School of Mathematics, Meteorology
and Physics

Department of Mathematics

Preprint MPS_2010-10

20 April 2010

Breakdown of hydrostatic balance at convective scales in the forecast errors in the Met Office Unified Model

by

Sanita Vetra-Carvalho, Mark Dixon, Stefano Migliorini, Nancy K Nichols, Susan P Ballard

Breakdown of h drostatic balance at convective scales in the forecast errors in the Met Office Unified Model

*Sanita Vetra-Carvalho *, Mark Dixon , Stefano Migliorini , Nancy K. Nichols *, Susan P Ballard*

* Department of Mathematics, University of Reading, UK

\# JCMM, Met Office, UK

% Department of Meteorology, University of Reading, UK

Abstract

P^f P^f P^f fl P^f ffi ffi fi

1 Introduction

Due to a continuing increase in computer power, it has become possible for meteorological centres to run high resolution models. These models are expected to produce more realistic forecasts because of their better representation of small-scale forcing from orography and land use, as well as their explicit representation of convection. Modelling at high resolution should thus lead to more accurate forecasting of high impact weather events such as flooding, with potential social and economic benefits.

Another important advantage of high-resolution models is that these may be used to assimilate high-resolution observations, e.g. from radar. To do so, it is important that the assumptions made for assimilation are still valid at these scales. For example, data assimilation systems used by operational meteorological centres such as the Met O ce

2 E s p s s s s

used, which seeks to ensure that the perturbation spread matches the RMS of the mean forecast,

$$\Pi_m = \Pi_{m-1}\sqrt{\frac{((\ (\mathbf{d}_m\mathbf{d}_m^T) - \ (\mathbf{R}))\ (\mathbf{S}_{m-1}))^{\frac{1}{2}}}{\ (\mathbf{S}_m)}} \tag{9}$$

where $\mathbf{d}_m = \mathbf{y} - \overline{\ (\mathbf{X}^f)}$ is the `average' innovation vector and $\mathbf{S}_m = \mathbf{H}\mathbf{P}^f\mathbf{H}^T$ is the spread of the forecast ensemble in observational space for forecast cycle . (Here $(\mathbf{X}^f)$ denotes the matrix with columns given by $(\mathbf{x}_i^f)$ and $\overline{\ (\mathbf{X}^f)}$ denotes the vector average of these columns.)

2.2 The Met Office Unified Model

The 1.5 km version of the UM used in this work uses non-hydrostatic deep atmosphere equations with a hybrid height/terrain-following vertical coordinate [4]. The model has staggered grids in the horizontal and the vertical. The Arakawa C-grid is used for horizontal staggering where the zonal velocity component is east-west staggered, and the temperature and the meridional velocity component are north-south staggered. The Charney Phillips grid is used for vertical staggering, where potential temperature is on the same levels as the vertical velocity.

The high resolution model has a grid length of 1.5 km with 360 grid points in latitude and 288 in longitude covering Southern England and Wales. The model has a grid with 70 vertical levels, where only about 50 lowermost levels are a ected by orography. The 24 km NAE model has 360 grid-points in latitude, 215 in longitude and 38 vertical levels.

The current data assimilation (DA) system for the 1.5 km model is similar to that used with the operational UK 4 km Met O ce model, which is discussed in detail in [6]. In summary, the DA combines a 3D-Var scheme, used to assimilate the conventional observations producing the large scale analysis, and nudging systems used to update the high resolution moisture and surface precipitation data. The system uses a cloud nudging (CN) procedure to nudge humidity increments, whereas surface precipitation rates are assimilated via latent heat nudging (LHN). The increments produced by 3D-Var and nudging procedures are used to correct the model trajectory at each time step during the DA window. The main di erences between the 1.5-km and the 4-km model DA configurations are as follows. The former uses hourly assimilation cycles, rather than the three-hourly used in the operational 4-km system, starting from fields interpolated to the southern England and Wales 1.5-km grid from the operational UK 4-km forecast. The 1.5-km DA system also uses more frequent cloud (hourly) and precipitation (every 15 minuntes) observations.

2.3 Generation of high-resolution perturbations

To obtain the initial condition ensemble at 1.5 km resolution, the following steps are performed (see figure 7 in Appendix 1). First, an ensemble of atmospheric fields is formed by adding an ensemble of operational MOGREPS NAE analysis perturbations at 24 km resolution valid at a given time (18Z), denoted as $\mathbf{X}'_{24km}$, to the operational 4D-Var atmospheric analysis at 18Z over the NAE domain, denoted as $\overline{\mathrm{x}}_{24km}$, resampled at a resolution of 24 km from its original resolution of 12 km. This 24 km resolution ensemble, denoted as X_{24km}, and the 4D-Var control, $\overline{\mathrm{x}}_{24km}$, are interpolated to a 1.5 km resolution grid to obtain $\mathrm{X}_{1:5=24km}$ and $\overline{\mathrm{x}}_{1:5=24km}$, respectively. The ensemble of perturbations at 1.5 km resolution, denoted as $\mathrm{X}'_{1:5=24km}$, is then obtained by subtracting the reconfigured 4D-Var analysis $\overline{\mathrm{x}}_{1:5=24km}$ from each column of the ensemble matrix $\mathrm{X}_{1:5=24km}$. Finally, to obtain the ensemble of initial conditions at 1.5 km resolution at 18Z, denoted as $\mathrm{X}_{1:5km}$, the 1.5 km ensemble perturbations $\mathrm{X}'_{1:5=24km}$ are combined with a 3D-Var high-resolution analysis { including nudging of precipitation and cloud observations { valid at 18Z from the 1.5 km model and data assimilation system over the southern UK [6]. The 18Z analysis was generated as part of a 1.5-km assimilation experiment with hourly DA cycles that where $_p$ = 1005 J kg^{-1}K^{-1} is the specific heat at constant pressure, $_0$ = 1000 hPa is a reference pressure at the ground level

d[(0)]TJ /F912 8959.69Tf 9.1the21(+the23((4(lev)26(el)]TJ

where $'_i$, $= 1, \ldots, 24$ denote the ensemble members { vectors containing the fields at all vertical grid points at a given location.

Then the hydrostatically balanced potential temperature perturbation, $'_H$, is computed for each element of each ensemble member, $= 1, \ldots, 24$, leading to the ensemble vertical error covariance matrix

$$\mathbf{P}_e = \left\langle \, '_H \; '^T_H \right\rangle \tag{23}$$

where $'_H$ is the hydrostatically balanced potential temperature perturbation ensemble. The correlation matrix $\mathbf{C}_e$ is obtained by scaling matrix $\mathbf{P}_e$ by its own variance.

4 Results

This section discusses the results from applying the equation (19) to the data obtained from the ensemble of 1.5 km forecasts. The model was initialised with a set of ensemble atmospheric states valid at 18Z on 27/07/2008 and determined as explained in section 2 and Appendix 1. Ensemble forecasts were produced at each hour for the following 3 hours, 19Z, 20Z, and 21Z, on the same day, 27/07/2008. This case was selected as the convection had already occurred before 18Z and at the time of initialisation, 18Z, the system was fully convective with convection moving in the domain over the next three hours.

Although figures showing the degree of hydrostatic balance present in forecast perturbations were computed for various vertical columns, here only the "extreme" cases are investigated. The column for which the ensemble mean precipitation was zero over the entire 3 hours is labeled 'Non-Conv' and columns for which the ensemble mean precipitation was the highest for each hour [1] are labeled, 'Conv19Z', 'Conv20Z', and 'Conv21Z', respectively. Column location is indicated in the figure 1.

To find the degree to which hydrostatic balance holds in the perturbations as a function of horizontal scale the original 1.5 km resolution data were aggregated into boxes ranging from sides 3 km up to 90 km resolution around the non-convective point 'Non-Conv' and the convective column 'Conv19Z'.

The following quantities were calculated from both the 1.5 km and the coarsened resolution data:

- Correlation matrices for $'$, $'_{\mathbf{H}}$, see figures 2, 3.
- Explained variances computed at each vertical level and 19Z, 20Z, 21Z, see figures 4, 5.
- Root mean square (RMS) errors between $'$ and $'_{\mathbf{H}}$ computed for each of the four columns as a mean error over all members, levels and the three hour forecast window, see table 1.
- Mean ensemble time-dependent error between $'$ and $'_{\mathbf{H}}$ computed for each of the four columns at each vertical level and for each hour as a mean error over all ensemble members, see figure 6.

[1]

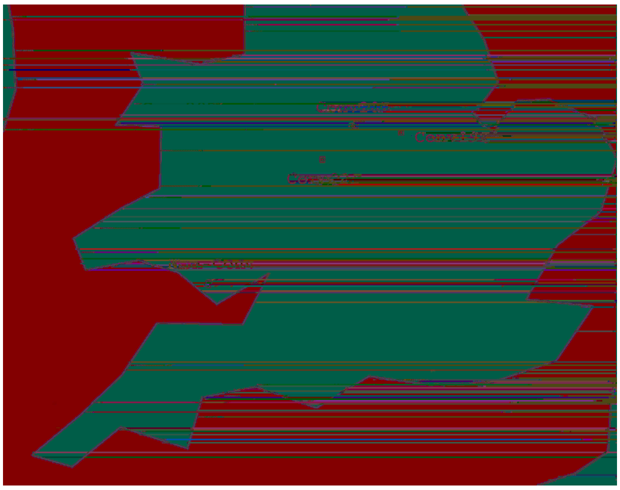

Fig. 1: 1.5 km resolution domain and the chosen vertical columns for testing

The dependence of hydrostatic balance can be ascertained in a quantitative way by means of the two error measures and the explained variance and in a more subjective way by inspection of the correlation matrices.

4.1 Correlation matrices for ^{0} and ^0_H

The balanced and "raw" data ensemble correlation matrices of $'$ and $'_H$ for the columns 'Non-Conv' and 'Conv19Z', at 1.5 km resolution, are shown in figure 2 for 19Z. As discussed in section 3, the balanced variable $'_H$ is computed using equation (19) for each ensemble member and each vertical level. From figure 2 we can see that at 1.5 km resolution in the case of no convection (figures 2(a) and 2(b)) the ensemble correlation matrices for $'$ and $'_{\mathbf{H}}$ are indistinguishable, meaning that hydrostatic balance holds very well in the perturbations when convection is not present. However, in the presence of convection (figures 2(c) and 2(d)) the ensemble correlation matrices for $'$ and $'_H$ are clearly di erent, especially just above the boundary layer (between vertical levels 20 and 40) where the convection is the strongest. Hence, in the cases (vertical levels) where the convection is present, the balance is no longer valid in the perturbations.

At 1.5 km resolution when convection is present, the $'$ is less correlated in the boundary layer (vertical levels 0 to 20) than $'_H$. However, $'$ is more correlated than $'_H$ right above the boundary layer at vertical levels 20 to 30.

By coarsening the grid we expect the perturbations to become more hydrostatically balanced. This is visually confirmed in figure 3, where correlation matrices of $'$ and $'_H$ for convective columns at 4.5 km and 12 km resolutions are shown. Notice, that even though at these resolutions the perturbations are not in hydrostatic balance in the mid-atmosphere (levels 20 - 40), they appear to be much more in the balance in the boundary layer.

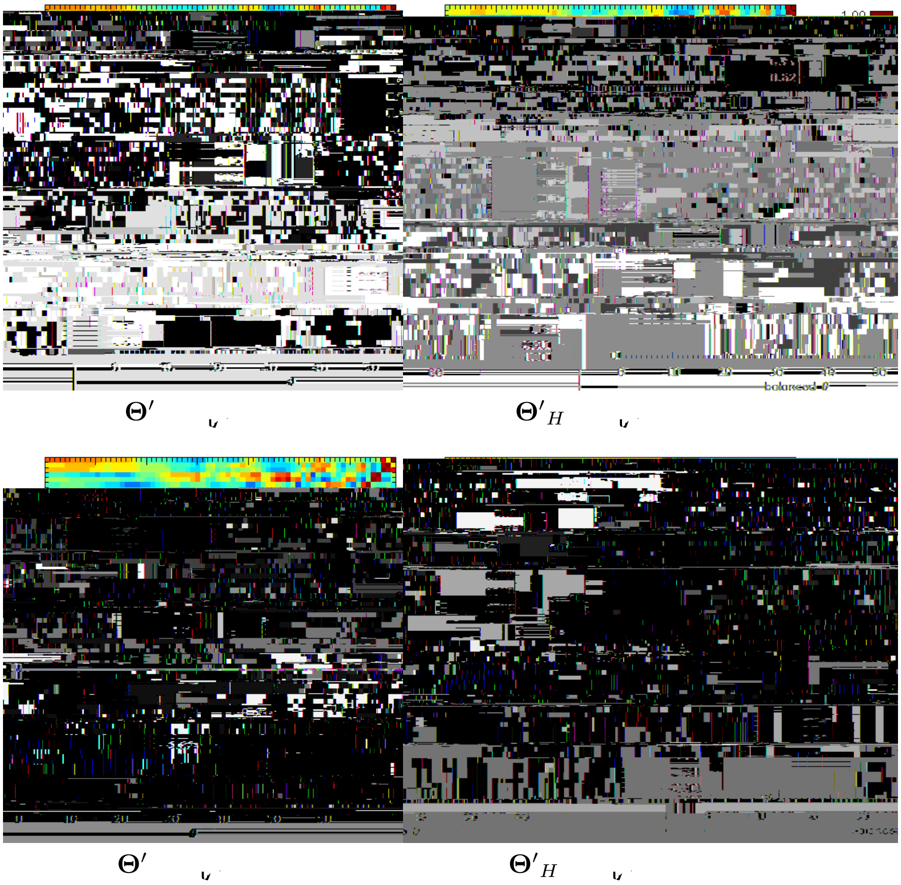

Fig. 3: Auto-correlations for $'$ and auto-correlations for the corresponding $'_H$ at 19Z for the convective column 'Conv19Z4km' at 4.5 km resolution and the convective column 'Conv19Z12km' at 12 km resolution

tions in all of the non-convective columns are very well explained by hydrostatic balance. Figure 5 shows the explained variance as a function of height for the convective columns from 3 km up to 22.5 km coarsened data at 19Z, 20Z, and 21Z. This shows the degree to which hydrostatic balance (in convective regions) increases as a function of horizontal

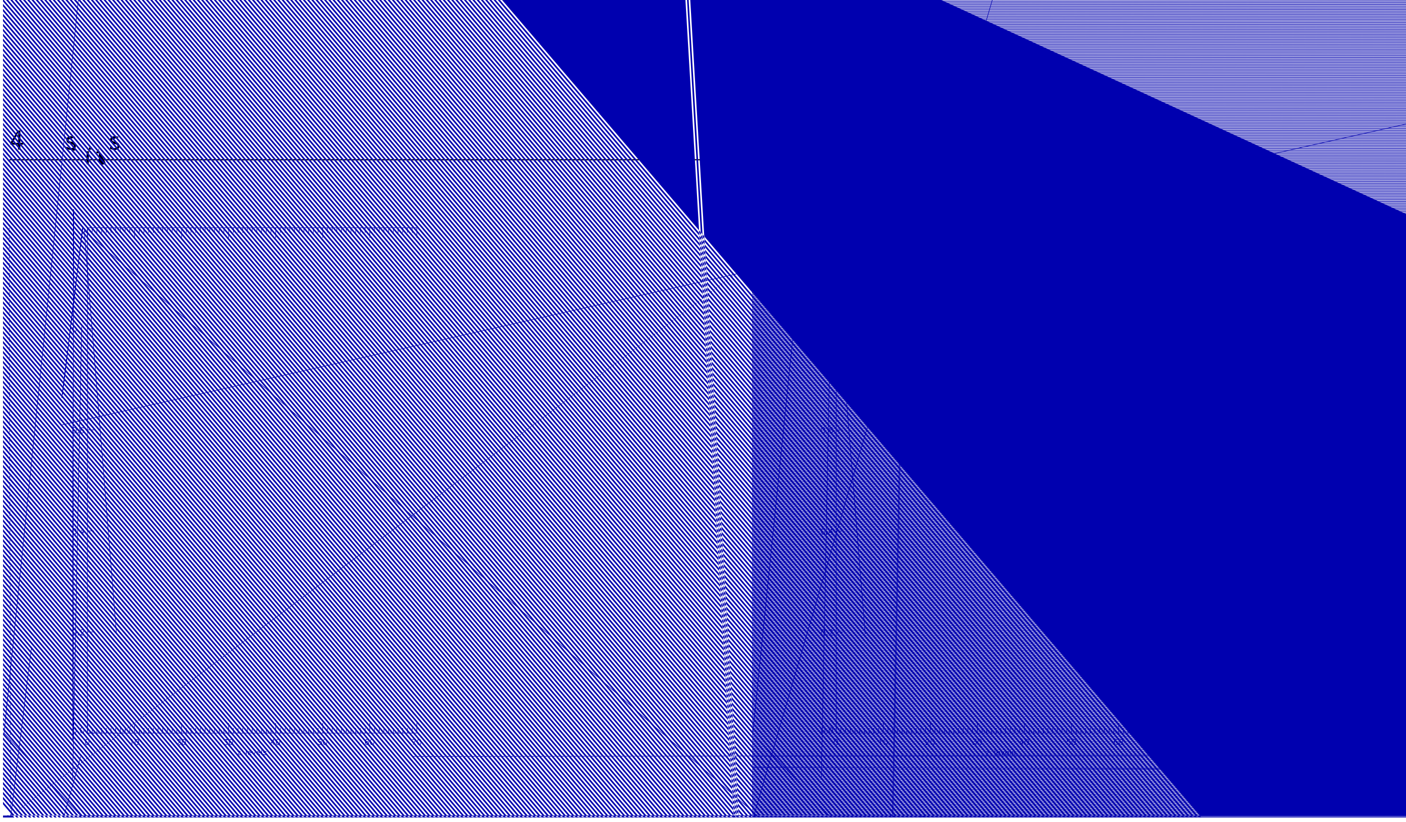

Fig. 4: Explained variance ( ) at 1.5 km for each vertical level at a) 19Z, b) 20Z and c) 21Z

For the 1.5 km data the 'Conv20Z' and 'Conv21Z' are more in hydrostatic balance initially at 19Z, and they become more unbalanced as convection becomes stronger in these columns at 20Z and 21Z, respectively.

4.3 High resolution RMS errors between $'$ and $'_H$

Another measure of the hydrostatic balance in the perturbations is the RMS error between $'_H$ and $'$, using the standard formula for the RMS,

$$\sqrt{\overline{(\ '_{H} - \ ')}}$$

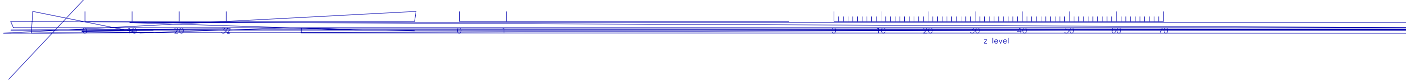

Fig. 5: Explained variance ( ) for resolutions of 3 km to 22.5 km of convective 'Conv19Z'

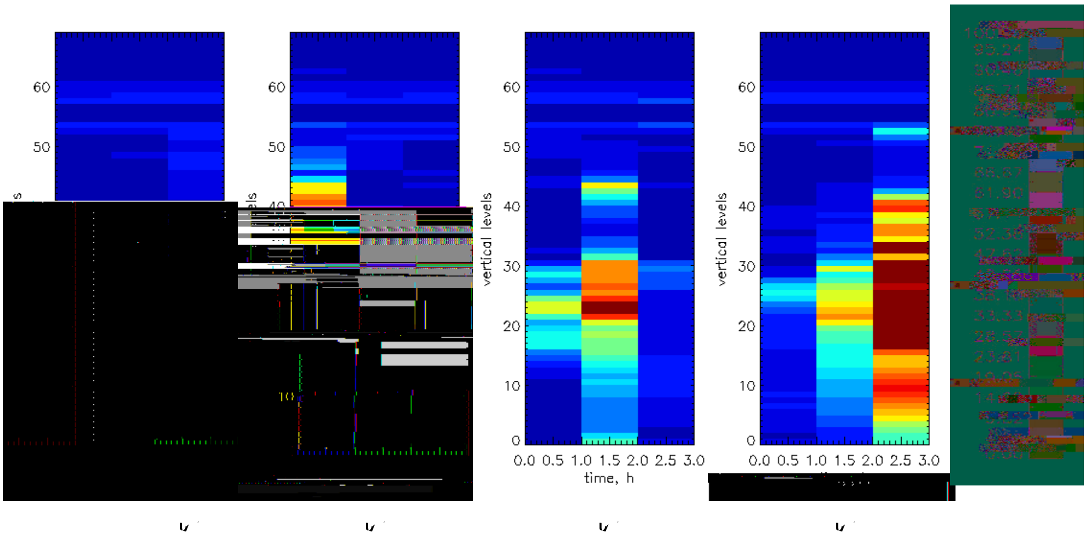

Fig. 6: Relative error between $'_H$ and $'$ averaged over all ensemble members in vertical space and time at 1.5 km resolution

5 Summar and conclusions

To investigate how well the hydrostatic balance holds for forecast errors at convective scales we used an ensemble with 24 members obtained from running the 1.5 km resolution version of the UM that was initialised according to the procedure described in section 2.3 and in Appendix 1. The ensemble was initialised at 18Z on 27/07/2008 when convection was fully developed and data for analysis were available at 19Z, 20Z, and 21Z on the same day. In this paper we focused on the vertical analysis of forecast errors. Four vertical columns from the whole domain were selected for testing purposes for each forecast hour: a column with no precipitation and three columns with highest rain rates in the ensemble mean, for each forecast hour 19Z, 20Z and 21Z, respectively. Data around two columns, one non-convective and one convective were aggregated from 3 km up to 90 km resolution. For each of these columns the hydrostatically balanced potential temperature perturbations, $'_H$, were calculated using the approximated hydrostatic equation for perturbations, expressed in the terms of the available fields - potential temperature , Exner pressure Π and specific humidity .

By constructing the correlation matrices for these columns it was shown that at 1.5 km resolution the hydrostatic balance does not hold in the perturbations in the regions of convection but does hold in the regions where convection is not present. Note, that from the explained variances and mean error in the vertical we establish not only the extent to which the perturbations are not in hydrostatic balance but also, that perturbations are in balance at all resolutions in all columns in the stratosphere (above vertical level 55) where the atmosphere is dry. Also, from the explained variances we see that at the vertical levels 10 { 30 the perturbations are very far from being balanced. This suggests that the hydrostatic balance should be relaxed around these columns and levels in the correlation matrices at 1.5 km resolution. This would require a redesign of the control variable transform in variational data assimilation system used by the Met O ce.

We also showed using the explained variances that 20 km horizontal resolution is the

1 Appendi : Flow of the high resolution EPS

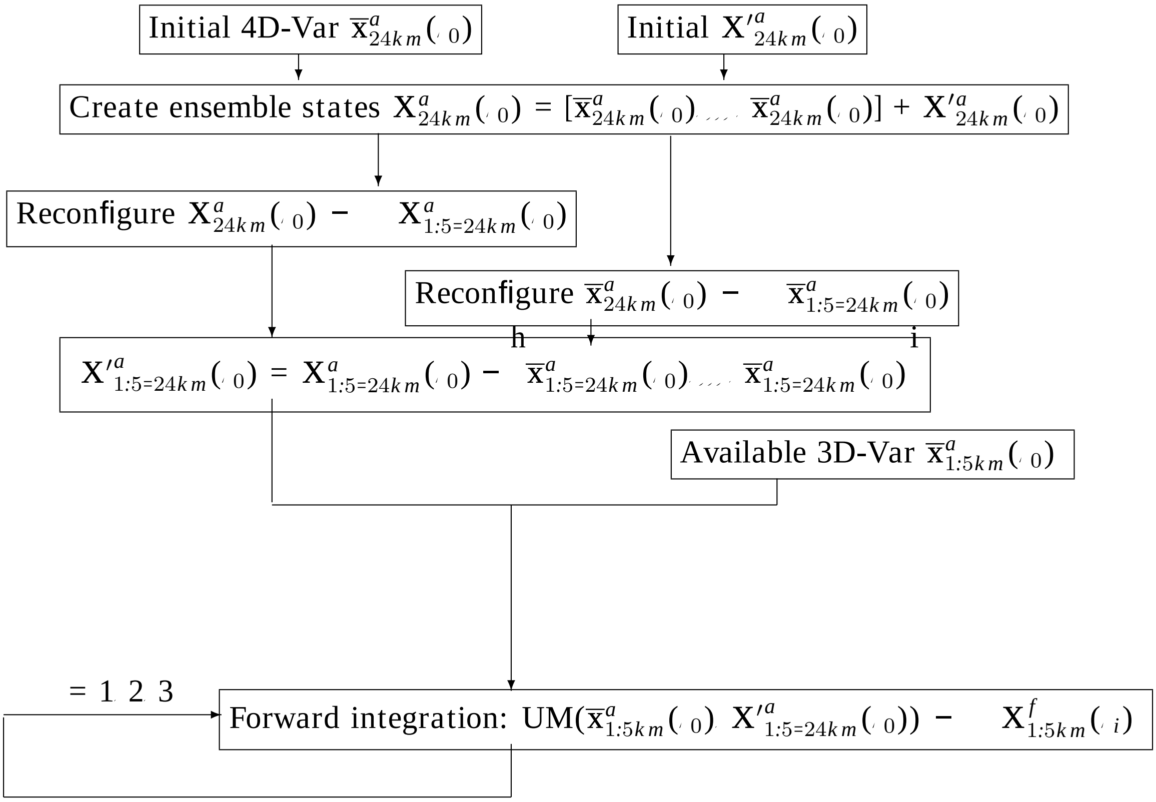

Fig. 7: Setup and flow of the reconfigured ensemble prediction system. Here $_0$=18Z, $_1$=19Z, $_2$=20Z, and $_3$=21Z all on 27/07/2008. The UM forward integration step at 1.5 km includes the ada8QBT/F17(adt;0T[=)-2(erturbn)-28511.955Tf5955Tfturbn

2 Appendi : RMS error of coarsened data