beta_{\text{ens}} = 4/3$ (25 % ensemble, 75 % static cov.), two outer loops, 66 and 44 km increment resolution. Ensemble perturbations scaled by a factor 4 in the covariance estimation.
4DEnVar-Hyb50-Sc	4D-En-Var, $\beta_{\text{var}} = \beta_{\text{ens}} = 2$ (50 % ensemble, 50 % 3D-Var FGAT cov.), one outer loop, 33 km increment resolution. Ensemble perturbations scaled by a factor 4 in the covariance estimation.
4DEnVar-Hyb50-NoSc	As 4DEnVar-Hyb50-Sc but without any scaling of the ensemble perturbations in the covariance calculation.
4DEnVar-Hyb90-NoSc	As 4DEnVar-Hyb50-NoSc with the exception that $\beta_{\text{var}} = 11$ , $\beta_{\text{ens}} = 11/10$ ( $\approx 90$ % ensemble, $\approx 10$ % 3D-Var FGAT cov.).
VertLoc	As 4DEnVar-Hyb50-Sc but with a vertical localization of ensemble-based covariances added

until 7 February, 03:00 UTC. Analysis fields for 7 February 2008, 03:00 UTC, taken from a 4D-En-Var Hybrid assimilation experiment (experiment name 4DEnVar-Hyb50-Sc, see Table 1), are shown for mean sea level pressure and 700 hPa temperature in Fig. 2a and for 500 hPa geopotential height and wind in Fig. 2b. Note the strong small-scale surface pressure development south of Iceland, the sharp frontal zones in the 700 hPa temperature field and the associated small-scale trough in the 500 hPa geopotential height field.

To investigate how the standard 4D-Var develops implicit flow-dependent assimilation structure functions, a simulated wind observation was inserted into the assimilation process at 7 February 2008, 02:00 UTC, 5 h into the assimilation window. The horizontal position of the simulated observation was  $58^\circ\text{N}$ ,  $15^\circ\text{W}$  and the vertical level was 500 hPa. The corresponding observation increment was approximately southwesterly  $12\text{ m s}^{-1}$  ( $10\text{ m s}^{-1}$  in the model geometry  $u$  component and  $5\text{ m s}^{-1}$  in the  $v$  component).

500 hPa geopotential height and wind assimilation increments induced by the single simulated observation on 7 February 2008, 02:00 UTC, are presented in Fig. 3 for the following three assimilation configurations: (1) standard 4D-Var with a single outer loop; (2) hybrid 4D-Var ensemble assimilation with a single outer loop, with equal weights given to the static background error covariance and the ensemble-based background error covariance and with the tuning factor  $\kappa = 4.0$ ; (3) 4D-En-Var Hybrid with a single outer loop, with equal weights given to the 3D-Var FGAT and the four-dimensional ensemble parts of the background error covariance and with the tuning factor  $\kappa = 4.0$  as in the 4D-Var Hybrid. All experiments applied a horizontal grid resolution of 33 km for the assimilation increments. The 20 ensemble members from a 4D-En-Var Hybrid experiment (experiment name 4DEnVar-Hyb50-Sc) were utilized in experiments (2) and (3). It should also be emphasized that the HIRLAM implementation of the 4D-Var, 4D-Var Hybrid, and 4D-En-Var Hybrid algorithms are completely consistent and based on the same computer code.

The single simulated observation impact assimilation increments clearly illustrate the utilization of flow-dependent structure functions in all three experiments (4D-Var, 4D-Var Hybrid, and 4D-En-Var Hybrid), in particular with regard to the sharpening of the small-scale trough in the 500 hPa geopotential height and wind fields, needed in order satisfy the simulated single observation. The flow dependency is more accentuated for the 4D-Var Hybrid and 4D-En-Var Hybrid experiments than for the standard 4D-Var experiment. This result is consistent with application of a static background error covariance at the start of the assimilation window in the standard 4D-Var experiment. In contrast, there are contributions from an ensemble of forecast perturbations to this background error covariance in the 4D-Var Hybrid and 4D-En-Var Hybrid experiments. Furthermore, in the 4D-En-Var Hybrid experiment there is an additional contribution from the ensemble of background error perturbations valid



**Figure 2.** Mean sea level pressure and 700 hPa temperature (a), 500 hPa geopotential height and wind (b), 7 February 2008, 03:00 UTC. Analyses fields are from one of the 4D-En-Var Hybrid experiments (experiment name 4DENVar-Hyb50-Sc). The position of the single simulated observation at 58° N, 15° W is indicated by a blue square. Isoline spacing is 5 hPa, 2°K and 40 m for mean sea level pressure, 700 hPa temperature and 500 hPa geopotential height, respectively. A reference wind arrow is provided to the right of the 500 hPa map.

at the time of the simulated observation (7 February 2008, 02:00 UTC). It is satisfactory to notice that the assimilation increments from 4D-Var Hybrid and the 4D-En-Var Hybrid are quite similar, taking into account that the derivations of the two sets of increments are quite different. A tangent linear (TL) model integration of the increment over 5 h from the start of the assimilation window is applied in 4D-Var Hybrid, while in the case of the 4D-En-Var Hybrid there is no such TL model involved. The time evolution of the increments is provided by the four-dimensional background error

covariance estimated from an ensemble on non-linear model integrations over the assimilation window.

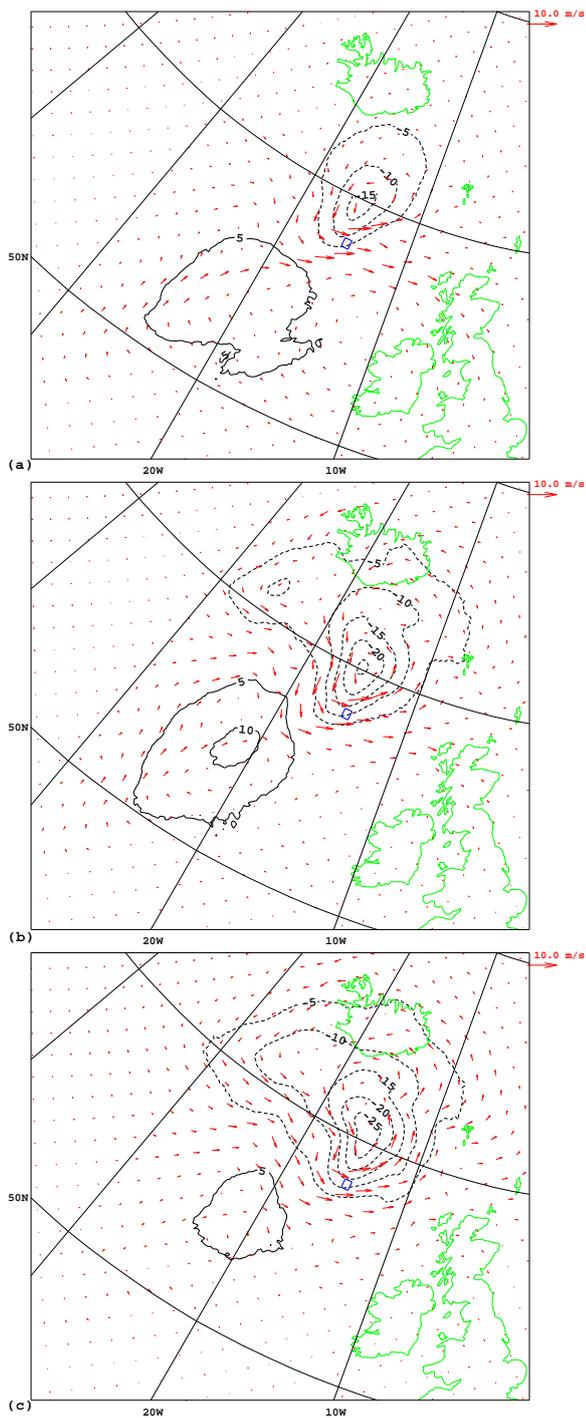
In order to illustrate the effects on cross-correlations between different model variables, assimilation increments of surface pressure and 700 hPa temperature due to the single 500 hPa simulated wind observation are presented in Fig. 4. First of all we may notice that the assimilation increments are more pronounced (amplified) in 4D-En-Var Hybrid experiment. The 4D-Var increments are much weaker and the 4D-Var Hybrid increments are somewhere in between the 4D-Var and the 4D-En-Var Hybrid increments. The shape of the surface pressure increments indicates that the low-pressure system is forced to move faster when the simulated 500 hPa wind is assimilated and, in particular in the case of the 4D-En-Var Hybrid, there is also a stronger deepening of the low. From the 700 hPa temperature assimilation increment fields we can notice a consistent faster propagation of the atmospheric frontal zones.

## 5.2 Real observation assimilation experiments

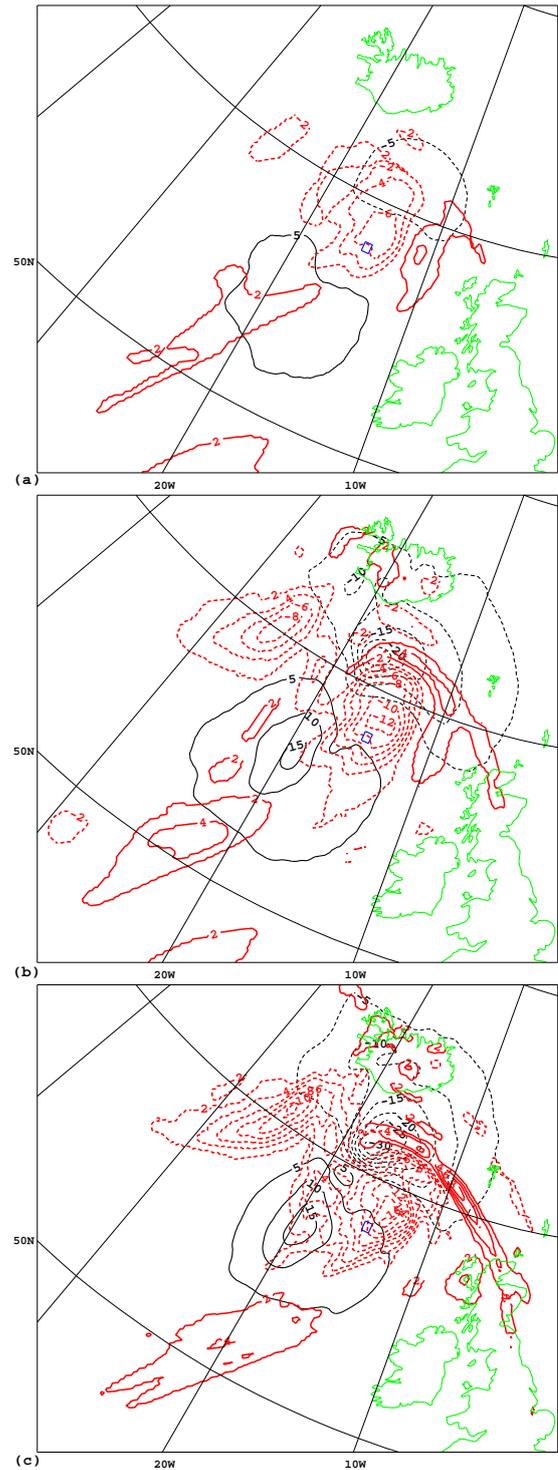
Real observation data assimilation and forecast experiments were carried out for the period 17 January–28 February 2008 and for the six data assimilation configurations described in Table 1. The same conventional observations (TEMP, PILOT, SHIP, AIREP, DRIBU, SYNOP) and satellite radiance data (AMSU-A) were used in all experiments. For each experiment, forecasts up to +30 h were produced and verified against radiosonde and surface observations. The first three days of the experiment period were excluded from the verification in order to allow a spin-up of the LAM ensemble perturbations. We will first describe and discuss forecast verification scores based on the three data assimilation schemes 4D-Var, 4D-Var Hybrid and 4D-En-Var Hybrid. Secondly, we will describe, compare and discuss forecast verification scores for three different versions of 4D-En-Var Hybrid with the aim of checking the sensitivity to two tuning parameters of the HIRLAM 4D-En-Var Hybrid. These tuning parameters are the magnitude of the scaling of the ensemble perturbations for estimation of the flow-dependent background error covariances (tuning factor  $\kappa$ ) and the relative weights given to the ensemble and the 3D-Var FGAT components of the background error covariance. Finally, we have also carried out a small impact study on the sensitivity of forecast verification scores to introduction of a very simple vertical localization of the ensemble-based background error covariance matrix in the HIRLAM 4D-En-Var Hybrid scheme.

### 5.2.1 Verification of forecasts based on 4D-Var, 4D-Var Hybrid, and 4D-En-Var Hybrid

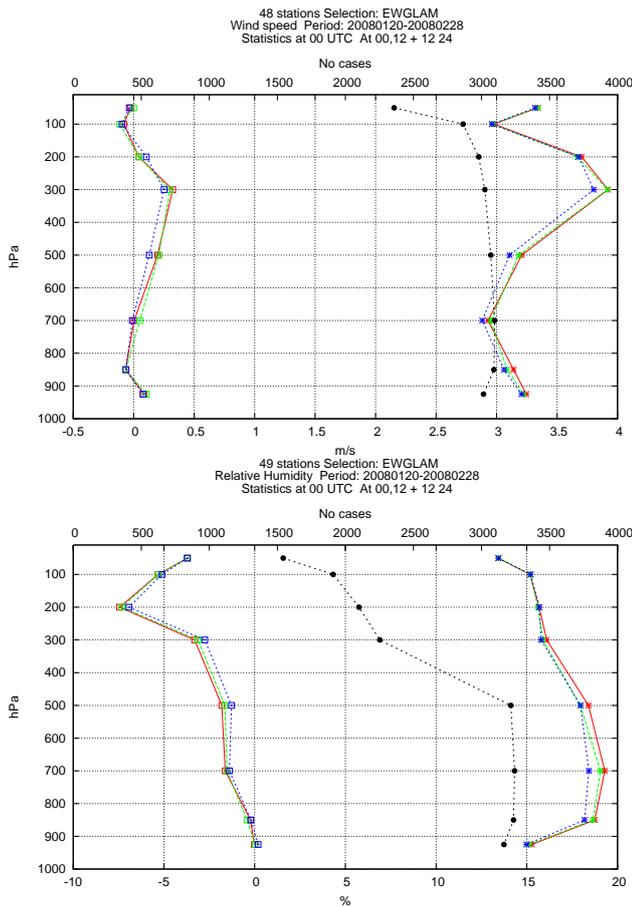
Standard deviation (STDV) and mean (bias) verification scores for forecasts of vertical profiles of +12 and +24 h wind speed and relative humidity valid at 00:00 UTC are presented in Fig. 5 for the 4D-Var, 4D-Var Hybrid and



**Figure 3.** 500 hPa geopotential height and wind analysis increments for 7 February 2008, 02:00 UTC, 5 h into the assimilation window between 6 February 2008, 21:00 UTC, and 7 February 2008, 03:00 UTC. 4D-Var (a), 4D-Var Hybrid (b), and 4D-En-Var Hybrid (c). Simulated single 500 Pa wind observation, with a  $10 \text{ m s}^{-1}$  observation increment at  $58^\circ \text{ N}$ ,  $15^\circ \text{ W}$  (marked with a blue square) in the model geometry  $u$  component and a  $5 \text{ m s}^{-1}$  observation increment in the  $v$  component. Isoline spacing is 5 m for the geopotential height and a reference wind error is provided to the right of the map.



**Figure 4.** Surface pressure and 700 hPa temperature analysis increments for 7 February 2008, 02:00 UTC, 5 h into the assimilation window between 6 February 2008, 21:00 UTC, and 7 February 2008, 03:00 UTC. 4D-Var (a), 4D-Var Hybrid (b), and 4D-En-Var Hybrid (c). Simulated single 500 Pa wind observation, with a  $10 \text{ m s}^{-1}$  observation increment at  $58^\circ \text{ N}$ ,  $15^\circ \text{ W}$  (marked with a blue square) in the model geometry  $u$  component and a  $5 \text{ m s}^{-1}$  observation increment in the  $v$  component. Isoline spacing is 5 hPa for the surface pressure and  $2^\circ \text{ K}$  for the 700 hPa temperature.



**Figure 5.** Verification of vertical profiles of wind speed and relative humidity forecasts against radiosonde data from stations in a European network. Standard deviation (STDV) and mean (bias) verification scores. Average verification scores for +12 and +24 h forecasts valid at 00:00 UTC over the period 20 January–28 February 2008; 4D-Var (experiment name 4DVar, red lines), 4D-Var Hybrid (experiment name 4DVar-Hyb25-Sc, green lines), and 4D-En-Var Hybrid (experiment name 4DENVar-Hyb50-Sc, blue lines). The black dotted line indicates the number of verification comparisons.

4D-En-Var Hybrid experiments. With regard to the standard deviation verification scores we can notice a small but consistent improvement of the scores for 4D-Var Hybrid in comparison with the standard 4D-Var, and a further larger improvement of the scores for 4D-En-Var Hybrid in comparison with 4D-Var Hybrid. These improvements are most significant for tropospheric winds and for relative humidity in the mid-troposphere (500–700 hPa). Vertical profiles of temperature forecasts were also verified against radiosonde data. The temperature verification scores show an almost neutral impact from the choice of assimilation method and for the current ensemble generation methods (results are not shown).

The improvements introduced by 4D-Var Hybrid in comparison with standard 4D-Var are certainly related to the use of an ensemble-based and flow-dependent component of the

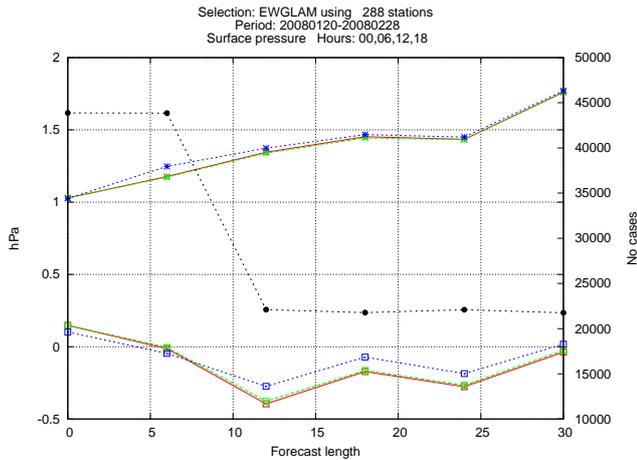
background error covariance at the start of the assimilation window – for a more detailed discussion see Gustafsson et al. (2014). The reasons for the clear improvement of forecast scores produced by 4D-En-Var Hybrid in comparison with standard 4D-Var and 4D-Var Hybrid are less obvious, and may be even a bit surprising. The main difference is the replacement of the tangent linear and adjoint models in 4D-Var and 4D-Var Hybrid with the use of a four-dimensional ensemble of non-linear model integrations (trajectories) over the data assimilation window. From a computing efficiency point of view, this result is highly satisfactory, since the computing time is significantly reduced when the integration of the TL and AD models is avoided (see below for details). The particularly improved forecast verification scores for relative humidity with 4D-En-Var Hybrid may possibly be explained by the too simplified TL and AD models for moist processes in HIRLAM 4D-Var (no condensation is included, for example).

Standard deviation (STDV) and mean (bias) verification scores for mean sea level pressure (MSLP) forecasts verified against surface observations are presented in Fig. 6 for the 4D-Var, 4D-Var Hybrid and 4D-En-Var Hybrid experiments. There are two features in these MSLP verification scores that need to be discussed and possibly also further investigated in order to arrive at plausible explanations. On one hand, there is a reduced bias in the MSLP forecasts based on 4D-En-Var Hybrid compared to the MSLP forecasts based on 4D-Var and 4D-Var Hybrid. On the other hand, the 4D-En-Var Hybrid forecasts are associated with larger STDV of the forecast errors for the shortest timescales +6 and +12 h, while the STDV of the scores are very similar for longer timescales, +18, +24 and +30 h.

Here we will only present an educated guess as to the behaviour of the verification scores for MSLP; further evidence from complementary investigations will be elaborated on below.

It is a general experience from the NWP practice that MSLP forecast biases are associated with physical inconsistencies causing, for example, adjustment processes in the form of mass fluxes between land and sea surfaces or, in the case of limited area models, mass fluxes over the lateral boundaries. One may thus speculate as to whether the time series of non-linear model perturbations, applied as assimilation basis functions in 4D-En-Var Hybrid, improves consistency with the non-linear model compared with assimilation increments derived from time integrations with a highly simplified tangent linear model.

There may be several reasons behind the degraded MSLP STDV verification scores for 4D-En-Var Hybrid for the shortest forecast timescales (+6 and +12 h). The HIRLAM 4D-Var and 4D-Var Hybrid assimilation algorithms include a weak digital filter constraint for control of high-frequency (gravity wave) oscillations. Such a constraint is not yet a part of the 4D-En-Var Hybrid algorithm with increased high frequency oscillations (noise) as a possible effect. This

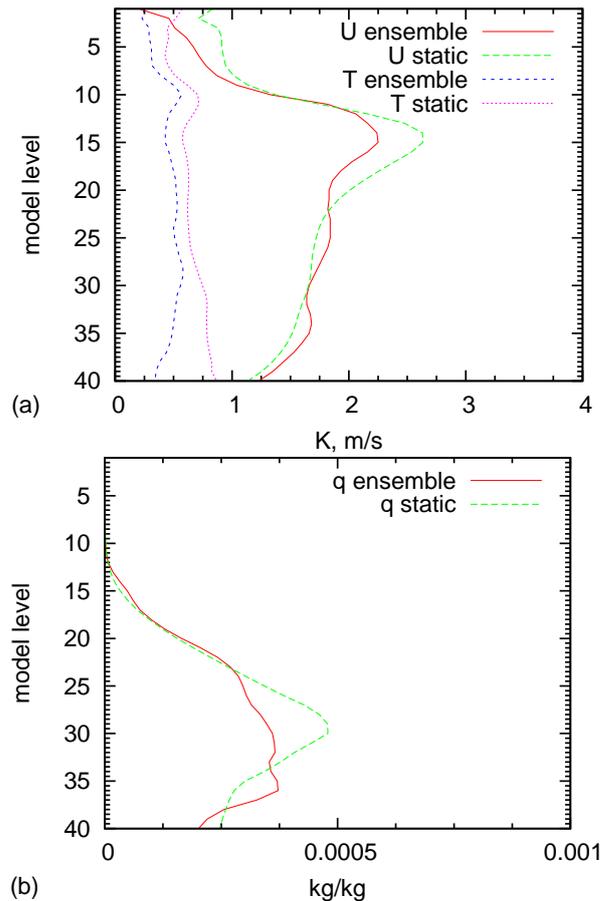


**Figure 6.** Verification of mean sea level pressure (MSLP) forecasts against surface (SYNOP) data from stations in a European network. Standard deviation (STDV) and mean (bias) verification scores. Average verification scores over the period 20 January–28 February 2008 as a function of forecast length. 4D-Var (experiment name 4DVar, red lines), 4D-Var Hybrid (experiment name 4DVar-Hyb25-Sc, green lines) and 4D-En-Var Hybrid (experiment name 4DEnVar-Hyb50-Sc, blue lines). The black dotted line indicates the number of verification comparisons.

possibility is further explored below. Furthermore, the rather primitive ETKF re-scaling scheme may also provide noisy surface pressure perturbations that may degrade their usefulness for assimilation purposes. Another explanation to the degraded MSLP STDV verification scores for the shortest timescales may be that the relatively small ensemble of background states (20 members) does not provide enough variability to fit the shortest timescale surface pressure variations equally well as with the tangent linear model in 4D-Var. Investigations of the fit between hourly surface pressure observations and the assimilation model states in the three different experiments seem to give some evidence for this hypothesis (see below).

### 5.2.2 Sensitivity of forecast verification scores to tuning parameters of the 4D-En-Var Hybrid scheme

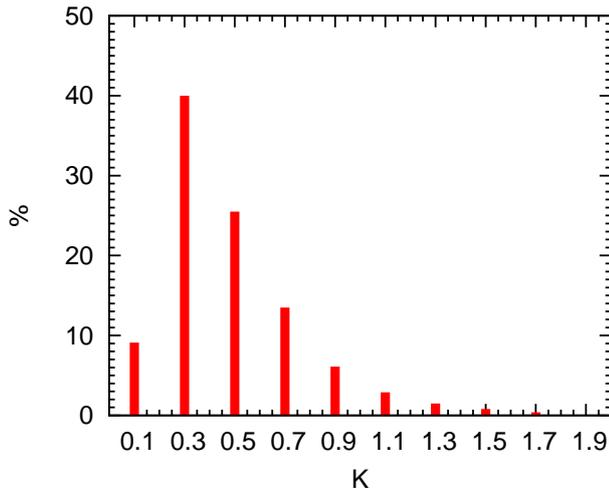
The HIRLAM 4D-En-Var Hybrid scheme includes several tuning parameters. Two of the most important ones are (1) the relative weights given to the ensemble and the static 3D-Var FGAT components of background error covariance (parameters  $\beta_{\text{ens}}$  and  $\beta_{\text{var}}$ ) and (2) the scaling factor  $\kappa$  applied to the ensemble perturbations that are used to estimate the flow-dependent (ensemble) component of the background error covariance. Data assimilation and forecast experiments were carried out over the time period 17 January–28 February 2008 to test the sensitivity of the forecast verification scores to different values of these 4D-En-Var Hybrid tuning parameters. It needs to be mentioned that the two tuning parameters are not completely unrelated. A scaling of the en-



**Figure 7.** Vertical profiles of climatological (static) and area-averaged ensemble-based background error standard deviations for one randomly selected case, 22 January 2008, 12:00 UTC + 6 h. (a)  $U$  component of wind (ensemble red curve, static green curve) and temperature (ensemble blue curve, static pink curve). (b) Specific humidity (ensemble red curve, static green curve).

semble perturbations is efficiently equivalent to changing the weight given to the ensemble part of the background error covariance while keeping the weight for the static part unchanged (neglecting Eq. 10).

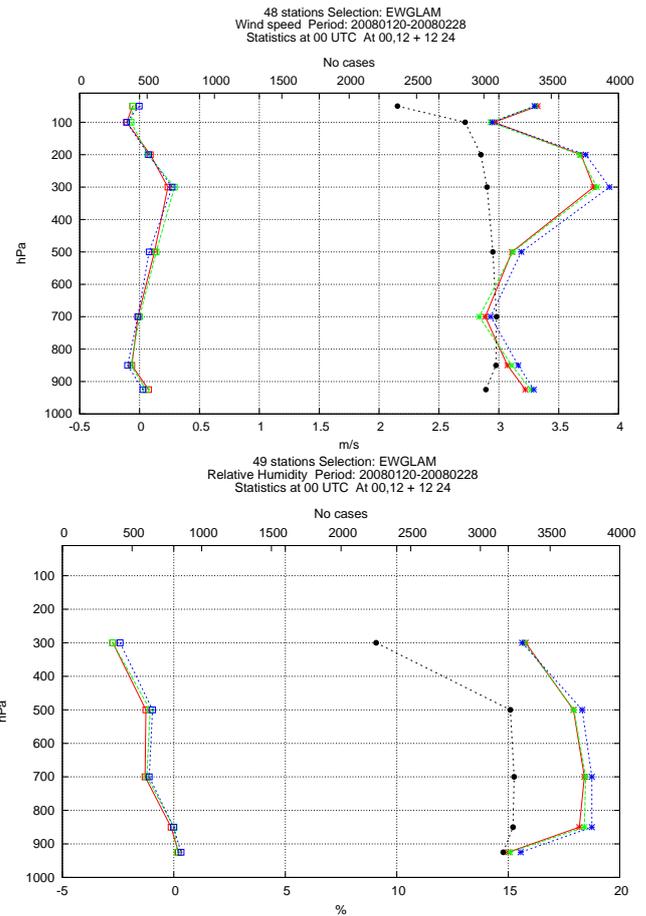
In order to understand better the results of these tuning experiments, it is necessary to examine the relative magnitudes of the contributing components of the background error covariance in the 4D-En-Var Hybrid experiments. Figure 7 shows vertical profiles of climatological (static) background error standard deviations and ensemble-based background error standard deviations, obtained by horizontal averaging of variances for one randomly selected case, 22 January 2008, 12:00 UTC, +6 h. It appears that the vertical variations and the magnitudes of these background error standard deviations are quite similar. It is encouraging that the simple tuning of the inflation factor in the ETKF re-scaling scheme, based on innovations (Bojarova et al., 2010), provides similar results as a more elaborate tuning of the climatological



**Figure 8.** Histogram of ensemble-based background error standard deviations for temperature on model level 30 ( $\approx 800$  hPa) for one randomly selected case, 22 January 2008, 12:00 UTC + 6 h.

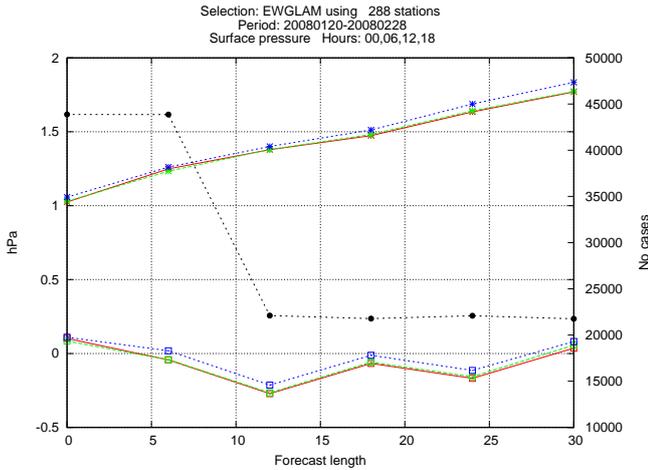
background error standard deviations (also based on innovations). One must also keep in mind that the ensemble based background error standard deviations in Fig. 7 represent horizontal averages – these standard deviations have significant spatial variations (Gustafsson et al., 2014). To further illustrate this, we present a histogram of individual grid point background error standard deviations for temperature at model level 30 valid on 22 January 2008, 12:00 UTC, +6 h in Fig. 8. The standard deviation calculated by horizontal averaging is 0.56 K. The frequency is highest for the interval 0.2–0.4 K, but contributions from much larger standard deviations in fewer grid points are significant. This result is consistent with the experience that large forecast uncertainty is associated with quite narrow zones of dynamical activity, in the present case frontal zones in the low tropospheric temperature field (see Gustafsson et al., 2014).

Figure 9 illustrates the sensitivity of verification scores of vertical forecast profiles of wind speed and relative humidity to the tuning of the scaling of the ensemble perturbations by comparing the verification results of experiment 4D-En-Var-Hyb50-NoSc (no scaling) with the verification results of experiment 4D-En-Var-Hyb50-Sc (scaling with a factor 4). The results indicate that the forecast quality is more or less insensitive to the scaling of ensemble perturbations used to estimate the forecast error covariance, see also the verification scores for MSLP in Fig. 10. One may argue that the HIRLAM ETKF rescaling scheme includes a multiplicative inflation scheme and that the background forecast perturbations are representative of background errors (Bojarova et al., 2010). Any further scaling is not needed. Furthermore, it was shown in experiments with the HIRLAM 4D-Var Hybrid assimilation (Gustafsson et al., 2014) that the *richness* in spatial structures of the ensemble is more important than the scaling of the amplitude of the ensemble perturbations.



**Figure 9.** Verification of vertical profiles of wind speed and relative humidity forecasts against radiosonde data from stations in a European network. Standard deviation (STDV) and mean (bias) verification scores. Average verification scores for +12 and +24 h forecasts valid at 00:00 UTC over the period 20 January–28 February 2008. 4D-En-Var Hybrid with scaling of ensemble perturbations and with equal weights given to the ensemble and 3D-Var FGAT components of the background error covariance (experiment name 4D-En-Var-Hyb50-Sc, red lines), 4D-En-Var Hybrid without scaling of ensemble perturbations and with equal weights given to the ensemble and 3D-Var FGAT components of the background error covariance (experiment name 4D-En-Var-Hyb50-NoSc, green lines). 4D-En-Var Hybrid without scaling of ensemble perturbations and with  $\approx 90\%$  weight given to the ensemble component of background error covariance (experiment name 4D-En-Var-Hyb90-NoSc, blue lines). The black dotted line indicates the number of verification comparisons.

Figures 9 and 10 also illustrate the sensitivity of the verification scores to the relative weighing between the ensemble and the 3D-Var FGAT contributions to the background error covariance by comparing the verification results of experiment 4D-En-Var-Hyb50-NoSc (equal contributions of the two components) and experiment 4D-En-Var-Hyb90-NoSc ( $\approx 90\%$  contributions of the ensemble component and  $\approx 10\%$  of the static 3D-Var FGAT component). The results show



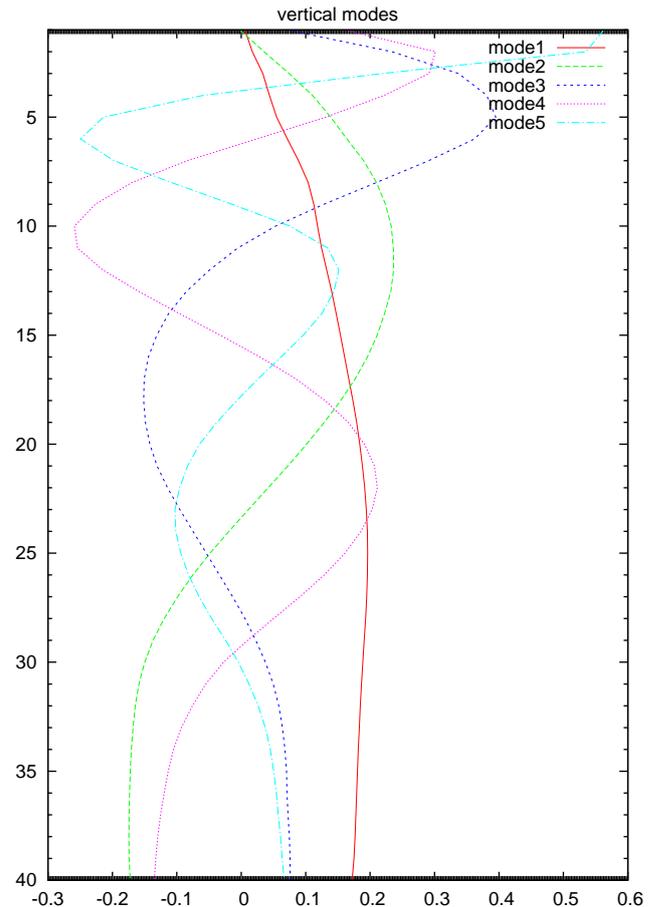
**Figure 10.** Verification of mean sea level pressure (MSLP) forecasts against surface (SYNOP) data from stations in a European network. Standard deviation (STDV) and mean (bias) verification scores. Average verification scores over the period 20 January–28 February 2008 as a function of forecast length. 4D-En-Var Hybrid with scaling of ensemble perturbations and with equal weights given to the ensemble and 3D-Var FGAT components of the background error covariance (experiment name 4D-En-Var-Hyb50-Sc, red lines), 4D-En-Var Hybrid without scaling of ensemble perturbations and with equal weights given to the ensemble and 3D-Var FGAT components of the background error covariance (experiment name 4D-En-Var-Hyb50-NoSc, green lines). 4D-En-Var Hybrid without scaling of ensemble perturbations and with  $\approx 90\%$  weight given to the ensemble component of background error covariance (experiment name 4D-En-Var-Hyb90-NoSc, blue lines). The black dotted line indicates the number of verification comparisons.

that the forecast quality, as given by the forecast verification scores, is consistently degraded when the static component of the background error covariance is reduced to 10% only. Our interpretation is that an ensemble of background model states with 20 members only does not provide variability enough (or, in other words, richness in structures enough) for an accurate estimation of the background error covariance.

### 5.2.3 The impact of a very simple vertical covariance localization on forecast verification scores

All results presented so far in this paper have been based on 4D-En-Var Hybrid with a horizontal localization only of the ensemble-based covariance matrix. In order to avoid spurious influence of, for example, near-surface observations in the stratosphere, and vice versa, it is necessary to introduce also a vertical localization, see Clayton et al. (2013).

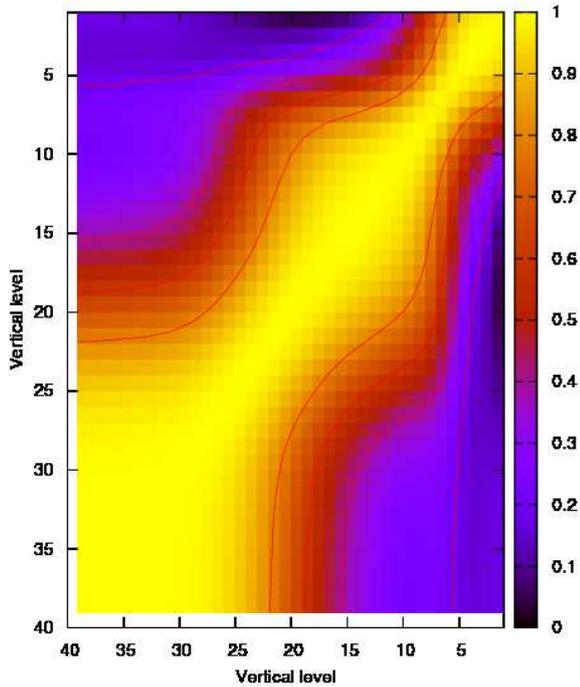
A first trial to introduce a vertical localization has been carried out also for the HIRLAM 4D-En-Var Hybrid. We simply started from the climatological background error vertical correlation matrix for vorticity for the largest horizontal scales (horizontal wave numbers 1–5). The eigenvectors of this vertical correlation matrix are presented in Fig. 11. The



**Figure 11.** First five eigenvectors (vertical modes) of the vertical correlation matrix for large horizontal scale vorticity (wave numbers 1–5) background errors, as applied in the HIRLAM variational data assimilation: mode 1 (red curve), mode 2 (green curve), mode 3 (dark blue curve), mode 4 (pink curve) and mode 5 (light blue curve).

first eigenvector (mode number 1) represents an average over the total atmospheric column, the second eigenvector represents a separation between the stratosphere and the low troposphere, while the third eigenvector has a maximum in the middle troposphere. A vertical localization correlation matrix, see Fig. 12, was constructed from these three vertical modes by setting eigenvalues of the remaining vertical modes (modes 4–40 in the present case with 40 vertical model levels) to zero and by re-normalizing the vertical correlations based on these three first vertical modes.

One real observation data assimilation experiment (4D-En-Var-Hyb50-NoSc) was re-run for the whole data period with vertical localization of the ensemble-based covariance added (experiment name Vertloc). Most forecast verification scores were slightly but consistently improved by this vertical localization, see for example verification scores for mean sea level pressure (MSLP) over a Scandinavian domain in Fig. 13. One exception to this positive impact of vertical



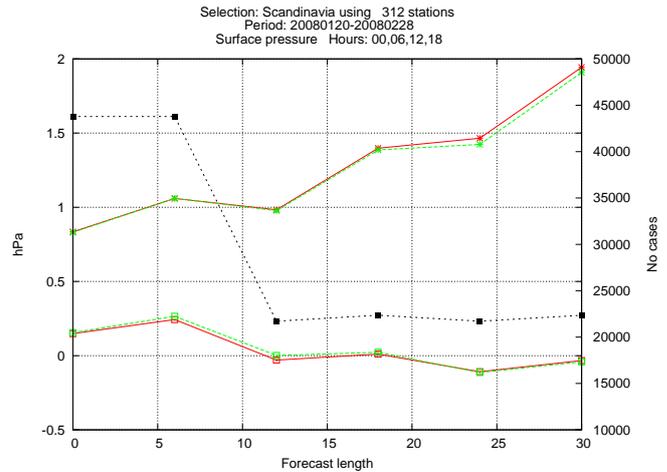
**Figure 12.** Vertical localization correlation matrix applied to ensemble-based background error covariances in the Vertloc assimilation experiment. The vertical localization matrix is based on the vertical correlation matrix for vorticity in the HIRLAM variational data assimilation, truncated to be valid for horizontal wave numbers 1–5 and to the three first vertical modes.

covariance localization on forecast verification scores was noticed for 700 hPa relative humidity forecasts (results are not shown). One likely explanation is that the vertical localization correlation matrix, derived by truncation from vorticity background error vertical correlations, does not fit very well the vertical variability of humidity background errors. Specific humidity background error vertical correlation patterns indicate two main vertical domains of activity, the planetary boundary layer and the tropospheric cloud layers above, and these are positioned vertically out of phase with the vertical correlation patterns seen in Fig. 12.

It need to be emphasized that this should be considered as an ad hoc trial to test the sensitivity of forecast verification scores to vertical localization of the ensemble-based background error covariances. Research is certainly needed to bring vertical localization, and localization in general, to a more scientifically founded level of development.

**5.2.4 Noise characteristics**

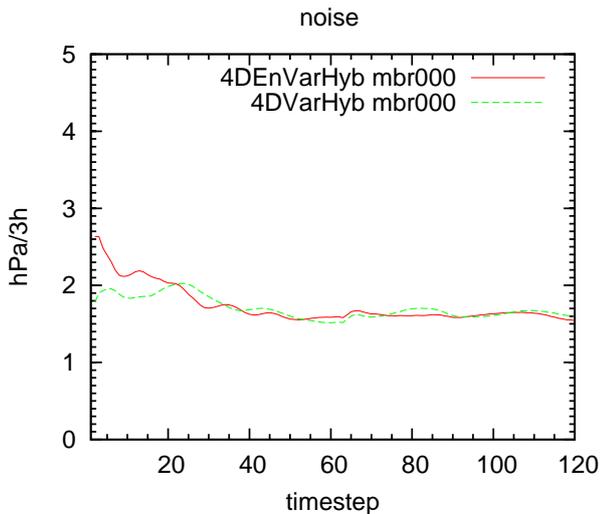
A weak digital filter constraint is included in the HIRLAM 4D-Var to control the level of high-frequency (gravity wave) oscillations during integration of the tangent linear and adjoint models over the data assimilation window. Although applied during the coarse-resolution tangent linear and ad-



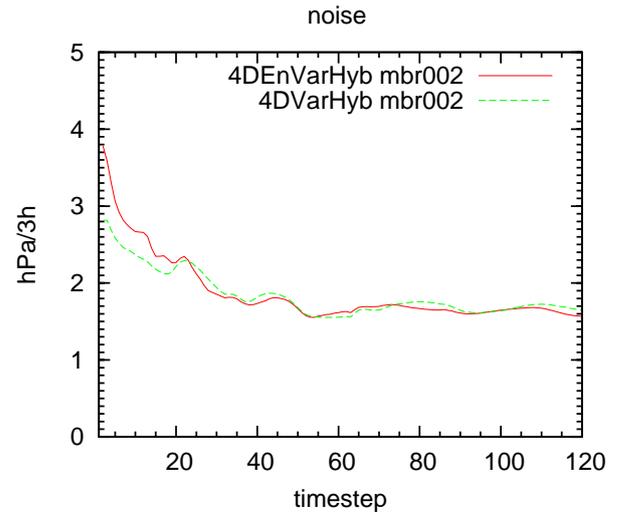
**Figure 13.** Verification of mean sea level pressure (MSLP) forecasts against surface (SYNOP) data from stations in a Scandinavian network. Standard deviation (STDV) and mean (bias) verification scores. Average verification scores over the period 17 January–28 February 2008 as a function of forecast length. 4D-En-Var Hybrid without vertical localization of the ensemble component of the background error covariance (experiment name 4D-En-Var-Hyb50-NoSc, red lines) and 4D-En-Var with vertical localization of the ensemble component of the background error covariance (experiment name Vertloc, green lines). The black dotted line indicates the number of verification comparisons.

joint model integrations, the application of this weak constraint during the data assimilation also efficiently prevents the presence of high-frequency (gravity wave) oscillations induced by the non-linear model during integrations from initial data created by the assimilation (Gustafsson et al., 2012). It is less straightforward to apply a similar digital filter constraint within the framework of 4D-En-Var since there is no model integration applied during the data assimilation process. The time resolution of the derived assimilation increments may not be sufficient for application of a digital filter. For this reason it is of interest to investigate the level of high-frequency oscillations in non-linear model integrations from initial model states produced by 4D-En-Var.

Figure 14 presents time series of the Sundqvist noise parameter, i.e. the horizontal average of the absolute value of the surface pressure tendency, for every time step during the first 12 h of the non-linear model integration, for the assimilation control (ensemble member 0) and for 4D-Var Hybrid and 4D-En-Var Hybrid. Figure 14 confirms that 4D-Var Hybrid, including a weak digital filter constraint, provides a quite noise-free non-linear model integration. It is also seen that 4D-En-Var Hybrid provides an initial state that causes a slightly increased noise level (from  $\approx 2.0$ hPa/3h for 4D-Var Hybrid to 2.5hPa/3h for 4D-En-Var Hybrid) during the first 1–2 h of model integration. We may consider that such a modest increase of the noise level hardly can harm the data assimilation process.



**Figure 14.** Time variation of the Sundqvist noise parameter, horizontal average of the absolute value of the surface pressure tendency ( $\text{hPa}/3\text{h}$ ), for each time step over the first 12 h of the non-linear model integration (model time step is 6 min). Ensemble control (mbr000). 4D-En-Var Hybrid (experiment name 4D-En-Var-Hyb50-Sc, red curve) and 4D-Var Hybrid (experiment name 4D-Var-Hyb25-Sc, green curve).



**Figure 15.** Time variation of the Sundqvist noise parameter, horizontal average of the absolute value of the surface pressure tendency ( $\text{hPa}/3\text{h}$ ), for each time step over the first 12 h of the non-linear model integration (model time step is 6 min). Ensemble member 2 (mbr002). 4D-En-Var Hybrid (experiment name 4D-En-Var-Hyb50-Sc, red curve), and 4D-Var Hybrid (experiment name 4D-Var-Hyb25-Sc, green curve).

For the ensemble members other than the ensemble control we do not apply data assimilation but rather an ETKF re-scaling of the forecast background perturbations to analysis perturbations representing analysis errors. The ETKF re-scaling may influence the balances of the forecast background model states and increase the noise level in non-linear model integration for these ensemble members. Figure 15 shows one example of the Sundqvist noise parameter from the time integration of ensemble member 2 (mbr002) and for both 4D-Var Hybrid and 4D-En-Var Hybrid. Indeed, the noise level, as measured by the Sundqvist parameter, is increased from  $\approx 2\text{hPa}/3\text{h}$  for the ensemble control to  $\approx 3\text{hPa}/3\text{h}$  for ensemble member 2 during the first few hours of non-linear model integration in the 4D-Var Hybrid experiment. Furthermore, it can be seen that the effects in the form of noise add up due to the imbalances caused by ETKF re-scaling and the 4D-En-Var Hybrid assimilation algorithm. With regard to possible negative effects of the combined imbalances created by the ETKF re-scaling and the 4D-En-Var Hybrid algorithms, one must keep in mind that the ensemble perturbations are only applied for estimation of background error covariances at the start of the next data assimilation cycle, thus at +3 h. At this forecast lead time model balances appear to have been stabilized through the adjustment processes of the high-frequency oscillations.

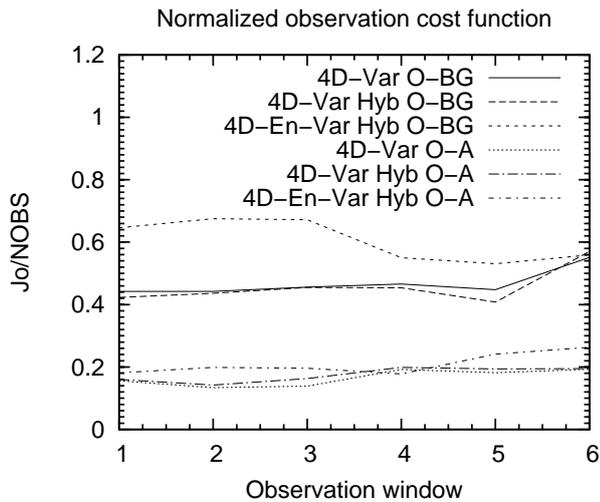
### 5.2.5 Observation fit statistics

It was noticed that MSLP forecasts for the shortest time range (+6 h) had slightly degraded Standard Deviation (STDV)

verification scores for the 4D-En-Var Hybrid experiment as compared to the 4D-Var and 4D-Var Hybrid experiments. Furthermore, it was also noticed that the level of high-frequency oscillations as measured by the surface pressure tendency was increased for the 4D-En-Var Hybrid, also compared with the 4D-Var and 4D-Var Hybrid experiments. In order to investigate further the use of surface pressure information in the 4D-En-Var Hybrid assimilation, we checked the behaviour of the surface pressure contribution to the observation part of the cost function during the minimization for one assimilation cycle, 22 February 2008, 12:00 UTC (the behaviour is similar for every assimilation cycle).

Figure 16 shows the SYNOP surface pressure contribution to the observation cost function, normalized with the number of surface pressure observations, as a function of observation time window within the 6 h data assimilation window for the 4D-Var, 4D-Var Hybrid and the 4D-En-Var Hybrid experiments. Separate curves are plotted before (O – BG, observation minus background) and after (O – A, observation minus analysis) the minimization. We may notice that the behaviour of these observation fit statistics are different for the 4D-En-Var Hybrid experiment as compared to the 4D-Var and 4D-Var Hybrid experiments.

The (O – BG) statistics indicate that the surface pressure forecasts for the observation windows 1–4 (corresponding background forecast length in the range 3–6 h) have larger errors for the 4D-En-Var Hybrid experiment than for the 4D-Var and 4D-Var Hybrid experiments. This confirms what we



**Figure 16.** The contribution of SYNOP surface pressure observations to the observation cost function ( $J_o$ ), normalized with the number of contributing observations (NOBS), as a function of the observation time window ( $-3, -2, -1, 0, +1, +2$  h). Assimilation cycle 22 February 2008, 12:00 UTC. Before minimization (O – BG) and after minimization (O – A). 4D-Var (experiment name 4DVar), 4D-Var Hybrid (experiment name 4DVar-Hyb25-Sc), and 4D-En-Var (experiment name 4DEnVar-Hyb50-Sc).

already have noticed with the forecast verification scores (see above).

The fit of the analysis to the surface pressure observations at the main observation hour (12:00 UTC, observation window 4) is closer for the 4D-En-Var Hybrid experiment as compared to the 4D-Var and 4D-Var hybrid experiments, while the fit is significantly worse for the 4D-En-Var Hybrid experiment for the other observation windows.

The cause of the different behaviour of surface pressure variations at short timescale in the different experiments is not obvious. At first sight, the increased level of high-frequency oscillations (measured by surface pressure tendencies) could possibly explain the worse surface pressure scores for the 4D-En-Var Hybrid experiment at short time range. However, this effect of imbalances in the initial 4D-En-Var Hybrid mainly affects forecasts in the range 1–2 h, while forecasts in the range 3–8 h, when the initial adjustments have already occurred, are applied as background states in the 4D-En-Var Hybrid.

Another possible explanation may be that the ensemble of 20 non-linear model trajectories in 4D-En-Var Hybrid does not provide the needed temporal variations to describe surface pressure variations with a timescale of 1 h, as provided by the observations, while the tangent linear model in 4D-Var is able to provide solutions that fit the observations more closely. This is not necessarily an advantage for the TL and AD model based HIRLAM 4D-Var. With a lack of a proper handling of biases and other representativity errors, the HIRLAM 4D-Var may over-fit the surface pressure data

when observations from the same stations are available for every hour. The increased surface pressure biases in longer forecasts for the 4D-Var based experiments and the recovery of quality of the 4D-En-Var Hybrid forecasts for the longer forecast may give some support to this hypothesis. Further investigations are needed to establish a better understanding.

## 6 Computational issues

The computing cost in 4D-En-Var Hybrid is quite different from the computing cost in 4D-Var. 4D-Var is dominated by heavy CPU utilization in the TL and AD models, while 4D-En-Var Hybrid is more dominated by input and handling of the ensemble of model trajectories needed for the estimation of the background error covariances. This can be seen in Table 2, which shows the measured wall clock computing times from a single assimilation cycle (minimization only) of 4D-Var, 4D-Var Hybrid and 4D-En-Var Hybrid. The computing times were measured from assimilation runs with 32 MPI tasks on a single compute node in an IBM Power 6-575 system. Since 4D-En-Var Hybrid is utilizing background error covariances based on non-linear model trajectories, 4D-En-Var Hybrid is using a single outer loop, while 4D-Var and 4D-Var Hybrid are using two outer loops in order to improve the linearization for the TL and AD models. Re-linearization for the observation operators could be motivated also for 4D-En-Var Hybrid, but the possibility for this has not yet been introduced in HIRLAM 4D-En-Var Hybrid.

The measured wall clock computing times of different components of 4D-En-Var Hybrid in Table 2 show that half of the wall clock time is spent in reading the ensemble of non-linear model trajectories over the data assimilation window (full model states for 20 ensemble members and for six observation time windows). In order to fully utilize the reduction in computing time that comes with avoiding the tangent linear and adjoint models in 4D-En-Var Hybrid, the reading of the model trajectories needs to be improved by making it parallel. This has not yet been implemented in the present version of HIRLAM 4D-En-Var Hybrid. Furthermore, the storing in memory of the ensemble of non-linear model trajectories has to be reduced by packing in fewer computer memory bits than occupied by double precision floating point numbers, taking the needs of higher model resolutions into account, both with regard to resolution in space and with regard to an increased number of observation time windows.

The longer computing time spent in reading and writing of observations, as well as in writing of field data, in 4D-Var and 4D-Var Hybrid, as compared to 4D-En-Var Hybrid, is explained by the two outer loops in 4D-Var and 4D-Var Hybrid, while a single outer loop is applied in 4D-En-Var Hybrid. The total number of inner loop minimization iterations is the same in all runs, as can be seen in the computing time for the observation operators.

**Table 2.** Wall clock computing time of different components of 4D-Var, 4D-Var Hybrid and 4D-En-Var Hybrid with one assimilation cycle, 22 February 2008, 12:00 UTC. 4D-Var and 4D-Var Hybrid are run with two outer loops, while 4D-En-Var Hybrid is run with a single outer loop. The computing times were measured from assimilation runs with 32 MPI tasks on a single compute node in an IBM Power 6-575 system (OpenMP was not utilized).

Assimilation component	4D-Var	4D-Var Hybrid	4D-En-Var Hybrid
Total	693 s	846 s	418 s
Read field data	46 s	56 s	216 s
Read obs. data	61 s	71 s	23 s
Minimization, total	373 s	339 s	91 s
TL model	141 s	109 s	–
AD model	191 s	146 s	–
Observation operators	20 s	20 s	20 s
Write field data	118 s	107 s	53 s
Write obs. data	44 s	30 s	15 s

## 7 Discussion and concluding remarks

A four-dimensional ensemble variational (4D-En-Var) data assimilation scheme has been developed for the HIRLAM forecasting system. The scheme is a hybrid between 4D-En-Var, with a flow-dependent background error covariance component estimated from an ensemble of non-linear model trajectories over the data assimilation window, and 3D-Var FGAT (first guess at appropriate time) with a static background error covariance component applied in the middle of the data assimilation window. The HIRLAM 4D-En-Var Hybrid is applied to the ensemble control only; initial states for other ensemble members are obtained by a basic ETKF (ensemble transform Kalman filter) re-scaling scheme. In order to prevent negative effects of the low rank of the ensemble-based background error covariance matrix, a localization of the ensemble covariance matrix based on a Schur product with a localization correlation matrix is applied.

Single simulated observation data assimilation experiments have been carried out to demonstrate the effects of flow dependency of the ensemble-based background error covariances. These experiments show an increased level of realistic flow dependency, both for 4D-Var Hybrid in comparison with standard 4D-Var and for 4D-En-Var Hybrid in comparison with 4D-Var Hybrid. The experiments also give support to the hypothesis that the ETKF re-scaling algorithm provides initial perturbations with a rich and physically relevant variability.

Real observation experiments over a period of 6 weeks to demonstrate the effects of the assimilation on forecast quality have also been carried out. Comparison is made with standard HIRLAM 4D-Var and with HIRLAM Hybrid 4D-Var ensemble data assimilation, where the ensemble based back-

ground error covariance is applied at the start of the assimilation window only. The real observation data assimilation experiments indicate that the 4D-En-Var Hybrid outperforms standard 4D-Var as well as 4D-Var Hybrid with regard to verification scores for forecasts based on initial data from the respective assimilation method. In particular, relative humidity verification scores are significantly improved by 4D-En-Var Hybrid, and this is likely to be linked to the too simplified tangent linear model applied in the HIRLAM 4D-Var.

In contrast to 4D-Var and 4D-Var Hybrid, no time integration of tangent linear and adjoint models is needed in 4D-En-Var Hybrid. This results in a significant reduction of computing time, and possibilities for better scalability, for 4D-En-Var Hybrid. The handling of the ensemble of non-linear model trajectories over the assimilation window needs to be designed carefully, however.

The successful implementation of HIRLAM 4D-En-Var Hybrid for synoptic-scale numerical weather prediction has encouraged us to develop a 4D-En-Var Hybrid also for numerical weather prediction at the convection permitting scales with model grid resolutions of a few km. The increased importance of moist processes and non-linearities at these smaller scales should be an advantage for 4D-En-Var since the assimilation increments are formed as localized combinations of non-linear model forecast perturbations rather than derived from time integrations with a simplified tangent linear model.

There are certainly many scientific uncertainties associated with the development of 4D-En-Var Hybrid, and data assimilation in general, for the convection permitting scales. Which spatial scales are described by present and future observing systems, and which spatial scales are meaningful to assimilate for convection permitting models? For synoptic-scale 4D-Var data assimilation, the application of incremental data assimilation with tangent linear model grid resolutions increased by a factor of 3–5 as compared to the non-linear model grid resolution, has been quite successful (Gustafsson et al., 2012). Does this also apply to the data assimilation at convection permitting scales? In 4D-En-Var for example, do we need an ensemble generated from full resolution non-linear model integrations in order to derive the appropriate background error covariances for data assimilation at convection-permitting scales? Since such a full resolution convection-permitting ensemble of model integrations would be computational demanding, it might be tempting to derive the ensemble needed for 4D-En-Var with a slightly reduced non-linear model resolution.

The 4D-En-Var method described here can easily be generalized to a simple method for nowcasting by extending the same weighting of ensemble background forecast perturbations beyond the data assimilation window.

Most results presented here were based on HIRLAM 4D-En-Var Hybrid using horizontal localization of ensemble-based background error covariances only. The results from an ad hoc trial with a simple vertical localization have

encouraged us to continue research on horizontal as well as vertical localization and, in order to permit a weak constraint formulation, also on temporal localization.

*Acknowledgements.* We would like to thank Dale Barker, Andrew Lorenc and Neil Bowler at the UK Met Office for information and discussions that helped us to develop the HIRLAM 4D-En-Var. Lars Axell inspired us with his pioneering work to develop 4D-En-Var for oceanographic models and discussions with Tomas Landelius and Magnus Lindskog were also helpful. Two anonymous reviewers helped to improve this paper. This work was supported by the Swedish National Space Board (Contract Dnr 105/08:3). The HIRLAM (HIGH Resolution Limited Area Model) project is a joint effort among the weather services in Denmark, Estonia, Finland, France, Iceland, Ireland, Lithuania, the Netherlands, Norway, Spain and Sweden for development of short-range numerical weather prediction techniques.

Edited by: O. Talagrand

Reviewed by: two anonymous referees

## References

- Bergthorsson, P. and Döös, B.: Numerical weather map analysis, *Tellus*, 7, 329–340, 1955.
- Berre, L.: Estimation of synoptic and mesoscale forecast error covariances in a limited area model, *Mon. Weather Rev.*, 128, 644–667, 2000.
- Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects, *Mon. Weather Rev.*, 129, 420–436, 2001.
- Bojarova, J., Gustafsson, N., Johansson, A., and Vignes, O.: The ETKF rescaling scheme in HIRLAM, *Tellus*, 63A, 385–401, 2010.
- Buehner, M.: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational nwp setting, *Q. J. Roy. Meteor. Soc.*, 131, 1013–1043, 2005.
- Buehner, M., Houtekamer, P., Charette, C., Mitchell, H., and He, B.: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments, *Mon. Weather Rev.*, 138, 1550–1566, 2010a.
- Buehner, M., Houtekamer, P., Charette, C., Mitchell, H., and He, B.: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations, *Mon. Weather Rev.*, 138, 1567–1586, 2010b.
- Clayton, A. M., Lorenc, A. C., and Barker, D. M.: Operational implementation of a hybrid ensemble/4D-Var global data assimilation at the Met Office., *Q. J. Roy. Meteor. Soc.*, 139, 1445–1461, 2013.
- Cohn, S. E. and Todling, R.: Approximate data assimilation scheme for stable and unstable dynamics, *J. Met. Soc. Jpn.*, 74, 63–75, 1996.
- Courtier, P., Thepaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. Roy. Meteorol. Soc.*, 120, 1367–1387, 1994.
- Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and Fisher, M.: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation, *Q. J. Roy. Meteorol. Soc.*, 124, 1783–1808, 1998.
- Cuxart, J., Bougeault, P., and Redelsperger, J.-L.: A turbulence scheme allowing for mesoscale and large-eddy simulations, *Q. J. Roy. Meteorol. Soc.*, 126, 1–30, 2000.
- Desroziers, G., Camino, J.-T., and Berre, L.: 4D-EnVar: link with 4D state formulation of variational assimilation and different possible implementations, *Q. J. Roy. Meteorol. Soc.*, doi:10.1002/qj.2325, in press, 2014.
- Evensen, G.: Sequential data assimilation with a nonlinear QG model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10143–10162, 1994.
- Gandin, L. S.: Objective analysis of meteorological fields, Leningrad Hydromet. Press, 1963.
- Gustafsson, N.: A review of methods for objective analysis, in: *Dynamic Meteorology: Data Assimilation Methods*, edited by: Bengtsson, L., Ghil, M., and Källén, E., Springer Verlag, New York, 17–76, 1981.
- Gustafsson, N.: Discussion on “4D-Var or EnKF”, *Tellus A*, 59, 774–777, 2007.
- Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K. S., and Thorsteinsson, S.: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint, *Tellus A*, 53, 425–446, 2001.
- Gustafsson, N., Huang, X.-Y., Yang, X., Mogensen, K., Lindskog, M., Vignes, O., Wilhelmsson, T., and Thorsteinsson, S.: Four-dimensional variational data assimilation for a limited area model, *Tellus A*, 64, 14985, doi:10.3402/tellusa.v64i0.14985, 2012.
- Gustafsson, N., Bojarova, J., and Vignes, O.: A hybrid variational ensemble data assimilation for the HIGH Resolution Limited Area Model (HIRLAM), *Nonlin. Processes Geophys.*, 21, 303–323, doi:10.5194/npg-21-303-2014, 2014.
- Hamill, T. M. and Snyder, C.: A hybrid ensemble Kalman filter-3D variational analysis scheme, *Mon. Weather Rev.*, 128, 2905–2919, 2000.
- Haugen, J. E. and MACHENHAUER, B.: A spectral limited-area model formulation with time-dependent boundary conditions applied to the shallow water equations, *Mon. Weather Rev.*, 121, 2618–2630, 1993.
- Heemink, A. W., Verlaan, M., and Segers, A. J.: Variance reduced ensemble Kalman filtering, *Mon. Weather Rev.*, 129, 1718–1728, 2001.
- Houtekamer, P. and Mitchell, H.: A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, 129, 123–137, 2001.
- Kain, J. S.: The Kain-Fritsch convective parameterization: an update, *J. Appl. Meteorol.*, 43, 170–181, 2004.
- Kalman, R. E.: A new approach to linear filter and prediction theory, *J. Basic. Eng.-T. ASCE*, 82, 35–45, 1960.
- Kalman, R. E. and Bucy, R.: New results in linear filtering and prediction theory, *J. Basic Eng.-T. ASCE*, 83D, 95–108, 1961.
- Kalnay, E., Li, H., Miyoshi, T., Yang, S., and Ballabrera-Poy, J.: 4D-Var or ensemble Kalman filter?, *Tellus A*, 59, 758–773, 2007.

- Kawabata, T., Seko, H., Saito, K., Kuroda, T., Tamiya, K., Tsuyuki, T., Honda, Y., and Wakazuki, Y.: An assimilation and forecasting experiment of the Nerima heavy rainfall with a cloud-resolving nonhydrostatic 4-dimensional variational data assimilation system, *J. Met. Soc. Jpn.*, 85, 255–276, 2007.
- Keper, J. D.: Covariance localisation and balance in an ensemble Kalman filter, *Q. J. Roy. Meteorol. Soc.*, 135, 1157–1176, 2009.
- Le Dimet, F. X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations. Theoretical aspects, *Tellus A*, 37, 309–322, 1986.
- Lewis, J. M. and Derber, J. C.: The use of adjoint equations to solve a variational adjustment problem with advective constraints, *Tellus A*, 37, 309–322, 1985.
- Lindskog, M., Gustafsson, N., Navascués, B., Mogensen, K. S., Huang, X.-Y., Yang, X., Andrae, U., Berre, L., Thorsteinsson, S., and Rantakokko, J.: Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments, *Tellus A*, 53, 447–468, 2001.
- Liu, C. and Xiao, Q.: An ensemble-based four-dimensional variational data assimilation scheme. Part III: Antarctic applications with advanced WRF using real data, *Mon. Weather Rev.*, 141, 2721–2739, 2013.
- Liu, C., Xiao, Q., and Wang, B.: An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test, *Mon. Weather Rev.*, 136, 3363–3373, 2008.
- Liu, C., Xiao, Q., and Wang, B.: An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with advanced research WRF (ARW), *Mon. Weather Rev.*, 137, 1687–1704, 2009.
- Lorenc, A.: A global three-dimensional multivariate statistical interpolation scheme, *Mon. Weather Rev.*, 109, 701–721, 1981.
- Lorenc, A.: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var, *Q. J. Roy. Meteorol. Soc.*, 129, 3183–3203, 2003.
- Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterization scheme, *Global Planet. Change*, 13, 145–159, 1996.
- Parrish, D. F. and Derber, J. C.: The National Meteorological Centre's spectral statistical interpolation analysis system, *Mon. Weather Rev.*, 120, 1747–1763, 1992.
- Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J. F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics, *Q. J. Roy. Meteorol. Soc.*, 126, 1143–1170, 2000.
- Rasch, P. J. and Kristjánsson, J. E.: A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations, *J. Climate*, 11, 1587–1614, 1998.
- Rutherford, I. D.: An operational three-dimensional multivariate statistical objective analysis scheme, GARP WGENE Report No. 11, 98–121, 1976.
- Sasaki, Y.: An objective analysis based on the variational method, *J. Met. Soc. Jpn.*, 36, 77–88, 1958.
- Savijärvi, H.: Fast radiation parameterization schemes for mesoscale and short-range forecast models, *J. Appl. Meteorol.*, 29, 437–447, 1990.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF version3, NCAR Tech. Note NCAR/TN-4751 STR, 125 pp., 2008.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S.: Ensemble square root filters, *Mon. Weather Rev.*, 131, 1485–1491, 2003.
- Trémolet, Y.: Accounting for an imperfect model in 4D-Var, *Q. J. Roy. Meteorol. Soc.*, 132, 2483–2504, 2006.
- Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Lenderlink, G., McDonald, A., McGrath, R., Navascués, B., Woetman Nielsen, N., Odegaard, V., Rodriguez, E., Rummukainen, M., Rööm, R., Sattler, K., Hansen Sass, B., Savijärvi, H., Wichers Schreur, B., Sigg, R., The, H., and Tijm, A.: HIRLAM-5 Scientific Documentation, HIRLAM-5 project, c/o Per Undén, SMHI, SE-60176 Norrköping, Sweden, 2002.
- Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M.: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part I: Observing system simulation experiment, *Mon. Weather Rev.*, 136, 5116–5131, 2008a.
- Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M.: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part II: Real observation experiments, *Mon. Weather Rev.*, 136, 5132–5147, 2008b.