TECHNICAL MEMORANDUM

510

Observing system impact assessment using a data assimilation ensemble technique: Application to the ADM-Aeolus wind profiling mission

David G.H. Tan, Erik Andersson, Michael Fisher and Lars Isaksen

Research Department

To appear in Q. J. R. Meteorol. Soc.

January 2007

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

Abstract

Ensembles of parallel 4D-Var data assimilation cycles have been used to assess the impact of two observing systems: (1) the exisiting network of radiosonde and wind profilers and (2) the future spaceborne ADM-Aeolus wind-profiling lidar. We demonstrate that this new technique for impact assessment provides a practical alternative to the traditional observing system simulation experiments (OSSE), with the particular advantage that real existing observations are assimilated exactly as in operational practise, and do not need to be simulated artificially. It is only the future observing system(s) under test (ADM-Aeolus in our case) that are generated through simulation. Unlike OSSEs, there is also no need to generate an artificial reference atmosphere ('proxy truth' or 'Nature Run'), and the problems normally associated with identical-twin experiments are thus avoided. Based on detailed simulation of the ADM-Aeolus wind-measuring capabilities and expected data quality, our results show that ADM-Aeolus will provide benefit comparable to the radiosonde and wind-profiler network, with analysis impact particularly over ocean, and in the tropics. The impact is retained to the medium range (e.g. day 5) of forecast. Our results for radiosonde and wind-profiler impact qualitatively agree with those obtained with the well-established observing system experiment (OSE) technique, which gives reason for some confidence in the usefulness of the ensemble-based technique for impact assessment.

Contents

1 Introduction

Data assimilation systems provide estimates of the atmosphere or ocean state (the analysis). There is also a need to provide estimates of the analysis uncertainty, which amongst several factors depends on the accuracy and coverage of the observing systems. Ensemble data assimilation methods naturally provide estimates of analysis uncertainty based on the spread between members of the ensemble of analyses (e.g. the ensemble Kalman filter, Houtekamer and Mitchell 2001; Houtekamer *et al.* 2005). However, most operational numerical weather prediction (NWP) centres currently produce their analyses using variational data assimilation techniques (3D-Var or 4D-Var schemes), for which analysis uncertainty is not so readily available. Here we will use a combined 4D-Var and ensemble approach to generate estimates of 4D-Var analysis and forecast uncertainty, with varying observing systems. The combined approach was proposed and used by Fisher (2003a) for calibration of background error covariances (see Žagar *et al.* 2005 for explanation and justification). In these ensembles, the assimilations use the static backgound error specifications of standard 4D-Var. No use is made of the ensemble's ability to cycle covariance information (as would be the case with an ensemble Kalman filter). Its purpose is purely diagnostic, associating ensemble spread with analysis uncertainty of the 4D-Var.

We investigate whether the difference between spreads in two different data assimilation ensembles, one with and one without the observing system of interest, can usefully be interpreted as a measure of impact of that observing system. We apply the technique within the framework of the operational 4D-Var system of the European Centre for Medium-Range Weather Forecasts (ECMWF): first to the existing radiosonde and wind-profile network, and second to simulated space-based wind profile data of ESA's future ADM-Aeolus mission. The radiosonde and wind-profile experiment is performed to enable a qualitative comparison with established observing system experiment (OSE, Dumelow 2003) results by Kelly *et al.* (2004). Qualitative agreement with radiosonde data denial OSE results will give reason for confidence in assessing the potential impact of ADM-Aeolus.

We present our method as a practical alternative to the established OSSE technique. In OSSEs (Arnold and Dey, 1986; Atlas 1997), an estimated absolute observation impact is obtained through comparison with a simulated, and therefore perfectly known, reference atmosphere (the proxy truth, or Nature Run, NR). From the NR, simulated 'true' observations are generated (Becker *et al.* 1996) by interpolation to actual or predicted observation locations, followed by the application of instrument-specific observation models (variously called forward models, or observation operators). Examples of such forward models are the radiative transfer models that simulate satellite radiance measurements, given the NR. Another example is the LIPAS forward model (Marseille and Stoffelen, 2003) for the Aeolus instrument used here. In OSSEs, the simulated 'true' observations are perturbed by adding realistic observation errors, and assimilated with an advanced analysis/forecast system (AFS) (e.g. Masutani *et al.* 2004). The observation impact is judged by how closely AFS assimilations with varying observing systems approximate the NR. If the same AFS is used in the assimilation as was used for the NR, the experiment is said to be of the 'identical twin' type. To obtain reliable OSSE results, however, it is required that two different AFSs are used, and that the difference between them is similar to the difference between the AFSs and the real atmosphere. If the two AFSs are more similar to one another than either AFS is similar to the real atmosphere, we have a 'fraternal twin' experiment (Arnold and Dey, 1986, Stoffelen *et al.* 2006), in which the simulated observation impact is likely to be overestimated. In contrast to OSSEs, our ensemble method does not rely on comparison with a NR, and cannot therefore be classified as either an identical twin, or fraternal twin experiement, and does not therefore suffer from the same problems and shortcomings. On the other hand, to obtain a reliable estimate of the abolute observation impact, the ensemble method requires careful calibration of its spread/skill relationship. For relative comparison of two observing system impacts (e.g. radiosonde vs. ADM-Aeolus), the absolute calibration becomes unimportant, and can be seen as a straightforward rescaling.

The Atmospheric Dynamics Mission ADM-Aeolus is the second of ESA's Earth Explorer Core Missions (ESA

1999; Stoffelen *et al.* 2005a). ADM-Aeolus is scheduled for launch in September 2008 and has a projected lifetime of three years. Its objective is to demonstrate the capability to measure wind profiles from space using a Doppler Wind Lidar (DWL). The DWL will provide layer-averaged wind measurements with a 1000 m vertical resolution through most of the atmosphere (i.e. from 2 to 16 km), 500 m below 2 km, and 2000 m between 16 and 26 km (Fig. 1). The schematic in Fig. 1 shows the instrument viewing from a low-altitude ($\sim$ 400 km) polar orbit in the direction perpendicular to the satellite track. There is information on the horizontal line-of-sight (HLOS) wind component only (close to east-west except at high latitudes), and the unobserved wind component and the mass field will have to be statistically inferred within the data assimilation process (Žagar 2004). It is therefore considered essential to investigate prior to launch whether single-component ADM-Aeolus data can provide the expected benefit to NWP (Riishøjgaard *et al.* 2004, with comments by Stoffelen *et al.* 2005b), through simulations. The instrument will provide 50 km along-track average winds, separated by 200 km (Fig. 1); this is to ensure minimal error correlation between consecutive measurements (Stoffelen *et al.* 2005a). The accuracy of the ADM-Aeolus wind measurements will depend primarily on the intensity of the backscattered laser light, which in turn depends on the presence and thickness of clouds, and the concentration of aerosol (Marseille and Stoffelen, 2003). Good quality ADM HLOS wind retrievals are expected in the layers of clear air above clouds, from cloud-top layers, from layers in and below thin clouds, and from layers with sufficient aerosol in the lower parts of the atmosphere, which has been confirmed through detailed simulation (Tan and Andersson 2005). The ADM impact assessment presented here relies on similar detailed simulation of the ADM wind-profile yield and accuracy based on ECMWF model cloud and climatological aerosol (Vaughan *et al.* 1995) distributions.

In Section 2 we describe the data assimilation ensemble technique for observing system impact assessment, and give the outline of the experimentation. Results are presented in Section3, with discussion and conclusions in Section 4.

2 Description of method and experimentation

2.1 Ensembles of data assimilations

Ensembles are used in a variety of contexts to provide estimates of uncertainty from the spread between ensemble members. In particular, the ensemble Kalman filter (Evensen 1994; Houtekamer and Mitchell, 1998) has been the focus of much recent research. The approach of Evensen and van Leeuwen (1996), Houtekamer and Mitchell (1998; 2001), Hamill and Snyder (2002) and Keppenne and Reinecker (2002) is to perform separate analyses for each member of the ensemble. The analyses differ because random perturbations are added to the observations, and because each member provides a different background field. This is an approach that can be applied to any analysis method (such as 4D-Var, in our case). What would turn the ensemble of analyses into an ensemble Kalman Filter would be the use, by each analysis, of a background error covariance matrix constructed from the ensemble. Instead, we have used the standard 4D-Var method with a specified background error covariance matrix (Derber and Bouttier, 1999; Fisher 2003a) that does not depend on the ensemble. The 4D-Var ensemble is run with perturbed observations but is otherwise distinctly different from an ensemble Kalman filter. The error growth between analysis times is not assumed to be linear. The resulting spread between ensemble members can be interpreted as an estimate of 4D-Var analysis (and forecast) uncertainty.

The utility of ensembles of data assimilations to assess the uncertainty in short-range forecasts (for calibration of background-error covariances) has been demonstrated by Fisher (2003a) and Žagar *et al.* (2005) where a detailed description of the technique is given. Tan and Andersson (2005) used the method to assess the uncertainty in ECMWF wind analyses (their Fig.4). Here we extend the approach to assess analysis and forecast impacts of observing systems. This is shown schematically in Fig. 2, in which time is running from left to right and each row represents an independent, cycling 4D-Var assimilation ensemble member with its analysis and forecast steps shown as boxes. The blown-up portion of the figure shows that each input of the analysis is perturbed: the background $\mathbf{x^b}$, the observations $\mathbf{y}$, and the surface fields SST, with errors ε^b, ε^o and ε^{SST}, respectively. Provided the perturbations are drawn from distributions with the statistical characterists of the true random errors of the inputs then the output perturbations, ε^a and ε^f, will have the statistical characteristics of analysis and forecast error, respectively. Evolved surface fields such as soil moisture and land surface temperature that are computed outputs of the model also become perturbed. Constant surface fields (e.g. surface roughness and vegetation parameters) that are provided as input to the model have however not been perturbed in these experiments. Note that only the first background state has to be explicitly perturbed. Subsequent analyses use the output forecast from the preceding cycle. The members of the ensemble differ because the input perturbations are a different random draw (from the relevant distribution) for each member of the ensemble.

Development of the approach was motivated by two significant advantages over the traditional OSSE approach for assessing the impact of anticipated data. The first is that the method is based on real observations thus obviating the need to simulate observations other than the anticipated new data (ADM in this case). The second is that, whereas OSSE results are difficult to interpret because of uncertainties surrounding the role of simulation biases (Masutani *et al.* 2004), the situation is less severe in the assimilation ensemble approach; this is because the ensemble-based diagnostics are obtained from differences between ensemble members which gives considerable scope for cancellation of bias effects. We argue that it is safe to assume that biases in the model and observations (exisiting as well as simulated) have similar effect on all members of an ensemble and contribute little to the calculated spread.

In order to obtain comparable error statistics for the simulated ADM data and the existing observations, we have generated the ADM data from short-range forecast winds provided by the Met Office model, rather than from the own experiment or the Control. For real data, the innovations $y - H(x_b)$ are affected by errors with respect to truth (subscript t) due to errors in the observation $\varepsilon^o = y - y_t$ and errors in the background $\varepsilon^b = x_b - x_t$. For the simulation of ADM, only, a substitute for the truth x_t is needed. Using the Control here would be inadequate, as ε^b would then be underestimated due to the similarity between the Control and ADM experiments. Using the Met Office winds (subscript m), we have $y_t = H(x_m)$, where H represents LIPAS, and $\varepsilon^b = x_b - x_m$. The use of Met Office winds is thus appropriate under the assumption that $x_b - x_m$ is statistically comparable to $x_b - x_t$ with respect to random errors and bias. The simulated observation is $y = y_t + \varepsilon^o$ where ε^o is a random sample of the LIPAS ADM observation error (σ^o).

2.2 Definition of experiments and data

Three separate ensembles have been generated, which differ in the observing systems made available to assimilate. Each of the three ensembles consisted of four independent members. Each member is an independent, 12-hourly cycling 4D-Var assimilation over the 50-day period from 10 January to 28 February 2003. The analyses and forecasts were computed at spectral resolution T159 (120 km) with analysis increments at T95 (200 km), on 60 model levels using the spring 2005 version of the ECMWF forecasting system (known as IFS cycle 29r1, see www.ecmwf.int/products/data/operational_system).

The three ensembles are: **Control**: All observational data used as in the ECMWF operational system, which includes extensive use of satellite sounding data (Thépaut and Andersson, 2003), **ADM**: As Control with simulated ADM-Aeolus added, and **NoSondes**: As Control but radiosondes, PILOT and wind-profilers were removed - collectively referred to as the Sonde dataset in the following. In all cases, the observation perturbations (ε^o) were independent Gaussian random numbers with the standard deviation equal to the specified observation errors in the assimilation scheme.

The Sonde dataset comprises approximately 1,100 radiosonde and PILOT profiles per day, unevenly distributed over the globe with concentrations over parts of Europe, North America, East Asia and Australia, and 2,500 hourly or half-hourly wind-profiler reports per day from central United States, Western Europe and Japan. It is anticipated that the ADM instrument will generate around 2,400 HLOS wind profiles per day, with a regular 200 km spacing along the orbit tracks over land as well as ocean.

The ADM data were generated by LIPAS (Marseille and Stoffelen 2003) using the same settings as Tan and Andersson (2005). The yield of the ADM data depends primarily on the presence and distribution of clouds, which were obtained from a high-resolution (T511, 40 km) assimilation with the ECMWF model. The simulated accuracy of the ADM winds varies with the cloud and aerosol (Vaughan *et al.* 1995) distributions, and the density of air. Here, the simulated ADM HLOS winds were generated from short-range forecasts provided by the Met Office, rather than ECMWF's own model, in order to avoid spurious self consistency between model and simulated data, which could otherwise lead to overly optimistic ADM impact results.

Fig. 3 shows comparable statistics for the real Sonde u-component (a) and simulated ADM HLOS (b) wind profile data accumulated over a one-week period (20030116 – 2003122). For ADM, the plot is a result of the various assumptions made in the simulations (LIPAS) and the use of Met Office winds. The plot does not in itself verify that those assumptions are correct (see Tan and Andersson 2005), but is shown here to help describe the data set that is being assessed. We see that the Sonde dataset has higher data counts than ADM through most of the troposphere (850-300 hPa), which is due to the high sampling frequency of the profiler data. Through the combination of measurement in the two receiver channels, Mie for particulate and Rayleigh for molecular returns, ADM provides good coverage throughout the troposphere and lower stratosphere. Note that in these simulations no ADM data were generated in the stratosphere above 20 km, although it is now envisaged that the mission will provide data to 26 or 30 km (Fig. 1). The standard deviation of departures (observation minus background, full lines Fig. 3) show that the simulated ADM data deviate slightly more from the ECMWF model winds than the Sonde data. This is partly because of sampling differences in the two global datasets, with the ADM providing fairly even coverage also over the ocean areas where errors are large, whereas the Sonde statistics mostly represent the data dense regions in the Northern Hemisphere where model winds are more accurate. Within the context of this study we conclude that the simulated ADM data have essentially similar accuracy to the real Sonde wind observations. This is also reflected in the specified observation errors (σ^o) for the two data sets, which results in similar closeness of fit to the analysis (dashed lines, Fig.3).

2.3 Spread and impact assessment

Statistics, such as the spread s_i within a particular ensemble i, were compiled for the period 16 January to 28 February 2003, discarding the first 6 days due to common initial conditions (that is, $\varepsilon^b = 0$, as suggested by Fisher 2003a). We compute the spread s_i as the rms over K days of daily N-member ensemble spreads, that is:

$$s_i = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{N-1}\sum_{n=1}^{N}(x_i - \bar{x})^2{}_n\right)_k} \tag{1}$$

where x can represent either analyses or forecasts. In these experiments K=44 days and N=4. The impact I_{12} is obtained as the difference in spread between two ensemble experiments (subscripts 1 and 2), with a scaling factor (α)

$$I_{12} = \alpha(s_1 - s_2) \tag{2}$$

Beneficial impact of an observing system corresponds to a reduction in ensemble spread. In a well calibrated ensemble, the spread between short-range forecast members should agree with the background error standard deviation that can be deduced from statistics of observation minus background differences. Fisher (2003a) and Houtekamer and Mitchell (2001) have found that data assimilation ensembles tend to be under-dispersive, that is, the ensemble spread underestimates the true uncertainty. Comparing innovations and spread, we found that the Control ensemble underestimates the short-range forecast error by a factor close to 2, which we adopt as the scaling factor α and apply to the results in the following. In our application it is not essential to determine α exactly. It is assumed that α is the same for the three ensembles, which is reasonable given that the spread difference between them is small ($\sim$ 5%) compared to the spread itself. Comparison of the scaled Control and NoSondes ensemble spreads permits essential validation of the approach against OSE data denial experiments in the same period (see below). The ratio between ADM and NoSonde impacts (both with respect to Control) facilitates a relative assessment of ADM versus Sonde impacts which is independent of α.

3 Results

The focus of this study is to assess impact on hemispheric scales. This is because spatial averaging is required to obtain statistically significant results, given the small ensemble size.

3.1 Comparison of ensemble-based Sonde impact against OSE results

Figs. 4 (a) and (c) show the Northern and Southern Hemisphere NoSonde impacts with respect to Control as obtained with the data assimilation ensemble technique, based on 12-hour forecast differences of 500 hPa geopotential. Dark shading indicates reduced spread in the Control, that is, beneficial impact of the Sonde data set, and light grey shading is the opposite. The relatively small ensemble size results in a great deal of noise obscuring the signal, but the main patterns of positive impact can be seen over most land areas of both hemispheres where there is good coverage of Sonde data. The result can be compared with corresponding results obtained with the established OSE technique, in the same period. The OSE results for the two hemispheres are shown in Figs. 4 (b) and (d). The OSE results are far less noisy than the current ensemble results. Most of the main areas of Sonde impact identified by the OSE method do, nevertheless, correspond to similar impacts in the ensemble results. The smaller-scale discrepances between the two sets of results are largely due to the limited ensemble size in the present study. There is relatively poor correspondence e..g. over Brazil, South Africa and Alaska. We conclude that the present ensembles enable useful observing system impact assessment at continental and hemispheric scales.

3.2 Impact of simulated ADM-Aeolus wind profiles

As an example of ADM-Aelous impact we show here the difference between ADM and Control ensemble spread for 12-hour forecasts of the 300 hPa east-west wind component (Fig.5a). Light grey shading indicates reduced spread, that is, beneficial impact of ADM, and dark shading is the opposite. Fig. 5 (b) shows the corresponding Sonde impact for comparison.

Fig. 5 (a) suggest that the main benefits from ADM-Aeolus (compared to Control) for analysed wind fields will be found over ocean regions in both hemispheres and in the tropics, and over parts of central Asia. These regions have been identified as priority areas for improvement (Tan and Andersson 2005; Stoffelen *et al.* 2005a). The ADM impact is less in the relatively well-observed regions of North America, Europe, Asia and Australia, where the wind analysis is largely determined by available wind profiles from ascending and descending aircraft

(Cardinali *et al.* 2003), radiosondes and wind profilers (Kelly *et al.* 2004). A constant scaling factor of order 2 has been applied to these results to obtain values commensurate with actual 12 hour wind uncertainties in the ECMWF system. The large ADM impact in the tropics is noteworthy. It is likely that the ADM wind data help determine the moisture convergence at low levels in the region of the ITCZ, as well as the divergent outflow in the upper troposphere. It is hoped that the ADM data will help maintain a more correct intensity of the Hadley circulation than is currently the case at ECMWF in the operational system and ERA-40 climate re-analyses (Uppala *et al.* 2005).

There is inevitably some sampling noise, in maps such as these, but this can be reduced through spatial averaging. Fig. 6 shows profiles of 12-hour forecast spread for the Northern Hemisphere (a), Southern Hemisphere (b) and the Tropics (c), for each of the three ensemble experiments: Control (full line), NoSonde (dotted) and ADM (dashed). These results are less noisy and the difference from Control is statistically significant at the 95% level (throughout the profile). Comparing to Control, we can see that the withdrawal of sondes and wind profilers increases spread (12-hour forecast error) in all three regions. Similarly, the addition of ADM data reduces forecast error in all three regions. The impact of ADM is similar in magnitude to that of sondes in both hemispheres (extratropics), whereas in the tropics ADM has larger impact than sondes. The impacts of both observing systems carry through to the medium range, as shown in Fig. 6(d) for 120 hour forecast in the Northern Hemisphere. Also for this range we see that ADM provides similar (or slightly larger) benefit than the radiosonde and profiler networks in these simulations.

3.3 Information content

We have complemented the above result with an assessment in terms of information content. Practical methods for computing information content (degrees of freedom for signal - DFS) in the observations and analyses have recently been developed (Fisher 2003b; Cardinali *et al.* 2004; Chapnik *et al.* 2005) within the context of the global 4D-Var assimilation system. The profiling observing systems that ADM can most usefully be compared with are again the radiosondes and wind profilers. Using the method of Fisher (2003b) information content was computed for each of the experiments Control, ADM and NoSonde, defined above. It was found that in terms of wind information in these experiments (Table 1) radiosondes plus wind profilers provide 3707 DFS and ADM 2743 DFS. These results are in keeping with the expected quantity, accuracy and coverage of the simulated ADM-Aeolus data. Radiosondes and wind profilers provide slightly more information in absolute terms, but require proportionately more observations to achieve this. Aeolus observations provide more information per datum because they are able to contribute information in regions that are currently poorly observed, i.e. over ocean regions in both hemispheres and throughout the tropics. Each conventional wind vector observation counts as two scalar data, each simulated HLOS wind component counts as one datum.

Data	Radiosonde and profilers	ADM-Aeolus
Number of data	90,688	50,278
DFS	3,707	2,743
Data per DFS	24.5	18.3

Table 1: Information content: Degrees of freedom for signal (DFS) in radiosonde and wind profiler data from the surface to 55 hPa and in simulated ADM-Aeolus single line-of-sight data, in data assimilation experiments with the ECMWF 4D-Var system.

4 Discussion and conclusions

A new approach for assessment of observing system impact has been described. The method consists of running multiple data assimilation ensembles with varying observing systems, interpreting differences in analysis and forecast ensemble-spread as a measure of impact. Three four-member ensembles were run with ECMWF's operational 4D-Var system at T159 (120 km) resolution for 50 days, **Control**: All observational data, **NoSondes**: Radiosondes and wind-profilers removed, and **ADM**: Simulated ADM-Aeolus wind-profiles added. The NoSondes vs. Control results were compared with established OSE impact results in the same test period, and reasonable qualitative agreement was found. The ADM vs. Control results demonstrate that the ADM-Aeolus winds would have significant impact on present-day global NWP, of similar magnitude to that of the current Sonde dataset. Our estimates of information content in the 4D-Var data assimilation system indicate that ADM-Aeolus will provide similar degrees of freedom for signal (DFS) as the global Sonde observing system: 2743 and 3707, respectively. Based on the geographical distribution of available wind information and the expected yield of good-quality ADM data it is expected that the main analysis and short-range forecast impact will be obtained over the oceans in both hemispheres, and in all regions of the tropics (Fig.5a). This confims earlier results by e.g. Tan and Andersson (2005), and complements those of Stoffelen *et al.* (2006) andŽagar (2004).

Furthermore, our results confirm that a modern data assimilation system can extract useful information from single-component (line-of-sight) wind information, succesfully reducing wind analysis uncertainty and forecast error.

The ensemble method proposed here is suitable for observation impact assessment in a data assimilation system where biases are not significant. As mentioned in section 2a, this is seen as an advantage; in OSSE methods special efforts must be made to remove spurious model bias. However, in some parts of the atmosphere, especially the stratosphere (Polavarapu 2005), biases are a serious problem. A new observing system in such areas with known or low bias would have additional benefit not measured by this method.

The scaling factor α was introduced to compensate for insufficient spread in the ensembles. On-going research is aimed at increasing the ensemble spread through improved perturbations, which requires enhanced representation of the various sources of error that contribute to analysis uncertainty: (1) observation error and its correlation; (2) representation of model error; and (3) a description of the errors in the surface boundary conditions. In the current study, operationally specified observation errors were used, SATOB observations were assumed correlated (all other observations including ADM-Aeolus were assumed un-correlated; it is a design goal of ADM-Aeolus to provide uncorrelated errors), and the sea-surface temperature fields were perturbed. Based on the experience so far it is thought that the neglect of observation error correlation for satellite radiance data, and the neglect of model error in the perturbations, are two of the main reasons for the under-estimation of ensemble spread in these experiments. One way of incorporating the effect of model error in the ensembles is by adding a stochastic forcing term to the model equations, e.g. Buizza *et al.* (1999). This will be further explored in future work.

The current study also used operationally specified quality control. The rejection of radiosonde and DWL wind components were not increased greatly by the ensemble method's error perturbations, because the relevant rejection thresholds were set at 4 to 5 standard deviations. In general, the ensemble method is most suitable for assessing the impact of observation types with similar, relatively weak, quality controls.

Acknowledgements

We thank members of ESA's ADM mission advisory group for valuable comments and suggestions, and to R. Barnes and K. Robertson for supplying Met Office winds for our ADM-Aeolus simulations. We are also grateful to our colleagues within ECMWF's Data Assimilation Section for help and useful discussions. Rob Hine skilfully prepared the figures for publication. The study was partly funded under ESA contract No. 15342/01/NL/MM.

References

Arnold, C. P. Jr. and Dey, C. H., 1986: Observing-Systems Simulation Experiments: Past, Present, and Future. *Bull. Amer. Meteorol. Soc.*, **67,** 687–695

Atlas, R. A., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75,** 111–130

Becker, B. D., Roquet, H. and Stoffelen, A., 1996: A simulated future atmospheric observation database including ATOVS, ASCAT, and DWL. *Bull. Amer. Meteorol. Soc.*, **77,** 2279–2294

Buizza, R., Miller, M., and Palmer, T. N., 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **125,** 2887–2908

Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M. and Zhu, Y., 2005: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev.*, **133,** 1076–1097

Cardinali, C., Isaksen, L. and Andersson, E., 2003: Use and impact of automated aircraft data in a global 4D-Var data assimilation system. *Mon. Wea. Rev.*, **131,** 1865–1877

Cardinali, C., Pezzulli, S. and Andersson, E., 2004: Influence matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130,** 2767–2786

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O., 2005: Diagnosis and tuning of observational error statistics in a quasi operational data assimilation system. *Q. J. R. Meteorol. Soc.*, **132,** 543–565

Derber, J. and Bouttier, F., 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A,** 195–221

Dumelow, R., 2003: Overview of observing system experiments. Pp 95–111 in Proceedings of the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September, 2003, Reading, UK.

European Space Agency, 1999: The Four Candidate Earth Explorer Core Missions – Atmospheric Dynamics Mission. *ESA Report for Mission Selection*, **ESA SP-1233(4),** 157 pp

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte-Carlo methods to forecast statistics. *J. Geophys. Res.*, **99(C5),** 10143–10162

Evensen, G. and vanLeeuwen, P., J., 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasi-geostrophic model. *Mon. Wea. Rev.*, **124,** 85–96

Fisher, M., 2003a: Background error covariance modelling. Pp 45–63 in Proceedings of the ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September, 2003, Reading, UK.

Fisher, M., 2003b: Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. *ECMWF Tech.Memo.*, **397,** 18 pp

Hamill, T. M. and Snyder, C., 2002: Using improved background error covariances from an ensemble Klaman filer for adaptive observations. *Mon. Wea. Rev.*, **130,** 1552–1572

Houtekamer, P. L. and Mitchell, H. L., 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126,** 796–811

Houtekamer, P. L. and Mitchell, H. L., 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129,** 123–137

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L. and Hansen, B., 2005: Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. *Mon. Wea. Rev.*, **133,** 604–620

Kelly, G., McNally, T., Thépaut, J.-N. and Szyndel, M., 2004: 'OSEs on all main data types in the ECMWF operational system'. Pp 63–94 in Proceedings of the 3rd WMO Workshop on The impact of various observing systems on numerical weather prediction, 9–12 March 2004, Alpbach, Austria.

Keppenne, C. and Reinecker, H., 2002: Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.*, **130,** 2951–2965

Marseille, G.-J. and Stoffelen, A., 2003: Simulation of wind profiles from a space-borne Doppler wind lidar. *Q. J. R. Meteorol. Soc.*, **129,** 3079–3098

Masutani, M., Lord, S. J., Woolen, J. S., Yang, W., Sun, H., Kleespies, T. J., Emmitt, G, D., Wood, S. A., Katz, B., Treadon, R., Greco, S. and Terry, J., 2004: Global OSSE at NCEP. Proceedings of the 8th AMS Symposium on Integrated observing and assimilation systems for atmosphere, ocean and land surface, 10–14 January, Seattle, U.S.

Polavarapu, S., 2005: The challenges of middle atmosphere data assimilation. *Q. J. R. Meteorol. Soc.*, **131,** 3513–3528

Riishøjgaard, L.-P., Atlas, R. and Emmitt, G. D., 2004: The impact of doppler lidar wind observations on a single-level meteorological analysis. *J. Appl. Meteorol.*, **43,** 810–820

Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J. M., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., Meynart, R., Endemann, M. and Ingmann, P., 2005a: The atmospheric dynamics mission for global wind measurement. *Bull. Amer. Meteorol. Soc.*, **86,** 73–87

Stoffelen, A., Marseille, G.-J., Andersson, E. and Tan, D. G. H., 2005b: Comments on "The impact of doppler lidar wind observations on a single-level meteorological analysis" by Riishøjgaard, L.-P., R. Atlas and G. D. Emmitt. *J. Appl. Meteorol.*, **44,** 1276–1277

Stoffelen, A., Marseille, G.J., Bouttier, F., Vasiljević, D., de Haan, S. and Cardinali, C., 2006: ADM-Aeolus doppler wind lidar observing system simulation experiment. *Q. J. R. Meteorol. Soc.*, **132,** 1927–1947

Tan, D. G. H. and Andersson, E., 2005: Simulations of the yield and accuracy of wind profile measurements from the Atmospheric Dynamics Mission (ADM-Aeolus). *Q. J. R. Meteorol. Soc.*, **131,** 1737–1757

Thépaut, J.-N. and Andersson, E., 2003: Assimilation of high-resolution satellite data. *ECMWF Newsletter*, **97,** 6–12

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljević, D., Viterbo, P. and Woollen, J., 2005: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131,** 2961–3012

Vaughan, J. M., Brown, D. W., Nash, C., Alejandro, C. B. and Koenig, G. G., 1995: Atlantic atmospheric aerosol studies 2. Compendium of airborne backscatter measurements at 10.6 μm. *J. Geophys. Res.*, **100**(D1), 1043–1065

Žagar, N., 2004: Assimilation of equatorial waves by line of sight wind observations. *J. Atm. Sci.*, **61,** 1877–1893

Žagar, N., Andersson, E. and Fisher, M., 2005: Balanced tropical data assimilation based on study of equatorial waves in ECMWF short-range forecast errors. *Q. J. R. Meteorol. Soc.*, **131,** 987–1011

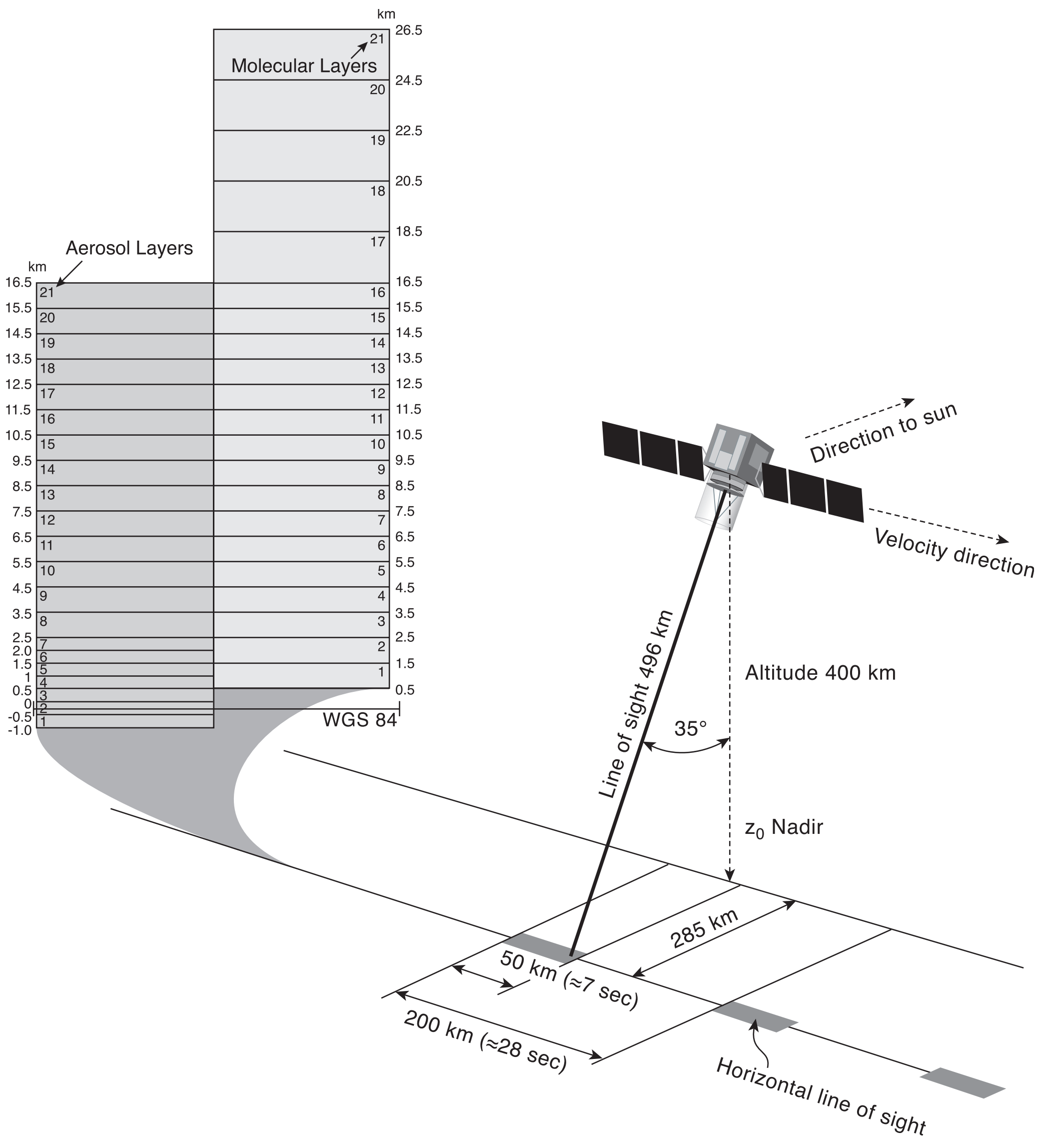

Figure 1: ADM-Aeolus viewing geometry and vertical resolution (Source: ESA website http://www.esa.int). Laser light is emitted from the satellite. A small fraction of the light is backscattered and its frequency shift (which is wind-induced) is measured in two channels by the onboard receivers: Mie for aerosol returns and Rayleigh for molecular returns. The combination of the two channels will enable good quality line-of-sight wind retrieval from 26 km to the surface in 0.5, 1.0 and 2.0 km layers as indicated, except where the view is restricted by thick clouds.

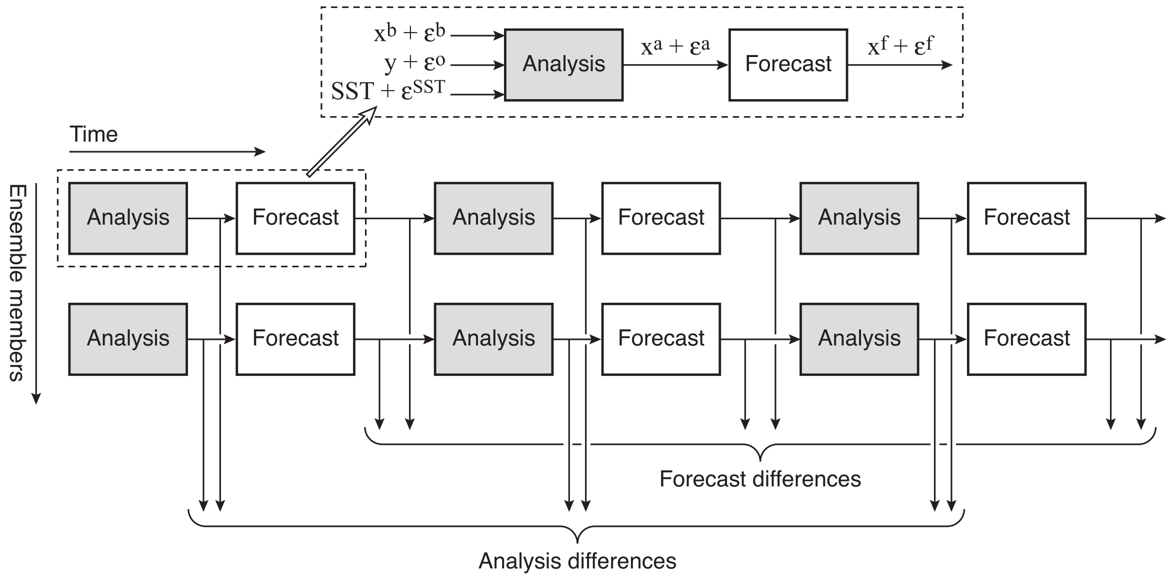

Figure 2: Ensemble of data assimilation cycles, schematic. See main text for definitions.

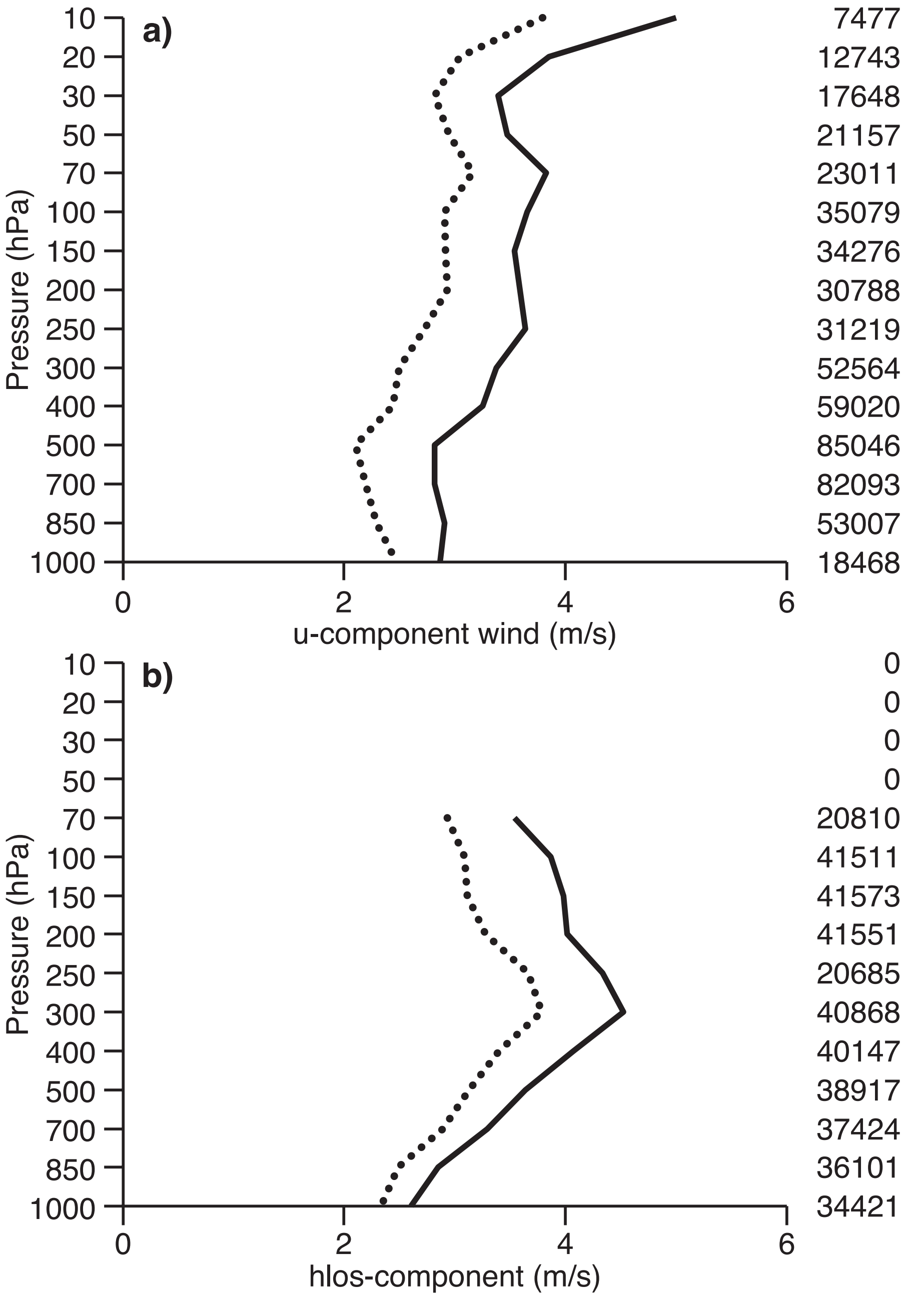

Figure 3: Standard deviation of departures (observation minus background, full lines) and residuals (observation minus analysis, dashed) for (a) the Sonde data set and (b) the simulated ADM-Aeolus data assimilated in the experiments. The column on the right shows the number of observations in each layer. This plot is based on global data in the period 20030116 – 20030122. Below 300 hPa the DWL statistics are dominated by Mie channel data (errors increase with height). Above 300 hPa the DWL statistics are dominated by Rayleigh data (errors decrease with height).

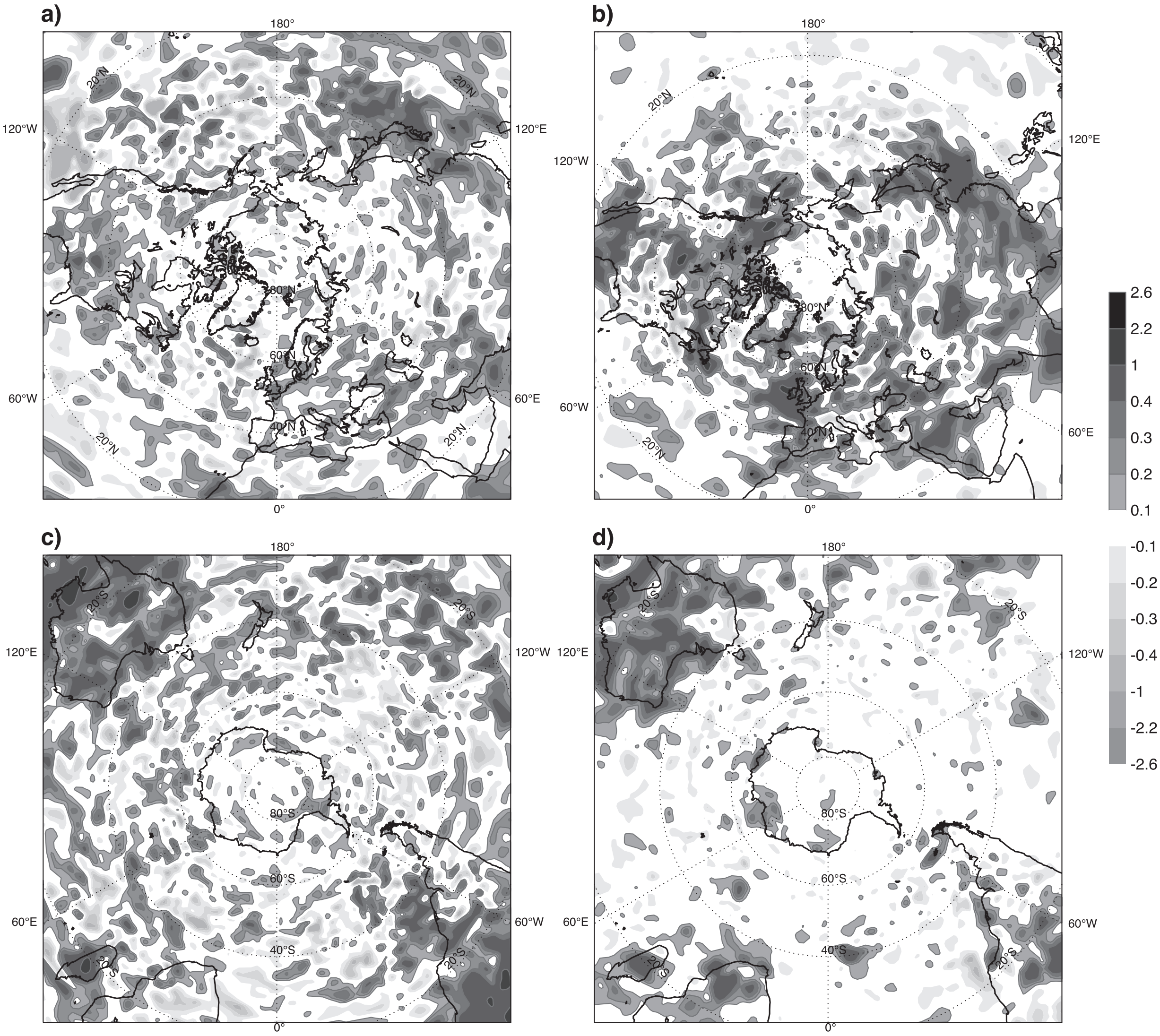

Figure 4: Impact in terms of 500 hPa zonal wind component (ms^{-1}) of radiosonde and wind profiler data as deduced from (a) the difference in scaled 12 hour forecast spread between two ensembles: Control and NoSonde and (b) the corresponding OSE for the same period, for the Northern Hemisphere; and (c) and (d) for the Southern Hemisphere. Dark (light) shading (see legend) indicates that the ensemble spread, i.e. the uncertainty, is reduced (increased) using the Sonde dataset. Shading in the range -0.1 to 0.1 is suppressed.

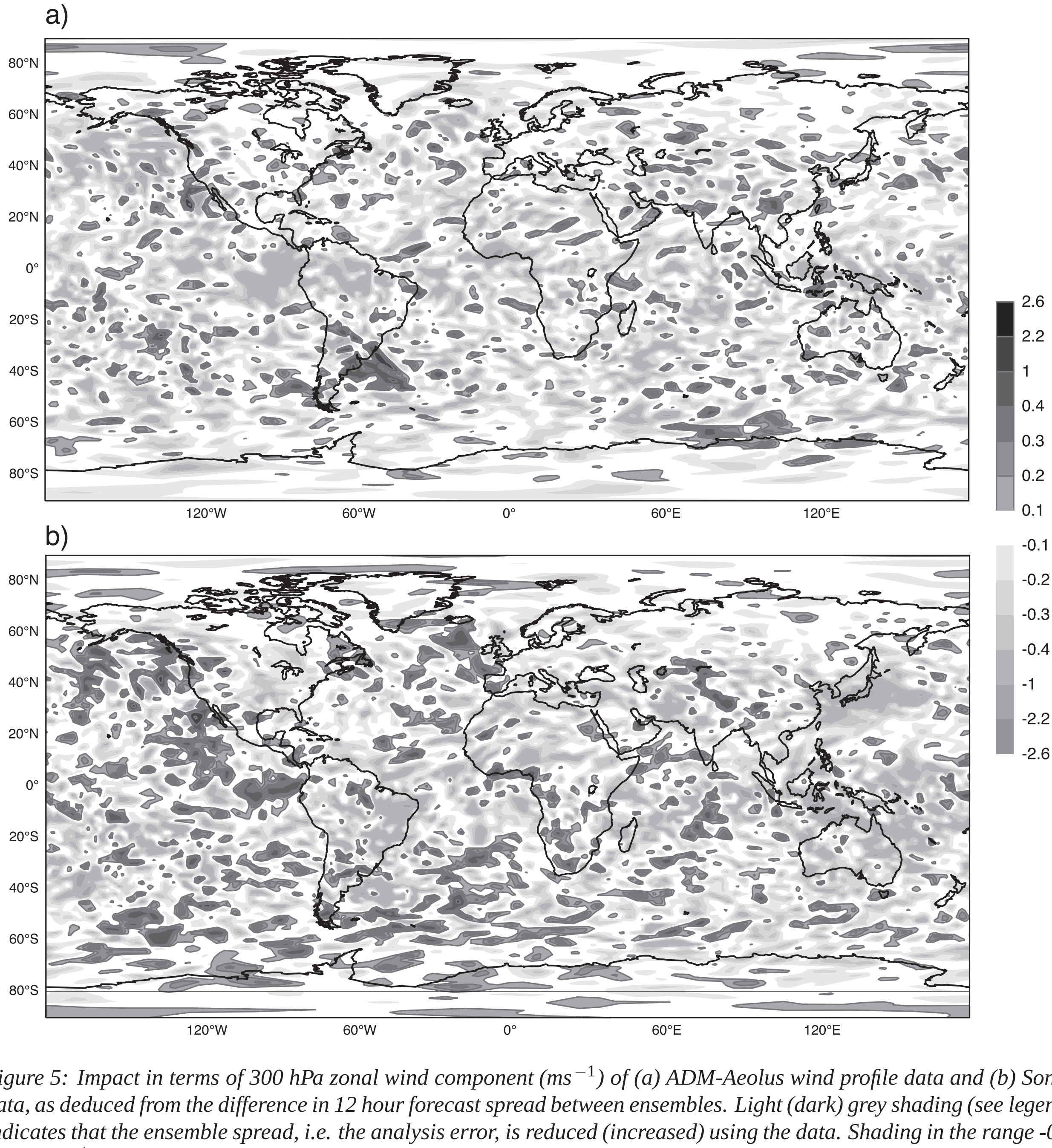

Figure 5: Impact in terms of 300 hPa zonal wind component (ms^{-1}) of (a) ADM-Aeolus wind profile data and (b) Sonde data, as deduced from the difference in 12 hour forecast spread between ensembles. Light (dark) grey shading (see legend) indicates that the ensemble spread, i.e. the analysis error, is reduced (increased) using the data. Shading in the range -0.1 to 0.1 ms^{-1} is suppressed.

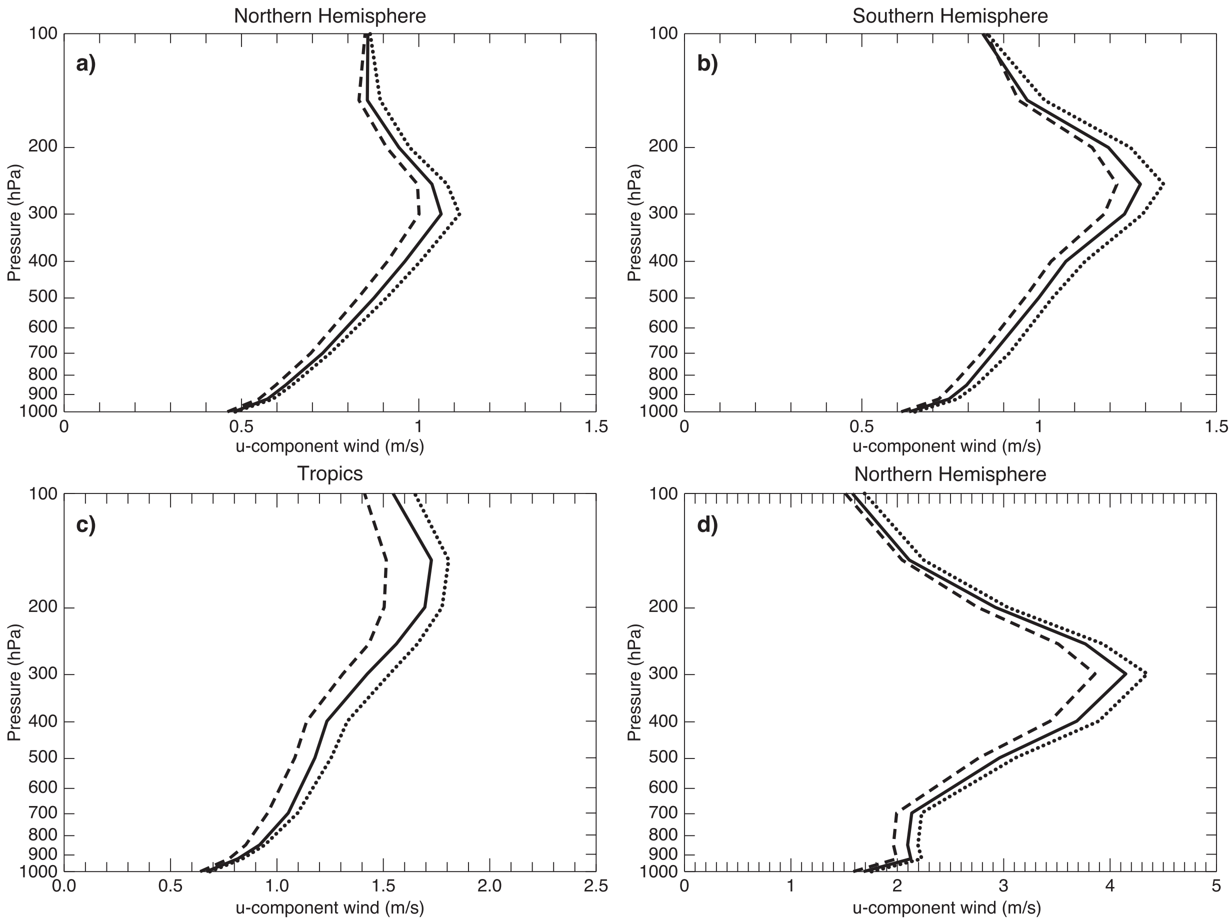

Figure 6: R.m.s. east-west wind component ensemble spread, or forecast error, (ms^{-1}), for (a) the extratropical Northern Hemisphere, (b) Southern Hemisphere, and (c) the tropics, at 12-hours, and (d) for the Northern Hemisphere at 120 hours, for each of the three experiments: Control (full line), NoSonde (dotted) and ADM-Aeolus (dashed). Note the scale runs from 0 to 1.5 in (a) and (b), to 2.5 in (c) and to 5.0 ms^{-1} in (d).