General comment

"The paper presents a rather standard 3D-VAR implementation as developed and well documented by the referred authors about 15 years ago. There is nothing new in the data assimilation development, but a documentation of still significant work to implement a basic variational analysis for RAMS. It is not clear to me what RAMS had before, but I assume and from what I can find, that it had no standard 3D-VAR on its model grid at all (in stead a lot of ad hoc methods like nudging and dynamical adaptation). Under this assumption, the author has made significant progress and the paper serves as a documentation to go with the RAMS system. The results are far to basic to be of scientific interest today, and the experimentation results are extremely basic. It shows nothing of relevance to mesoscale modelling and only that a basic DA gives better fit to observations than no DA.

There are also a few misunderstandings and one mistake I believe, in the implementation of the transform (full T or un-balanced T)."

About this general comment I point out that there is no standard variational system for the RAMS model, and few researchers (cited into the discussion paper) implemented an owned variational data assimilation system for this model. This paper is an effort in this direction.

I agree with the reviewer that the results presented are rather basic; nevertheless they should be of interest for the tropospheric profiling community because the results of the paper shows, for the particular settings used in the paper, the impact of the tropospheric profiling on the short-term wind forecast in Central Europe.

Review AMTD

P2r10: law->laws

toward - > towards analyses -> analysis

Frequently mixed up singular and plural !

I will carefully review the English of the paper to solve this issue.

3587 9 : innovations : are observation increments, you probably mean analysis increments here : specify

Right, I mean analysis increments. In the revised version of the paper I will write:

"The recursive filter performs the task of convolving a spatial distribution of the analysis increments with a smoothing kernel, which is the covariance function of the background error."

3588 : correlation between u-v must be respected see e.g. Daley, Hollingsworht, Undén....

In the revised version of the paper, the correlation between the zonal and meridional wind components is taken into account.

It should be emphasized that most of the research in preparing the revised version of the paper, was to better define the U_v transform. In doing that, several changes were introduced into the analysis package, both in the NMC method and in the 3D-Var scheme. So, the transform U_v that is used in the revised version of the paper is different (and better performing) compared to the discussion paper. This, of, course changed the results of the paper. All the simulations were redone and the new statistics are presented.

At the same time the paper will be rewritten for the part discussing the U_v transform and should be much clearer and detailed compared to the discussion paper. Indeed, as noticed by the reviewer, see also two comments below, the discussion paper was rather poor for the U_v transform. To avoid an excessive paper length, the Appendix B, reported at the end of this document, will be added to detail the practical computation of the U_v transform.

The comment that will be included in section 2 is the following:

"The vertical transform U_v is given by an empirical orthogonal function (EOF) decomposition of the vertical component of the background error correlation matrix ($\boldsymbol{C_z}$). To determine $\boldsymbol{C_z}$, the NMC method is applied, by averaging both is space (in longitude and latitude) and time, the difference between 24-h and 12-h forecasts valid at the same time (see the Appendix B for the details).
The cross-correlation between different variables is considered only between the two wind components. More in detail, the $\boldsymbol{C_z}$ matrix is a block matrix, where each block contains the vertical correlations between variables errors averaged in space and time:

$$C_z = \begin{pmatrix} r(T,T) & (0) & (0) & (0) \\ (0) & r(ur,ur) & (0) & (0) \\ (0) & (0) & r(u,u) & r(v,u) \\ (0) & (0) & r(u,v) & r(v,v) \end{pmatrix} \qquad (5)$$

In Eqn. (5), *r(var1,var2)* is a square-matrix whose dimensions are equal to the number of levels of the analysis grid (29, Table 2), containing the vertical error correlations between the variables *var1* and *var2*; *T, ur, u* and *v* are the temperature, relative humidity, the zonal and meridional wind components, respectively.
The $\boldsymbol{C_z}$ matrix is symmetric and positive-defined and can be decomposed in the eigenvalues and eigenvector matrices, i.e. $\boldsymbol{C_z=ELE^T}$, where $\boldsymbol{E}$ is the eigenvectors and $\boldsymbol{L}$ the eigenvalues matrix. Using this decomposition, the vertical transform U_v is written as $U_v\boldsymbol{=EL^{1/2}}$.
The above setting was derived after several numerical experiments using different settings of the U_v transform. In particular, considering the whole vertical component of the background error correlation matrix ($\boldsymbol{C_z}$) had worse performance compared to the setting used in this paper.
This issue is likely caused by the need of considering the local variation of the vertical component of the background error correlation matrix (Barker et al., 2003b). Moreover, forecast differences, used in

the NMC method, are only a proxy for the structure of forecast errors. These points will be further investigated in future studies."

So, considering also the Appendix B, the U_v transform should be defined in detail.

It is interesting to compare the statics obtained in the discussion paper with those of the revised paper. First let me introduce the statistics used.

The RMSE computed between the background run (i.e. the forecast not starting from the analyses produced by the analysis package presented in the paper) and the observations at a fixed time and for the whole period is referred as the background error (RMSE_b). Similarly, the RMSE computed between the forecast run (i.e. the forecast starting from the analyses produced by the analysis package presented in the paper) and the observations at a fixed time and for the whole period is referred as the forecast error (RMSE_f). For the computation of both RMSEs, the grid point nearest to each observation is considered and the statistics are computed for the whole domain.

The following Table shows the difference between RMSE_b and RMSE_f, averaged over the levels, for the wind components (m/s) for the results of the discussion paper and for the revised version. A positive difference means that the short-term forecast has a lower error than the background and the wind forecast is effectively improved by using the analyses as initial conditions. Each cell of the Table has a pair of values in the following order: (zonal wind component, meridional wind component).

TIME	RMSE_b-RMSE_f (discussion paper), m/s	RMSE_b-RMSE_f (revised paper), m/s
ANL	(1.00, 1.03)	(1.14, 0.94)
01	(0.35,0.45)	(0.41, 0.47)
02	(0.15, 0.16)	(0.17, 0.26)
03	(0.04, 0.02)	(0.06, 0.08)

It is noticed that the performance of the revised version is always better than the discussion paper, except for the analysis of the meridional wind component. Interestingly, the performance for the meridional wind component is improved for the one-, two- and three-hours forecasts, showing that the analyses filters more the data but are more balanced compared to the discussion paper.

3589 15 : T should here be the unbalanced T, since there is a large balanced component related to the geopotential Z. I believe this is a serious omission. It will work anyway but it is not correct.

Corrected (I assume T is ξ). In the revised version of the paper I will write:

"The geopotential increment is determined by the geostrophic equilibrium in pressure coordinates (the analysis system uses pressure as vertical coordinate):

$$\nabla_p^{\,2} Z' = \frac{f\xi'}{g} \qquad (3)$$

where ξ' is the vertical component of the perturbed relative potential vorticity computed from the increments of the zonal (u') and meridional (v') wind components, g is the gravity (m/s^2) and f is the Coriolis parameter (s^{-1})."

3590 the background error should not need to be specified as simply as that: It can be derived from the NMC method.

In the revised version of the paper a deeper discussion of the vertical transform U_v will be given into the text (see two comments above and Appendix B), and the background error should be clearly specified. Moreover, after the discussion about the observation and background errors (σ_b, σ_o) I will write:

"Considering the σ_b, U_p, U_v, and U_h , defined above, the $\boldsymbol{x'} = \boldsymbol{U\chi}$ transformation can be better explained as:

$$x' = U\chi = U_p\rho \qquad (4)$$

and:

$$\rho = U_v U_h \chi = U_v \sigma_b \hat{F} \chi \qquad (8)$$

where $\hat{F}$ denotes the application of the recursive filter to the control variable χ and the U_p transform has been separated from the others because it is applied after the minimization of the cost-function."

3592 QC check is very ad-hoc and no science behind it (no flow dependency, level, latitude etc., no cross check)

Considering the Quality Check of the observations, I agree with the reviewer that the one adopted into the paper is crude and was used to avoid introducing large unbalances in the analysed fields. Nevertheless, the adoption of a more sophisticated QC should require a larger dataset, i. e. the study of a longer period, compared to that reported in this paper. This is somewhat behind the purpose of this paper and will be considered in future works.

In the revised version of the paper I will include the following sentence after the description of the QC (Section 3) to stress this point.

"It is important to highlight that the QC adopted in this paper is crude and is used to avoid inserting large unbalances into the analysed fields. Future implementation of the data assimilation system will consider more sophisticated QCs, which, however, must be based on a number of simulations larger than that used in this paper, to compute reliable statistics of the innovation distributions. This is out of the scope of this paper."

3606 Fig. 4 b. Very strange increments. Looks like only 3 observations used. Not credible.

The problems with the understanding of Figure 4b were of two kinds: the figure didn't show the position of the observations, and the vertical impact of an innovation was limited inside 2500 m from the observation height. This filter of 2500 m was not applied in the other parts of the discussion paper, but, wrongly, its application was not stated into the text when discussing the Figure 4b.

To make the Figure 4b understandable, I will redo the Figure 4b to include the positions of the observational systems entering into the analysis (this issue was also raised by the reviewer #1), and by considering the U_v transform adopted in the paper. Moreover, to consider a level for which more data are available, I will show the analysis for the zonal wind component at 850 hPa for the first July at 12 UTC.

The new Figure 4 will be:

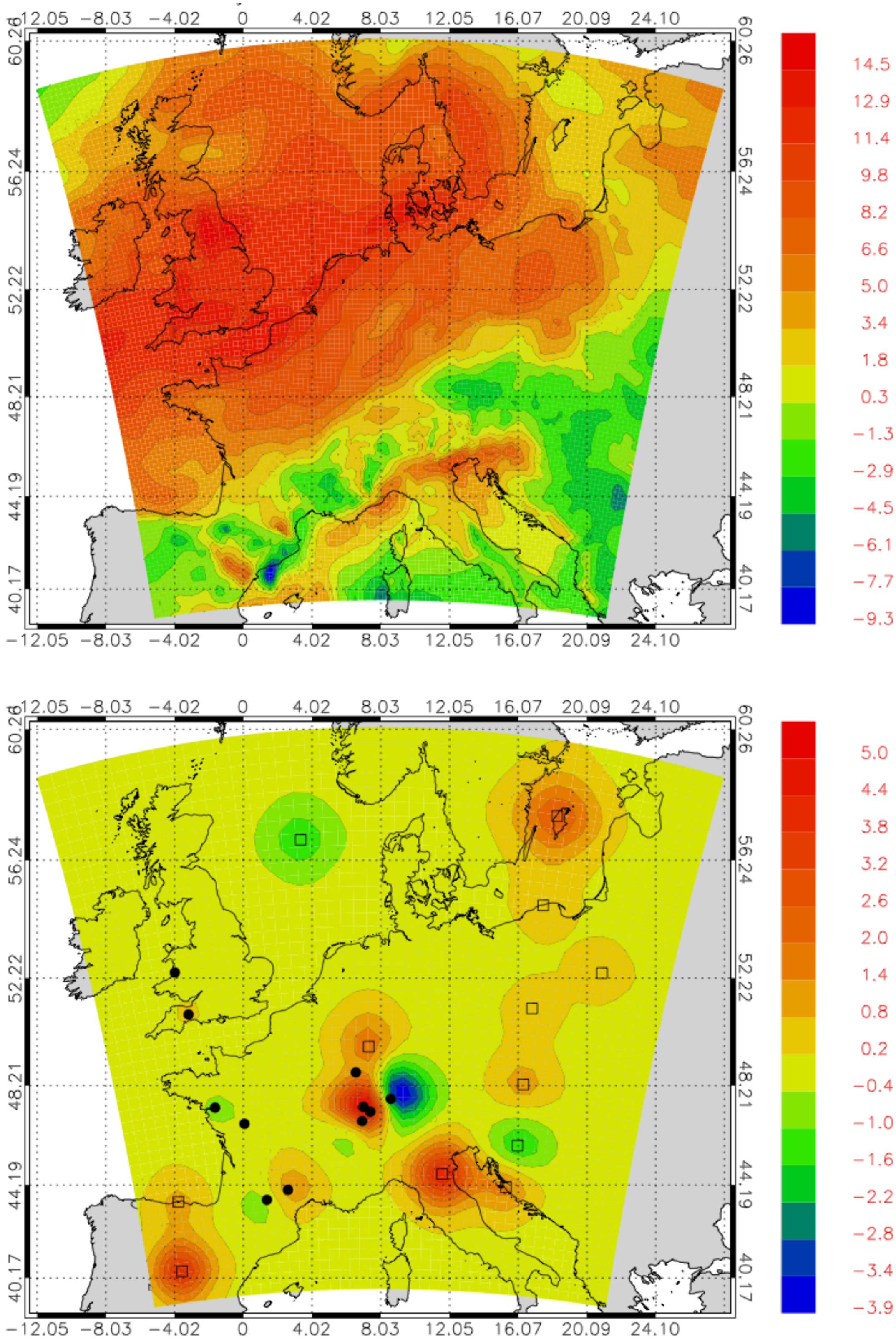


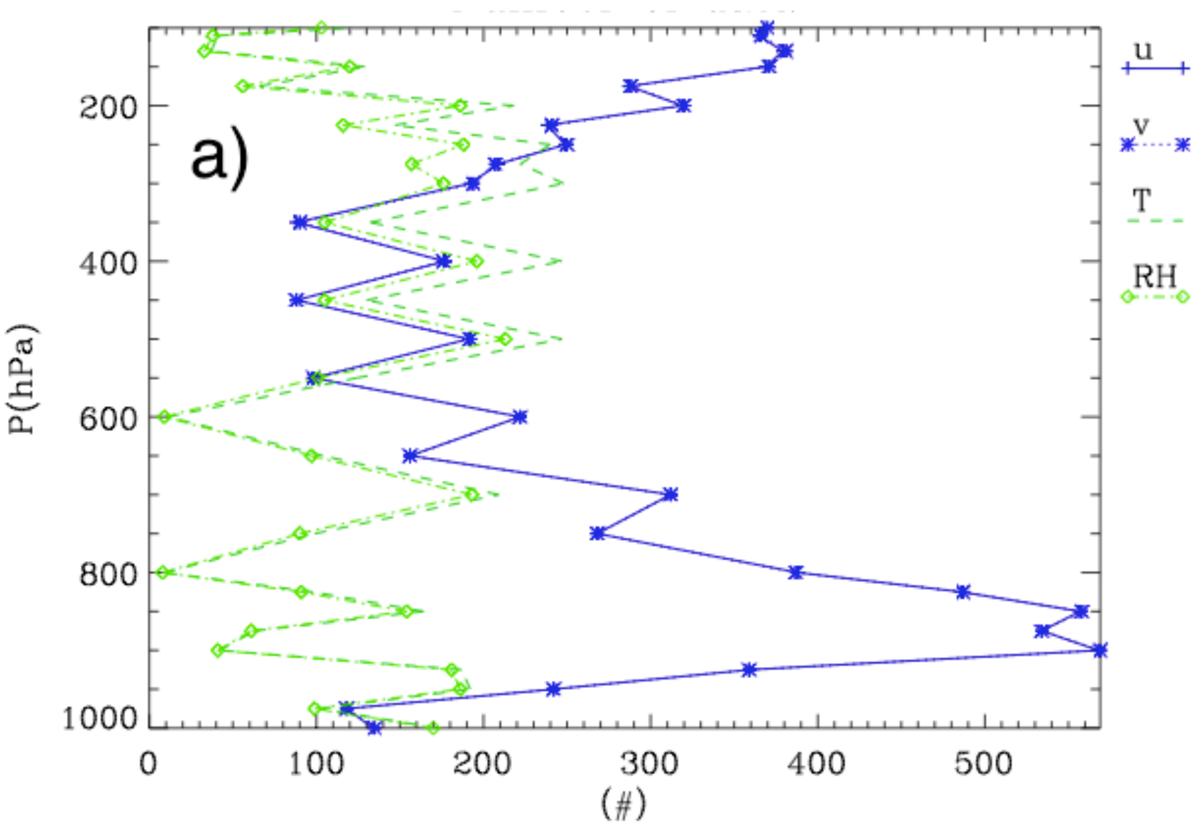
Figure 4: a) Background of the zonal wind component (m/s) at 850 hPa at 12 UTC on 01 July 2012; b) analysis increments (m/s) at the same time and level of a). The positions of the radiosoundings (open squares) and of the wind profilers (filled circles) used in the analysis are shown. The figure shows the horizontal domain used in this paper.

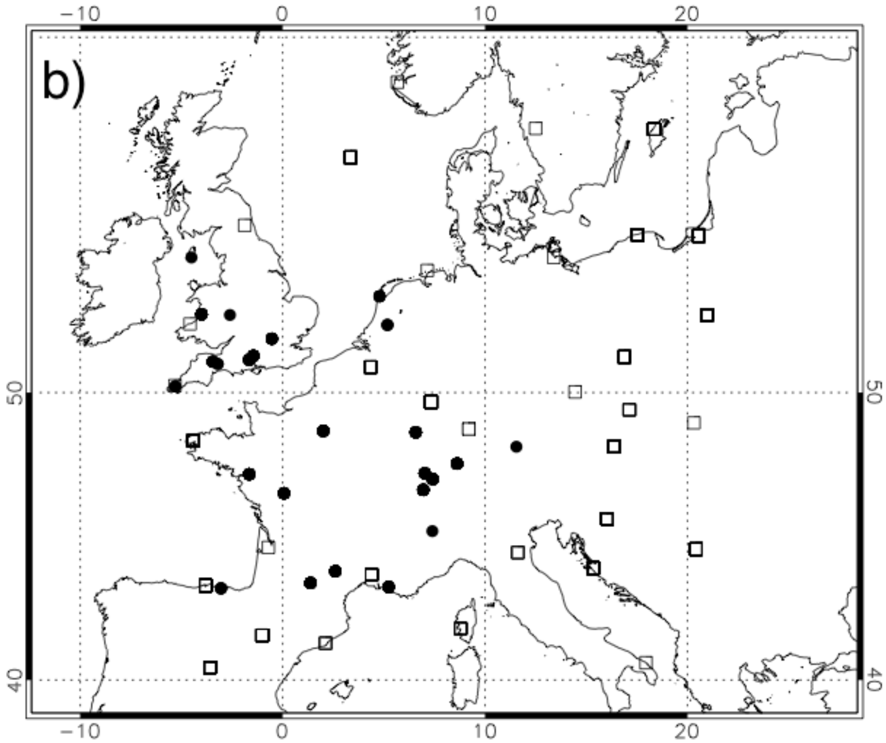
The comment to the Figure will be:

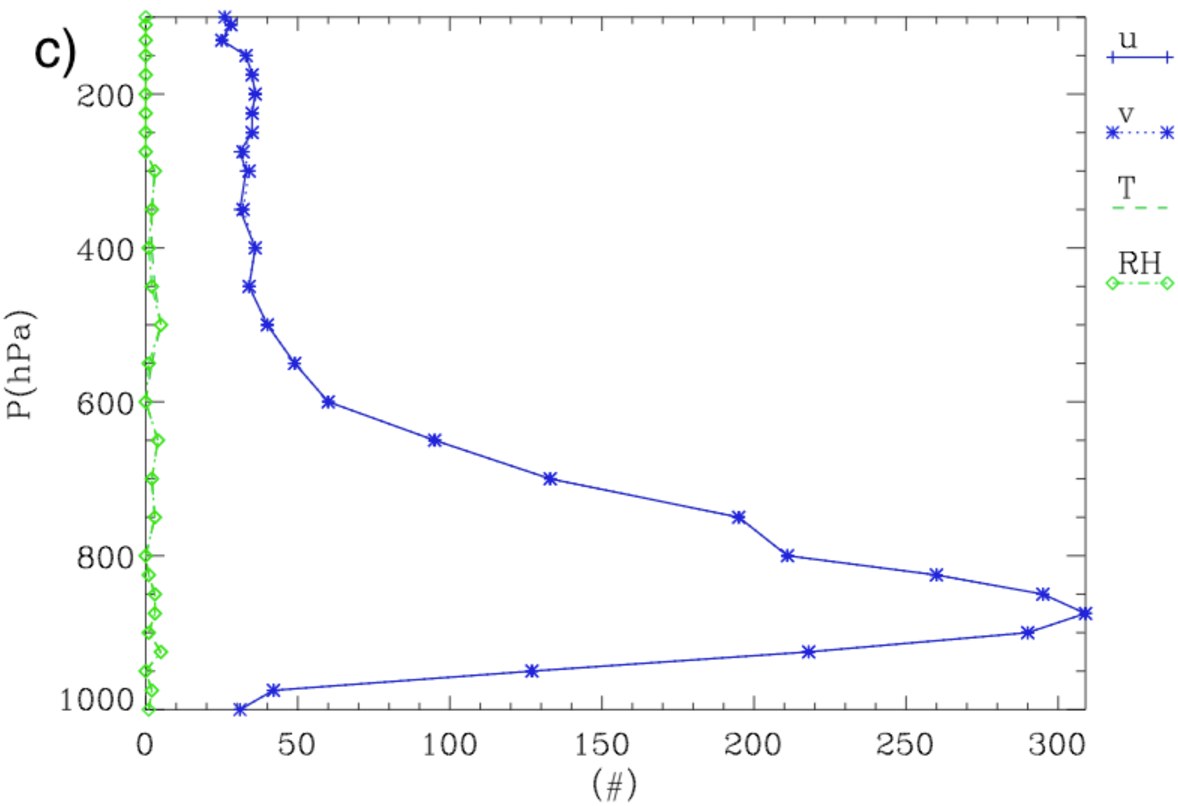
“Figure 4 shows, for example, the analysis for the zonal wind component at 850 hPa at 12 UTC on 1 July. The background has 10 km horizontal resolution and covers the central Europe. Its grid setting is shown in Table 2. The analysis increments are given on the same grid as the background but with halved horizontal resolution (20 km) to speed-up the analysis computation. Figure 4b, in particular, shows how the interaction among several observations impacts the analysis increments, as over Central Europe; it also shows the effects of isolated observations on the analysed field, as over the North Sea.“

3608-10 Fig. 6 - 8. None of the figures are legible. I cannot see any numbers, the font size is 5 times smaller than the text! If published need to be remade.

All Figures 6-8 will be remade to make the numbers readable. Moreover, the results of Figures 7-8 changed because of the different settings of the analysis system adopted in the revised version of the paper. It is important to highlight that Figure 6a of the first submission paper had a wrong cut-off of the data after 21 days. This resulted in less data number displayed in the old Figure 6a, compared to the new one. This change, however, does not affect the discussion that was made in the paper (sections 3 and 4) about Figure 6a. The new Figures are shown below:




b)
−10
0
10
20
50
40


c)
P(hPa)
200
400
600
800
1000
0
50
100
150
200
250
300
(#)
u
v
T
RH

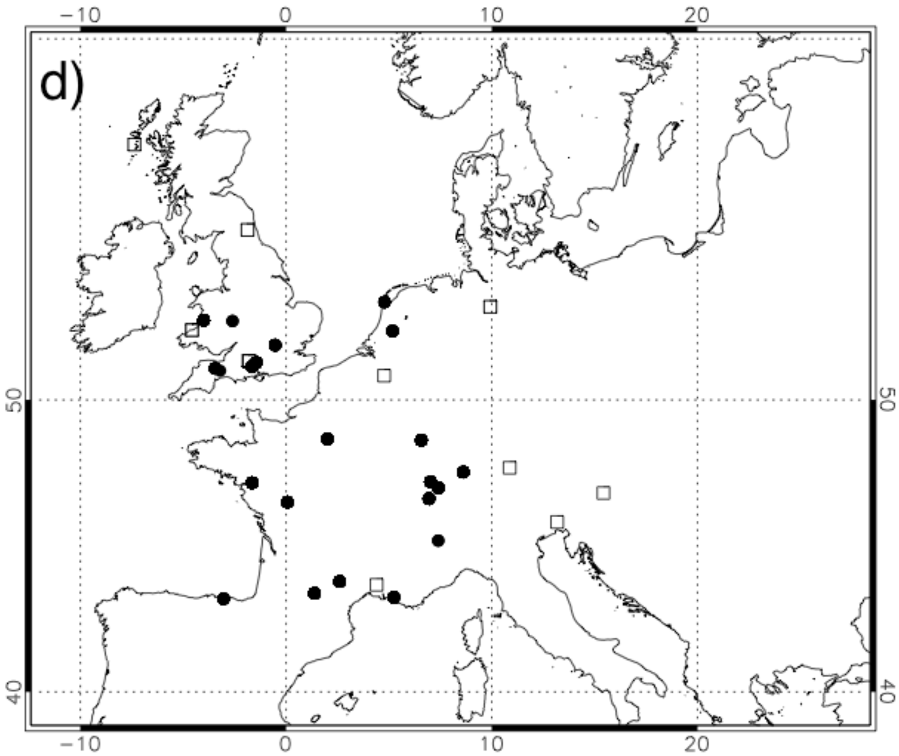
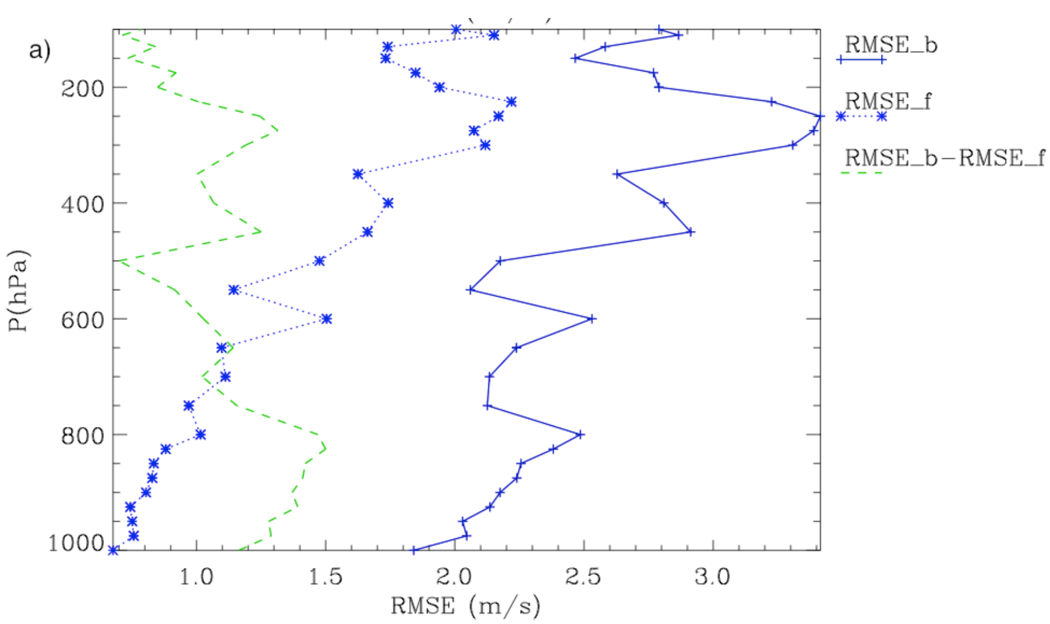


Figure 6: a) The number of data available at the analysis time (12 UTC) accumulated for the whole period and over the whole domain. *T* is for temperature, *RH* is for relative humidity, *u* and *v* are for the zonal and meridional wind components, respectively. The number of data for the wind components (*u, v*) is the same for all levels; b) positions of the radiosoundings (open squares) and radar wind profilers (filled circles) at 12 UTC considering the whole period. Not all radiosoundings and radar wind profilers are reporting data at a specific analysis time; c) as in a) for 13 UTC; d) positions of the radiosoundings (open squares) and radar wind profilers (filled circles) at forecasting times (13, 14 and 15 UTC). Not all radiosoundings and radar wind profilers are reporting data at a specific forecasting time.



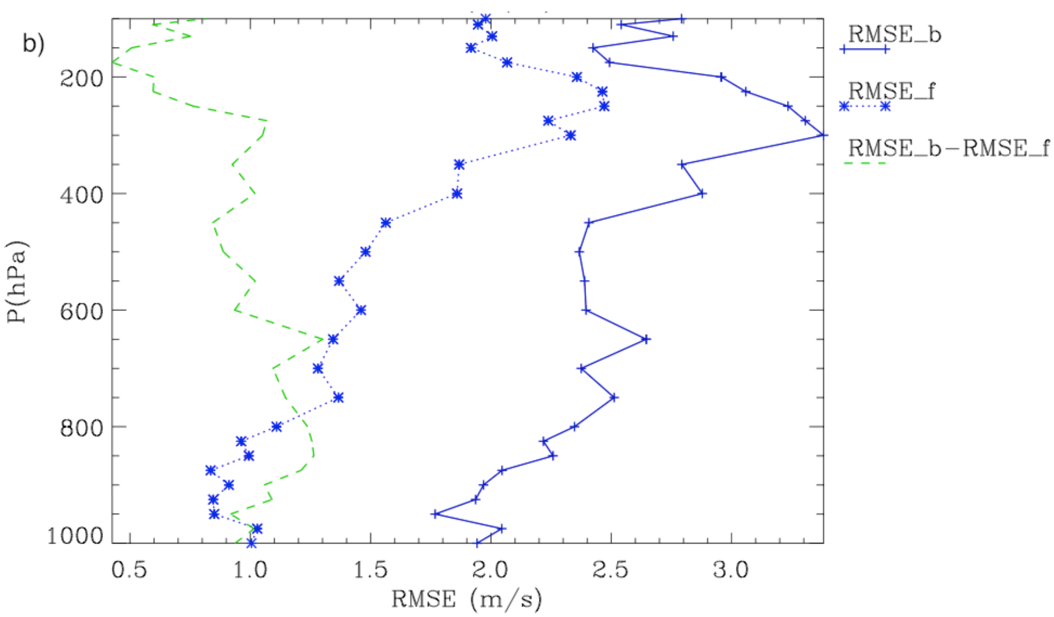
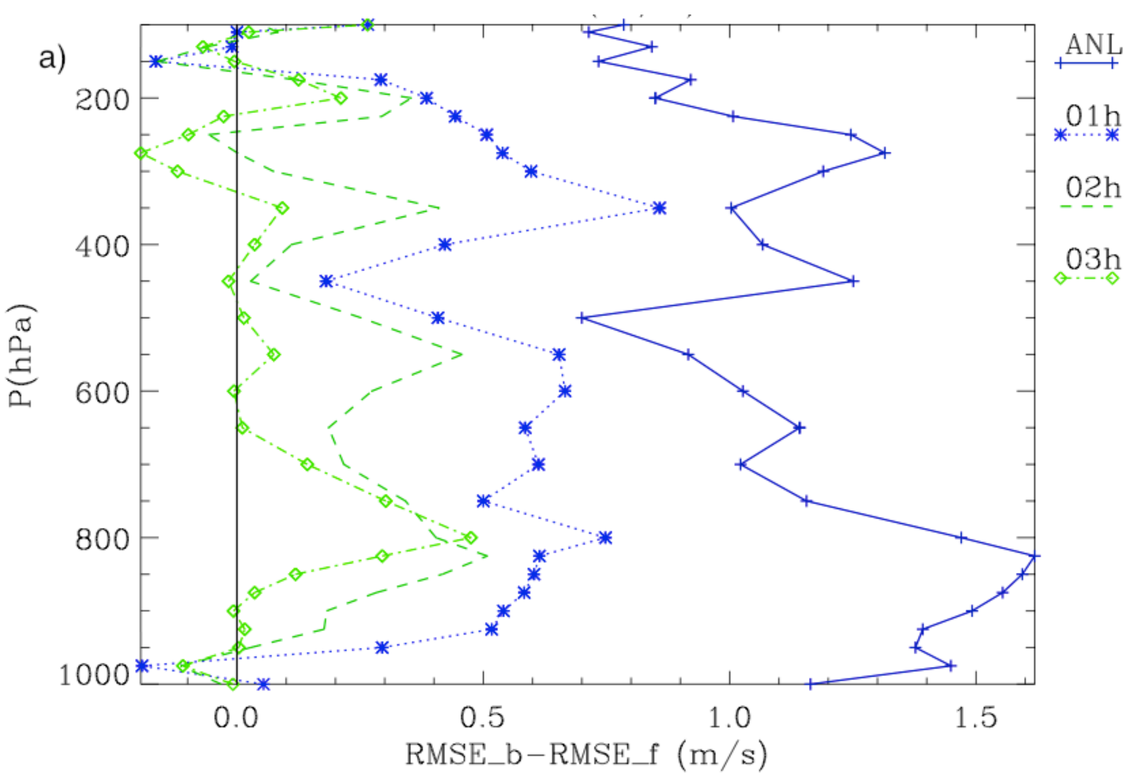


Figure 6: RMSE of the background field (RMSE_b), of the analyses (RMSE_f), and their difference (RMSE_b-RMSE_f) for: a) zonal wind component; b) meridional wind component. The RMSEs are computed for the whole period considering the grid-points nearest to the observations. The ANL statistics are computed after the RAMS model has been initialized by the analyses.



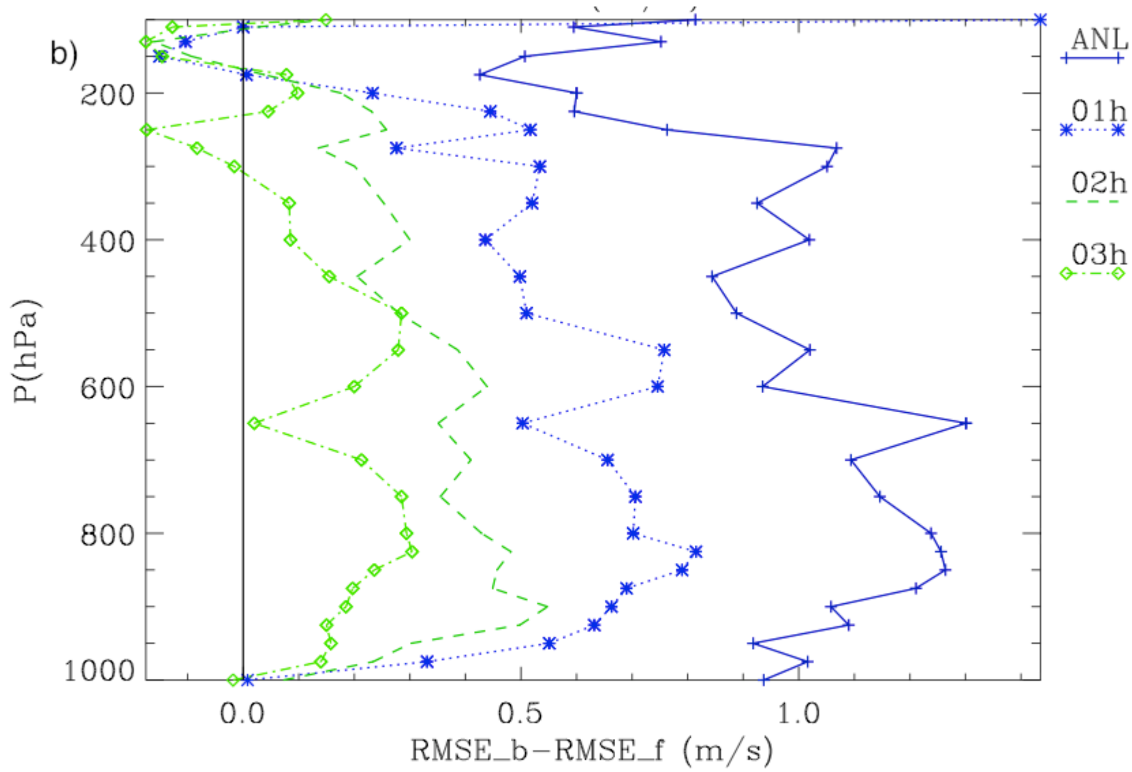


Figure 7: Differences between RMSE_b and RMSE_f for the analysis time (ANL), one- (01h), two- (02h), and three-hours (03h) forecast for the: a) zonal wind component; b) meridional wind component. The RMSEs are computed for the whole period considering the grid-points nearest to the observations. The analysis time is shown to better understand the behaviour of the performance with time.

It is also noticed that the Figure 3 of the discussion paper will be changed to match the settings of the analysis system used into the paper. The new Figure is:

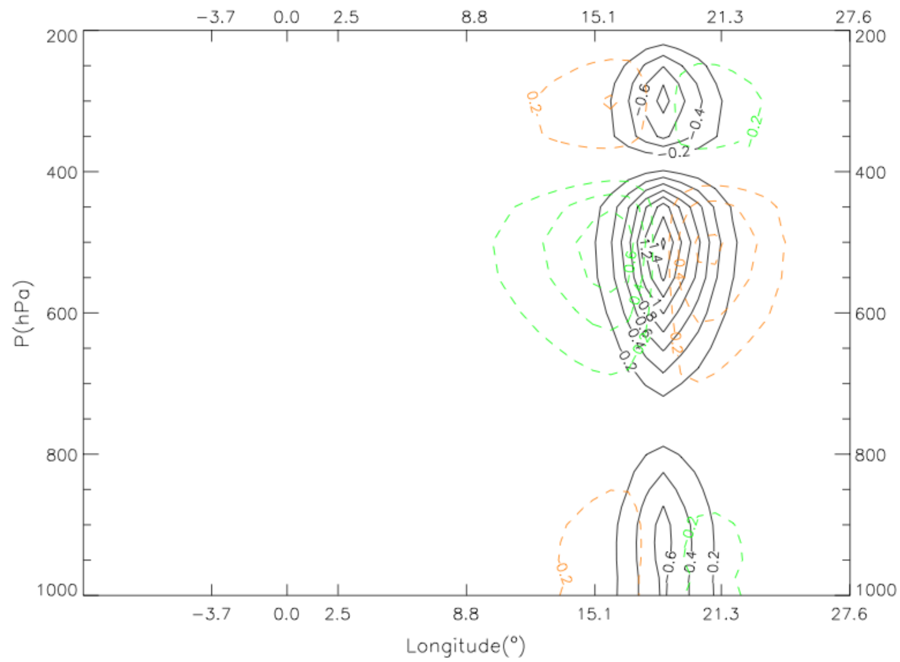


Fig. 3: The effect of the U_h , U_v and U_p transforms (see text for details). Solid lines are contours of meridional wind increments (*v'*, contours from 0. to 1.6 m/s with 0.2 interval); dashed lines are the geopotential height increments *Z'* (contours from -0.6 to 0.6 m with 0.2 m interval).

The comment to Figure 3 is:

Figure 3 shows the combined effect of the U_p, U_v and U_h transforms. It is shown the longitude-height cross section at 57.5 N latitude for the meridional velocity increments and for the geopotential height increments determined by a single meridional wind component observation innovation of 2.5 m/s, introduced over the Gotland Island ~(57.5 N, 18 E) at 500 hPa. The final increment is spread vertically by the U_v transform and horizontally by U_h. The U_p transform determines the increments of the geopotential height. It is worth noticing that a positive innovation of the meridional wind component at 500 hPa causes negative increments of the same variable in the upper and lower troposphere. This is caused by the negative correlation between the vertical errors at 500 hPa and those in the upper and lower troposphere for the meridional wind component, as shown in Appendix B.

Finally the Figure 2 of the discussion paper is removed in the revised version of the paper, because part of the discussion of the vertical background error correlation matrix, which is enlarged compared to the discussion paper, is done in Appendix B and the new "Figure 2" is put in this Appendix.

Appendix B

In this Appendix it is shown how the vertical component of the background error matrix $\boldsymbol{C}_z$ is derived. The methodology to compute $\boldsymbol{C}_z$ is detailed in the following points:

1) The difference between two short-term forecasts verifying at the same time is firstly computed ***x'****(i,j,k,t)*=$\boldsymbol{x}_{T1}$*(i,j,k,t)*-$\boldsymbol{x}_{T2}$*(i,j,k,t),* where *T1=12h* and *T2=24h*, and *i, j, k, t* show the dependence of ***x'*** on the three spatial dimensions and time. In this paper the whole month of July 2012 was considered and the differences ***x'****(i,j,k,t)* where computed between two short-term forecasts verifying at 00 UTC on each day;

2) For each vertical level, the average $\overline{x}(k,t)$ and the standard deviation $\boldsymbol{\sigma_x}(k,t)$ are computed to define standardized anomalies:

$$v'(i,j,k,t) = \frac{x'(i,j,k,t) - \overline{x}(k,t)}{\sigma_x(k,t)} \qquad \text{(B.1)}$$

3) The domain averaged vertical background error correlation matrix ***C****(k,k',t)* is derived for each day as:

$$C(k,k',t) = \sum_{i=1,I;\,j=1,J} \frac{v'(i,j,k,t)\,v'(i,j,k',t)}{IJ} \qquad \text{(B.2)}$$

4) The matrices ***C****(k,k',t)* are averaged in time to get the vertical component of the background correlation matrix $\boldsymbol{C_z}$:

$$C_z = C_z(k,k') = \sum_{t=1,T} \frac{C(k,k',t)}{T} \qquad \text{(B.3)}$$

For each variables pair, the $\boldsymbol{C_z}$ matrix is symmetric and positive-defined and can be decomposed in the eigenvalues and eigenvector matrices, i.e. $\boldsymbol{C_z=ELE^T}$, where ***E*** is the eigenvectors and ***L*** the eigenvalues matrix. Using this decomposition, the vertical transform U_v is written as $U_v=\boldsymbol{EL}^{1/2}$.
Using the whole $\boldsymbol{C_z}$ to define the vertical transform U_v, however, had a negative impact on the data assimilation and forecasting system. The reason for this negative impact is likely caused by the need of considering the local characteristics of the background error correlation matrix (Barker et al., 2003b). This problem is under investigation to better define the U_v transform.
Figure B.1 shows the vertical correlation profile of the meridional wind component for three levels, namely 900, 500 , and 250 hPa.

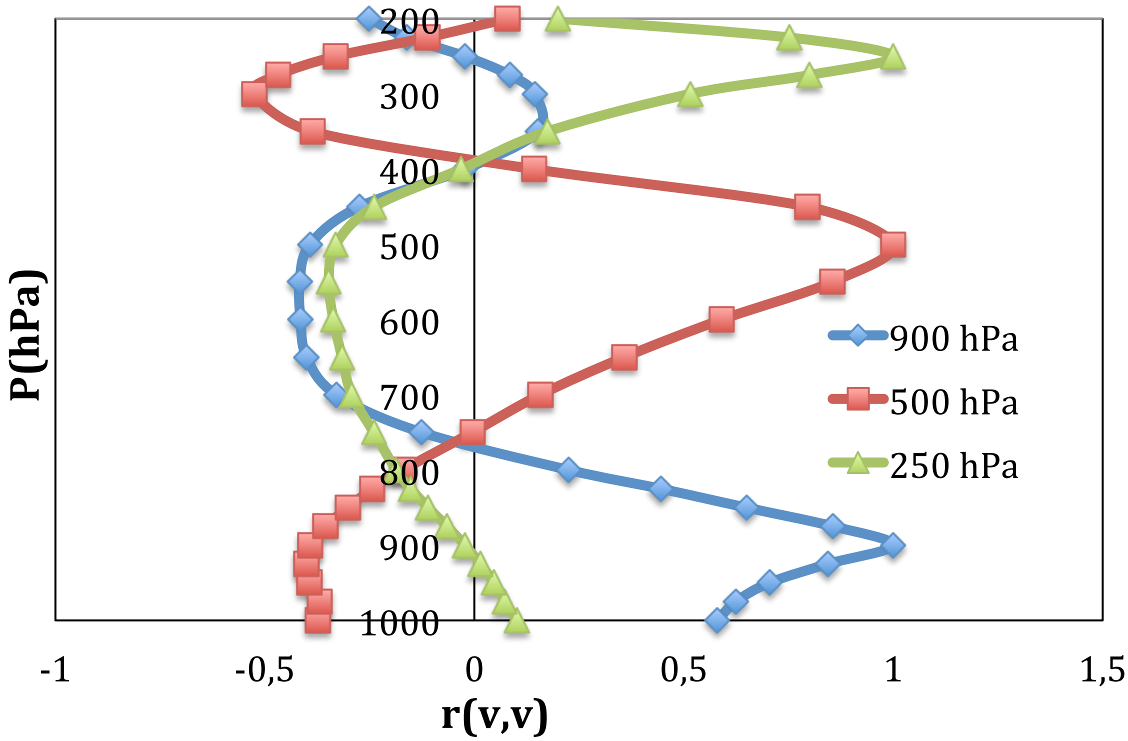


Figure B.1: Vertical error correlation function of the meridional wind component at 900 (diamonds), 500 (squares) and 200 (triangles) hPa.

The vertical error correlation shows a broader peak at 500 hPa compared to that in the upper and lower troposphere. Moreover, a positive increment of the meridional wind component at 500 hPa causes negative increments in the lower and upper troposphere, as shown by the negative vertical error correlations between these levels. This behaviour of the vertical error correlation is clearly shown in Figure 3.

Finally, Figure B.2 shows the cross-correlation between the meridional wind component at 900, 500 and 250 hPa, and the zonal wind component. There is a decrease of the correlation values compared to Figure B.1, as expected, and there is a rather complex interaction between the two wind components as a function of the level.

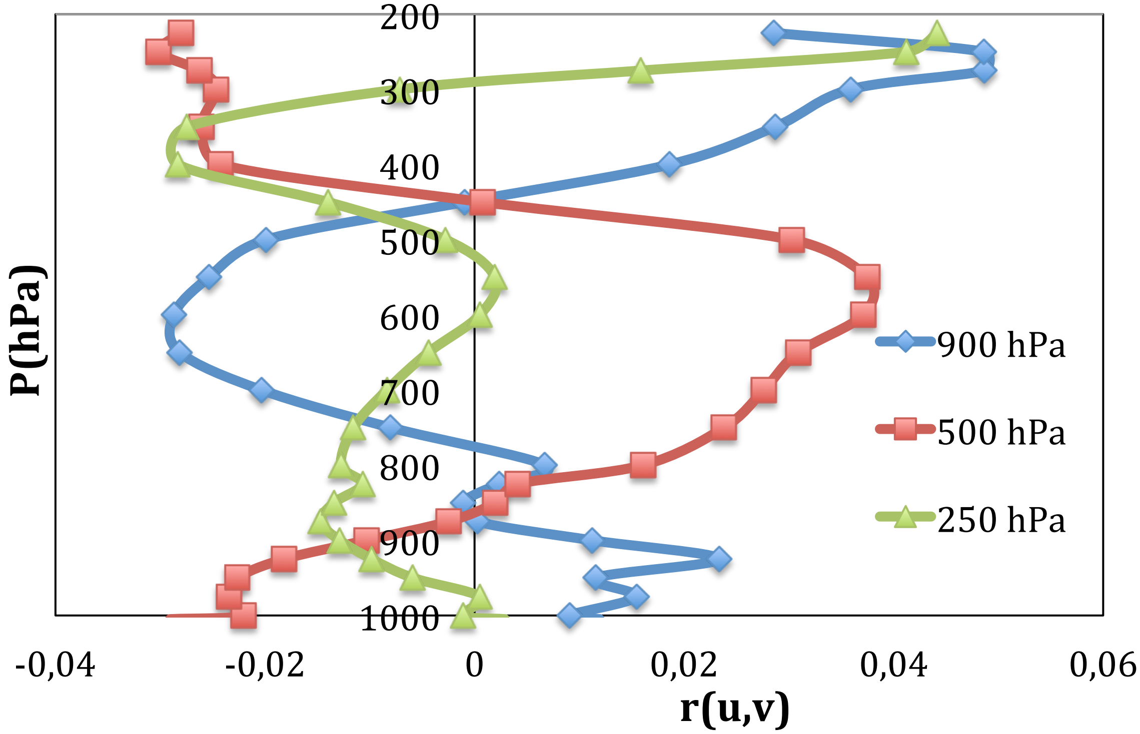


Figure B.2: Vertical error correlation function of the meridional wind component at 900 (diamonds), 500 (squares) and 200 (triangles) hPa and the zonal wind component.