Tellus (2005), 57A, 701–708
Printed in Singapore.

TELLUS

On the impact of humidity observations in numerical weather prediction

By LENNART BENGTSSON and KEVIN I. HODGES*, *Environmental Systems Science Centre (ESSC), University of Reading, Harry Pitt Building, Whiteknights, PO Box 238, Reading, RG6 6AL, UK*

(Manuscript received 19 October 2004; in final form 11 February 2005)

ABSTRACT

The impact of humidity observations on forecast skill is explored by producing a series of global forecasts using initial data derived from the ERA-40 reanalyses system, in which all humidity data have been removed during the data assimilation. The new forecasts have been compared with the original ERA-40 analyses and forecasts made from them. Both sets of forecasts show virtually identical prediction skill in the extratropics and the tropics. Differences between the forecasts are small and undergo characteristic amplification rate. There are larger differences in temperature and geopotential in the tropics but the differences are small-scale and unstructured and have no noticeable effect on the skill of the wind forecasts. The results highlight the current very limited impact of the humidity observations, used to produce the initial state, on the forecasts.

1. Introduction

Recent investigations making use of so-called reanalysis data have demonstrated that the accuracy of numerical weather prediction (NWP) has improved significantly over the last decades (Bengtsson, 2001). This improvement is due to a combination of the deployment of new observing systems, such as satellites which provide information with global coverage, and to more efficient ways of utilizing the available observations based on more sophisticated models and data assimilation systems (Simmons and Hollingsworth, 2002). The predictive skill of weather forecasts is inherently limited to about two weeks on average because of the rapid growth of errors in the initial state (Lorenz, 1982). The initial state is defined by the state of pressure, density, temperature, wind and humidity through the depth of the atmosphere. A key objective in NWP has therefore been to find ways to improve the initial state as much as possible. The predictive skill also depends on accurate information of sea surface conditions, such as temperature and sea ice, land surfaces and atmospheric composition including aerosols. Here it is assumed that this information is known and we will only consider the effect of atmospheric data on the prediction of weather.

To determine the initial state for a forecast integration, at the resolution required by present atmospheric global models, calls for the knowledge of all prognostic variables to be known every 20–100 km in the horizontal and at some 30–100 vertical levels from the surface to the top of the atmosphere (30–100 km above the surface). Fortunately, this extremely demanding requirement can be significantly relaxed.

First, pressure and density through the atmosphere can be accurately obtained from knowledge of surface pressure, humidity and temperature through the use of the equation of state and the hydrostatic relation. For the scales of motions considered in global weather models, these relations can be applied with high accuracy (Phillips, 1973). Secondly, fields of temperature, wind and humidity have characteristic structures in time and space, making it possible to determine a high-resolution field from a much smaller number of observational data than the number of grid points. Thirdly, fields of temperature together with surface pressure and wind are coupled through a semi-balancing relation such as the geostrophic relation, which means that the temperature field (mass field) can approximately be obtained from the wind field, and vice versa the wind field from the temperature field. The geostrophic adjustment is a more complex process than the hydrostatic adjustment as it depends on the latitude and on the scale of atmospheric fields. For large scales and high latitudes, the mass field has preference over wind, while the opposite is the case for smaller scales and lower latitudes (Temperton, 1976). The geostrophic adjustment has a time-scale which is determined by the inverse of the vertical component of the local Coriolis force, corresponding to about half a day at middle latitudes. The intrinsic adjustments depend on the actual weather systems and cover a broad range from hours to a few days and are largely unrelated to the time-scale of observations and data assimilation.

*Corresponding author.
e-mail: kih@mail.nerc-essc.ac.uk

Finally, the evolution of the state of the atmosphere as described by a complex model includes internal adjustments between the dynamical state of the model and internal physical processes influencing, for example, many aspects of the hydrological cycle. Present models handle the large-scale aspects of the hydrological cycle rather satisfactorily (Bengtsson and Arpe, 2000; Hagemann et al., 2004). Evaporation is largely determined by boundary layer fluxes and precipitation by the convergence of water vapour driven by dynamical processes in the atmosphere. This means, for example, that if a model integration begins with an incorrect humidity field, this will change to a new state which is consistent with the three-dimensional fields of wind and temperature and their evolution in time. In a dynamically active region, such as where a cyclogenesis is taking place, this process is fast, with a time-scale of several hours only, while in less active regions it may extend over a few days.

However, the reverse does not take place. Assume that a forecast is started with incorrect wind or temperature fields but with the correct humidity field. Areas of divergence and convergence will immediately start to change the humidity field, but the humidity field will not change the temperature and the wind field, as there are no immediate physical processes by which this can occur. It might be expected that the condensation of water vapour, creating a source of latent heat, would result in a change in the temperature distribution and gradually, through geostrophic adjustment, also the wind field. However, the latent heat sources are generally small-scale and the corresponding change in the wind field is generally slow and disorganized.

It has been suggested that more advanced data assimilation systems, such as four-dimensional variational (4DVar) assimilation should be able to handle the assimilation of humidity much better (Hólm et al., 2002). We believe that this will not necessarily be the case because of the slowness of the feedback, via latent heat sources to the three-dimensional wind field, which could be of the order of days. 4DVar assimilation cannot span an overly large time interval, as then the underlying assumptions behind the 4DVar assimilation, such as small incremental changes, become invalid.

Different techniques have been used to overcome this intrinsic problem. Krishnamurti et al. (1991) have introduced estimated heat sources based on synoptic and empirical rules. Mahfouf et al. (2003) have assimilated rainfall rates from the Special Sensor Microwave/Imager (SSM/I) and Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI). However, before the data assimilation community embarks on the development of methods for humidity assimilation, which unavoidably include strong empirical components of an ad hoc nature, we believe we need to better understand the impact of humidity observations on predictive skill. We are in no way questioning in this paper the effect of humidity observations on prediction but only the predictive impact of that part of the humidity field which cannot be formed within an advanced prediction model through the action of assimilated wind and temperature data. Recent work by Erik Anderson (private communication) at the European Centre for Medium-Range Weather Forecasts (ECMWF) using a new formulation for the assimilation of humidity (Hólm et al., 2002) has shown there is an impact of humidity observations in the short range, but found no impact in the medium range. Humidity observations have also been found to have an impact on forecasting tropical cyclones (Kamineni et al., 2003) and small-scale events, again in the short range. This is consistent with the previous discussion and our results presented here.

Early studies by Smagorinsky et al. (1970) investigated the relative importance of several variables in the initial conditions for dynamical weather prediction. Although the study was limited to the Northern Hemisphere (NH) and to one integration, it showed that initial humidity data had insignificant effect on the predictive skill. In order to explore further the impact of the humidity observations on the predictive skill of a more current model, a radical experiment has been performed whereby all the humidity data in the meteorological observations are rejected and instead the model and the subsequent assimilation of temperature and wind and surface data are allowed to determine the humidity field. Needless to say, this can only satisfactorily work when the atmospheric model is physically and dynamically realistic and capable of describing the different aspects of the hydrological cycle sufficiently well. Bengtsson et al. (2004a) have investigated this by removing all humidity data from the different observing systems used in the ECMWF 40-yr reanalysis (ERA-40) system (Simmons and Gibson, 2000) and then repeating the data assimilation. The result was that hardly any differences could be found in the extratropics, while minor differences could be identified in the tropics, although without any systematic organization. In order to explore further the significance of the assimilation without humidity observations, we have here undertaken a major prediction study using the operational ECMWF forecasting model and using the analyses from the original ERA-40 analyses as well as our previous ERA-40 experiment (Bengtsson et al., 2004a) for the initial states. We describe the new experiment in Section 2, and discuss the results in Section 3. In the final section we analyse the results of the study in a more general context and the possible consequences it may have for setting priorities in observing and monitoring the atmosphere and in the development of data assimilation systems.

2. The Experiment

We have undertaken global, 7-d forecasts starting from an analysis every 6 h during the period 1 December 1990 to 28 February 1991 (360 forecasts). Additionally, a further set of forecasts has been produced for the period 1 June 2000 to 31 August 2000 (368 forecasts). These two sets of forecasts for two different seasons and two different years allow us to exclude any dependency on the chosen period from our results and conclusions. For this experiment we have used the ECMWF operational model (for a detailed description of the model, see White, 2000) and the analyses are from our original no-humidity data assimilation experiment (Bengtsson et al., 2004a) as well as the original ERA-40 analyses. This model is one of the most advanced in operational use and capable of predicting the global atmosphere with an accuracy just barely less than what is theoretically possible (Simmons and Hollingsworth, 2002). We have used a later and further improved version of the ECMWF model, IFS 26R3, than that used in ERA-40 (IFS 23R4). The horizontal (spectral) and vertical resolution of T159L60 is identical to that used in ERA-40, so no interpolation was required. However, the horizontal resolution used here is significantly less than that used in the current ECMWF operational version (T511).

As well as producing new forecasts from the original ERA-40 analyses to exclude any dependency on the use of different models on our results and conclusions, we have also compared our results with the original ERA-40 forecasts, i.e. generated with the 23R4 version of the model. As measures of the predictive skill the root mean square error (RMSE) of the geopotential at 500 hPa (Z_{500}) is used for the extratropics normalized by the standard deviation (StD) of the ERA-40 analyses. In the tropics, the absolute error in the winds, $\langle[(\Delta U)^2 + (\Delta V)^2]^{1/2}\rangle$, at 850 and 250 hPa, is used.

3. Results

For convenience, the new prediction experiment using humidity fields generated by the assimilating model and without using humidity observations in the assimilation will be called NOHUM

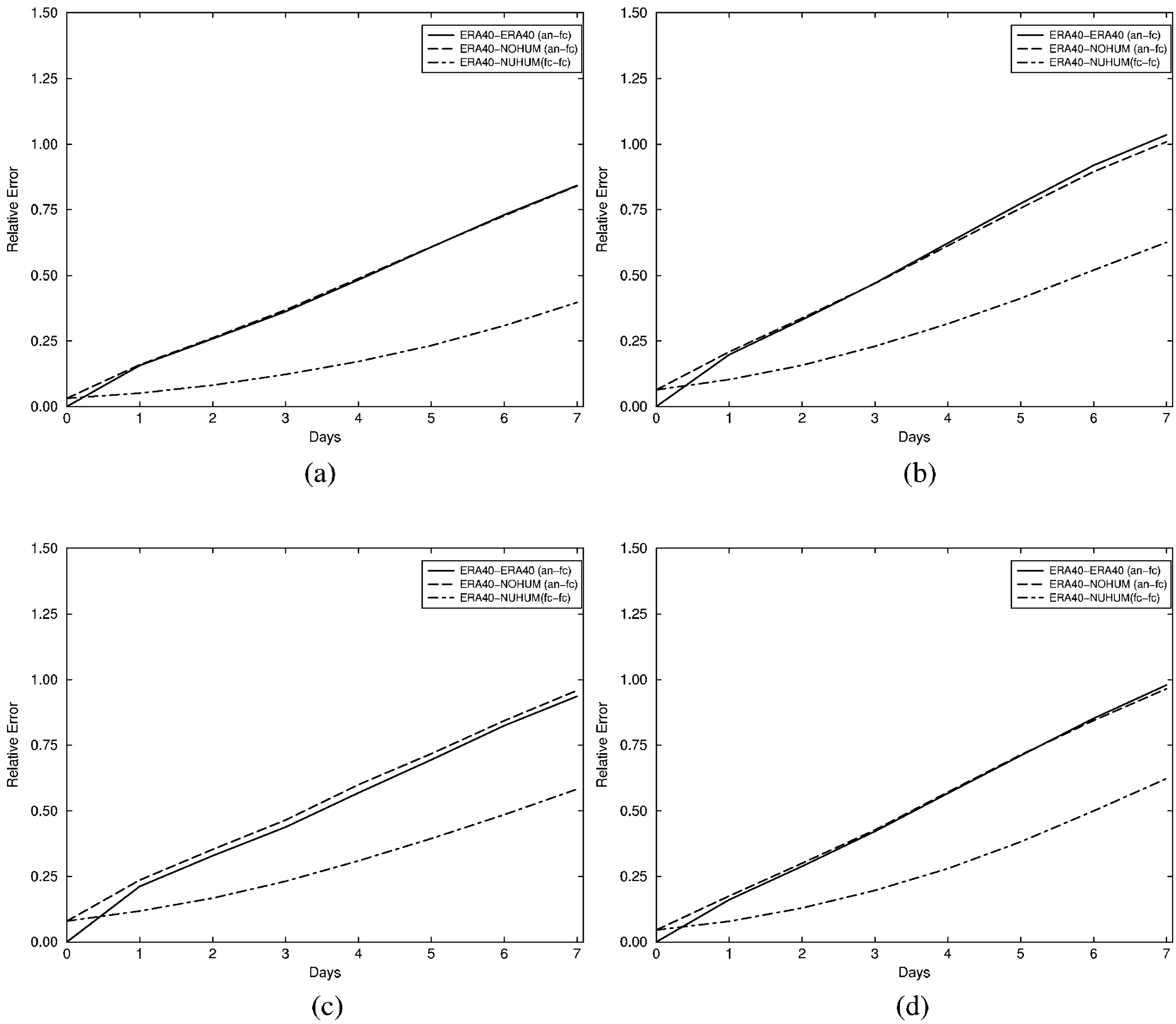

Fig 1. RMSE based on geopotential at 500 hPa for (a) the NH (20°–90°N), averaged for DJF 1990/1991, (b) the Southern Hemisphere (20°–90°S), averaged for DJF 1990/1991, (c) same as (a) but for JJA 2000, (d) same as (b) but for JJA 2000. Error is normalized by the standard deviation of the ERA-40 analysis before area averaging. The full line denotes ERA-40, the dashed line NOHUM, and the dash-dotted line the forecast–forecast comparison for NOHUM–ERA-40. Verification is against ERA-40.

to distinguish it from the ERA-40 forecasts (26R3 model) and analyses (23R4 model). The two sets of forecasts have been verified against the original ERA-40 analysis poleward of 20°N and 20°S, respectively, for the two chosen periods.

The two sets of predictions, from the ERA-40 and the NOHUM experiment, are for all practical purposes identical in quality in the extratropics of both hemispheres and for both periods, as shown in Fig. 1. The NOHUM experiment has in general slightly less error growth than ERA-40, but the difference in the skill of the forecasts for both hemispheres (Fig. 1) is so small that it can hardly be considered as significant. This is true for both periods, although in the NH the error growths are almost identical and for the JJA 2000 period the skill (Fig. 1c) is slightly worse than the corresponding DJF 1990/1991 period. The reason for this deterioration in the skill needs further study. One possibility is the indirect impact of some of the other satellite temperature channels on the moisture field. We have also considered the evolution in time of the difference between the two sets of predictions, indicating the predictability. These results are shown in Table 1 for the winter periods of both hemispheres. After an error saturation of some 25% (6% of the error variance), which occurs after about 5 d, the error growth is virtually identical to the predictability estimate that we have calculated from the differences of successive forecasts (Lorenz, 1982). Performing the same analysis with the original ERA-40 forecasts (produced with 23R4) produced nearly identical results. We have examined a number of other parameters, but there are no notable differences between the two sets of forecasts. Nor are there any systematic differences between the two sets of forecasts, as can be seen from Fig. 2 showing the mean absolute error at day 5 for Z_{500} for the winter periods of the two hemispheres for ERA-40 and NOHAM. The error patterns are virtually identical in both structure and amplitude. The analysis as shown in Fig. 2 has also been performed with the original ERA-40 forecasts with no noticeable differences from the results presented here.

As we expect humidity observations to have the largest effect in the tropics, we have validated the forecasts in the band 20°N–20°S. As the height field has very little variance in the tropics, we validate instead the wind field at 850 and 250 hPa. These results are shown in Fig. 3, for both periods, for the absolute error for ERA-40 and NOHUM, as well as the change in time of the difference between the two sets of forecasts. Both sets of forecasts have been validated against ERA-40, which means that the ERA-40 runs have unrealistic small errors in the short range. Comparison with National Center for Environmental Prediction (NCEP) analyses suggests that the ERA-40 predictions have much larger errors for the first 3 d or so. Independent estimates of the initial wind error are suggested to be of the order of 1–2 m s^{-1} at best. From day 3 onward, the errors of the two sets of forecasts are the same and so are the slopes of the error growth curves. As with the extratropics, these results are consistent for both periods and also for the original ERA-40 forecasts (not shown). However, for the JJA 2000 period there is again a slight deterioration in the skill, as seen in the NH Z_{500} comparison (Fig. 1). In Fig. 4 we have calculated the day-5 mean wind field error for 850 hPa for the two periods. The result is the same as for the extratropical 500-hPa height field, namely that the two sets of forecasts have almost identical error patterns.

Table 1. (a) Summary of the normalized error for the NH of 500-hPa height for DJF 1990/1991, averaged over 360 forecasts and the area poleward of 20°N. Column 2: growth of the difference between the two forecasts day by day. Column 3: growth of the difference between successive forecasts (same as Lorenz, 1982). Column 4: ERA-40 prediction error growth validated against ERA-40 analyses. Column 5: the same for the NOHUM forecasts, validated against ERA-40 analyses. All results have been normalized before area averaging by the standard deviation of the ERA-40 analysis for each grid point. Area averaged standard deviation for this period was 103 m. (b) The same as in (a) but for the Southern Hemisphere (SH) for JJA 2000, averaged over 368 forecasts and the area poleward of 20°S. Area averaged standard deviation for this period was 102 m.

Time (d)	NOHUM–ERA-40	NOHUM	ERA-40	NOHUM
(a) NH, DJF 1990/1991				
0	0.031	–	0.000	0.031
1	0.051	0.158	0.156	0.159
2	0.082	0.195	0.258	0.262
3	0.123	0.260	0.362	0.368
4	0.171	0.346	0.483	0.488
5	0.232	0.439	0.607	0.608
6	0.309	0.541	0.731	0.727
7	0.397	0.641	0.842	0.840
(b) SH, JJA 2000				
0	0.046	–	0.000	0.046
1	0.079	0.161	0.161	0.175
2	0.130	0.216	0.289	0.301
3	0.197	0.292	0.422	0.427
4	0.280	0.392	0.567	0.571
5	0.382	0.503	0.713	0.713
6	0.499	0.627	0.852	0.844
7	0.623	0.750	0.979	0.964

We conclude therefore that the observed humidity has a negligible effect on the forecast skill even in the tropics.

We have estimated the predictability of the model by this experiment and established that the error growth between days 4 and 7 is 1.2–1.5 d for variance doubling. The same error growth is found by Simmons and Hollingsworth (2002) comparing consecutive predictions as suggested by Lorenz (1982). However, because the original error level is higher with this approach (as it is identical to a 1-d forecast error) a similar variance doubling time in this case occurs between days 2 and 5 of the forecasts (Table 1). As pointed out by Simmons and Hollingsworth, the error growth is faster for smaller errors and thus faster in the early part of the forecast. The error growth is also faster than found by Lorenz. It is interesting to note that when the error has reached the same level in the two estimates of predictability, the error growth

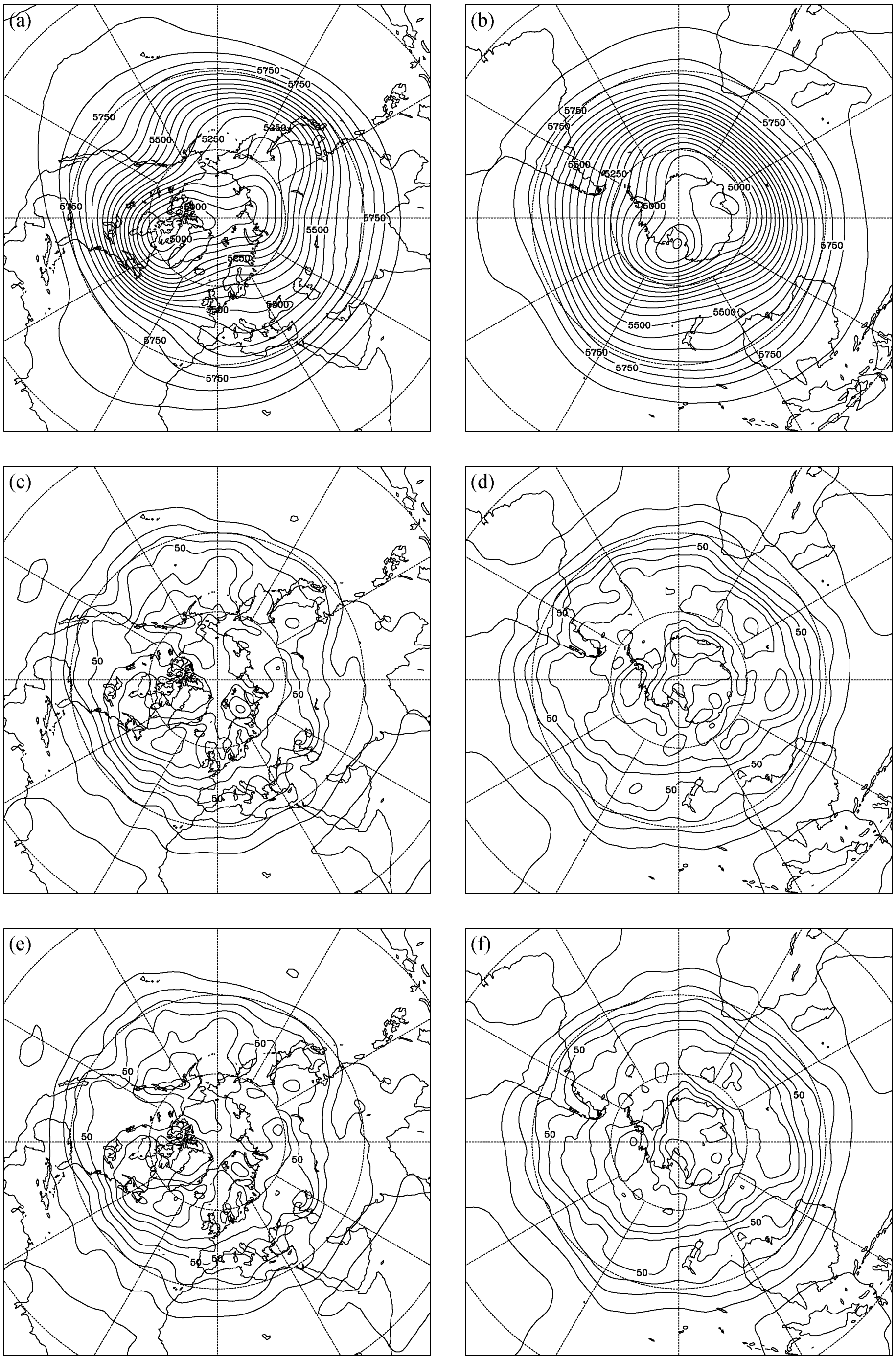

Fig 2. (a) Spatial distribution of the day-5 absolute errors based on geopotential at 500 hPa: (a) NH mean geopotential for DJF 1990/1991; (b) the same for the Southern Hemisphere for JJA 2000; (c) mean of the day-5 absolute error for the NH for ERA-40 for DJF 1990/1991; (d) mean of the day-5 absolute error for the Southern Hemisphere for ERA-40 for JJA 2000; (e) the same as (c) but for NOHUM; (f) the same as (d) but for NOHUM. Verification against ERA-40; units are m.

is also the same. When the model has reached a level of around 6% error variance saturation, the error growths in both ERA-40 and NOHUM are almost identical and only 10–15% higher than the estimated predictability. Following Lorenz, we consider the estimated predictability to be an upper bound on predictive skill. This means that, given no further reduction of the initial error, the possibility of further forecast improvements is limited. As both the ERA-40 and the NOHUM forecasts have the same

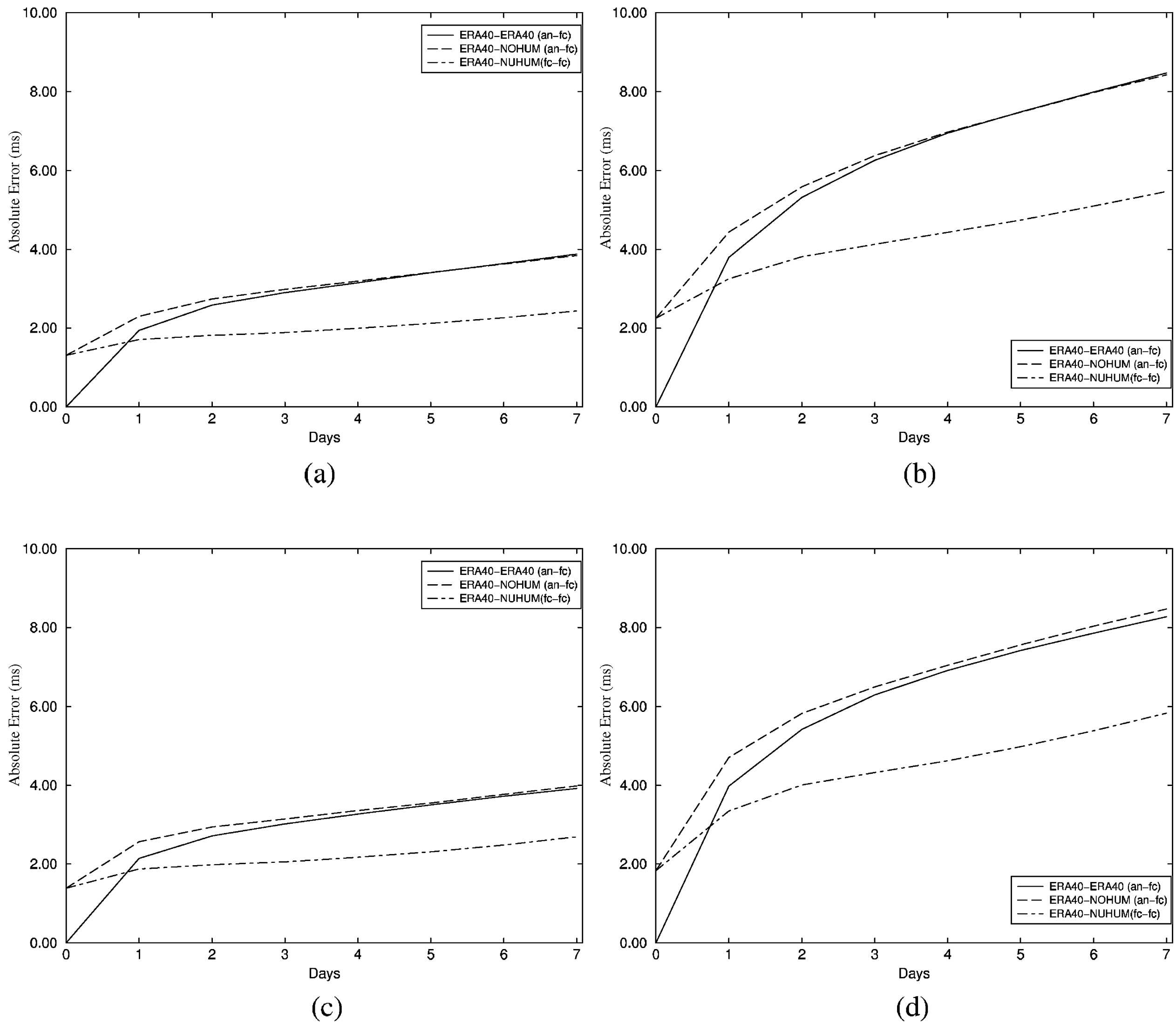

Fig 3. (a) Absolute error in the winds, determined as $\langle[(\Delta U)^2 + (\Delta V)^2]^{1/2}\rangle$, for the tropics (20°N–20°S) for (a) 850 hPa, DJF 1990/1991, (b) 250 hPa, DJF 1990/1991, (c) 850 hPa, JJA 2000, (d) 250 hPa, JJA 2000. The full line denotes ERA-40, the dashed line NOHUM, and the dotted line the forecast–forecast comparison for NOHUM–ERA-40. Verification against ERA-40.

error growth and consequently the same skill, we conclude that higher forecast skill from more accurate humidity observations is highly unlikely.

4. Discussion

The experiments demonstrate that humidity observations have very little impact on forecast skill beyond about 1–2 d. This is a robust result, as indicated by very similar results being produced from two widely separated periods and for two different seasons. On the other hand, the incorporation of additional observations for temperature and wind (Bengtsson et al., 2004b) have a distinct impact on the quality of the numerical forecasts. The reason is related to the way the different variables in the model influence each other. In an advanced numerical model, the hydrological cycle is calculated in considerable detail, including sources (through evaporation), sinks (through precipitation) and by transporting it vertically and horizontally. A crucial role is played by winds (all aspects of the hydrological cycle) and temperature essentially in affecting vertical stability. Assume that we start a numerical integration with a dry atmosphere. Evaporation will immediately commence as determined by the gradient of water vapour over wet surfaces. Water vapour will disperse from the sources and after a time period of some weeks the atmosphere will end up with a distribution of water vapour consistent with the three-dimensional circulation of the model. As demonstrated by different studies (Wang and Zwiers, 1999; Jacob, 2001; Hagemann et al., 2004) numerical models today are able to

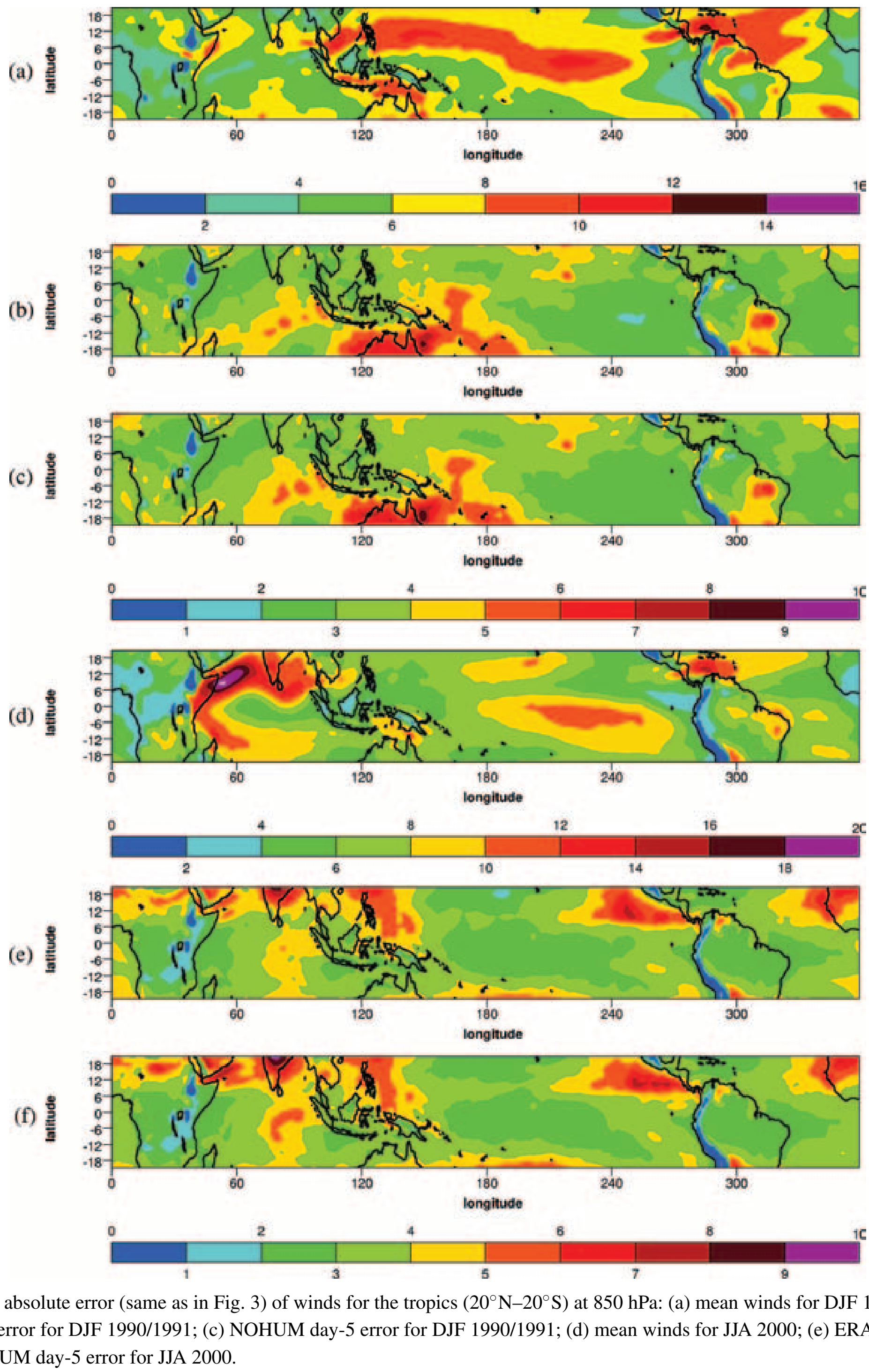

Fig 4. Mean day-5 absolute error (same as in Fig. 3) of winds for the tropics (20°N–20°S) at 850 hPa: (a) mean winds for DJF 1990/1991, (b) ERA-40 day-5 error for DJF 1990/1991; (c) NOHUM day-5 error for DJF 1990/1991; (d) mean winds for JJA 2000; (e) ERA-40 day-5 error for JJA 2000; (f) NOHUM day-5 error for JJA 2000.

generate a hydrological cycle in good agreement with climatology. In fact, such a generated hydrological cycle is in several ways comparable to what can be inferred from observations (Bengtsson and Arpe, 2000).

What will then happen if we have accurate data of humidity but poor wind and temperature data? Unfortunately, this is not very helpful because we can hardly reproduce wind and temperature from the humidity field as the humidity field itself is the result of changing temperature and wind over a period of time. It has been suggested that a variational approach in time and space, so-called 4DVar, could do so. However, a 4DVar system can only successfully be applied over a period during which non-linear processes can be ignored. This limits the time to the order of 12 h and hardly longer than 24 h, while the time-scale for water vapour in the atmosphere is at least one week. This means that even in cases where humidity data are more accurate than wind and temperature the influence on the forecast will be minor as the wind and temperature fields in any case will modify the humidity field through the dynamical circulation taking place during the forecast.

Our experiments show that very little additional skill, if any, is achieved by the use of observed humidity data beyond about 1–2 d. Instead we suggest that more accurate observations of wind and temperature are needed preferably in the upper troposphere where present temperature and wind observations have deficiencies.

Needless to say, as clouds and water vapour are generally more easily observed than at least the wind field, a practical question to ask is whether there are other ways we can make use of moisture data either prior to the assimilation process or during the assimilation. One idea recently investigated by Mahfouf et al. (2003) was to identify areas of precipitation rates as indicated from SSM/I and TMI. The success in this approach depends on how accurately the heat sources can be determined. Another possibility is to use the moisture data to infer the winds, which can then be assimilated, as is already done with cloud motion winds.

A valid question to ask, in view of the fact that this study underpins the early work by Smagorinsky et al. (1970), is whether this result is a consequence of principal limitations of present models and data assimilation in using moisture information or whether the result here is rather an indication of poor NWP models and assimilation methods? Based on these investigations we are not able to answer such a question. We suggest, though, that analysed water vapour may have a positive impact on the short-range prediction of precipitation, cloudiness and small-scale structures, but is less likely to have any significant impact on the longer forecasts in general beyond 2–3 d.

As always in model studies a general reservation is required whether the result is an artefact of the specific model being used. We do not expect this to be the case, as the ECMWF model is probably the most advanced model presently in operational use. Whether any other observations or future model will behave in a different way is an open question. We suggest, however, that similar studies as here are undertaken with future models as we consider it essential to demonstrate the need and positive impact of any set of observations in NWP.

5. Acknowledgments

The authors would like to thank the ECMWF for making the ERA-40 and operational systems available to us and for providing the computing resources for running the forecasts and analyses.

References

Bengtsson, L. 2001. Numerical modelling of the energy and water cycle of the Baltic Sea. *Meteorol. Atmos. Phys.* **77**, 9–17.

Bengtsson, L. and Arpe, K. 2000. Wasser im Klimasystem. In: *Wasser* (eds B. Busch, and L. Förster). Kunst- und Ausstellungshalle der Bundesrepublik Deutschland. Wienand Verlag, Bonn und Köln, pp. 116–123.

Bengtsson, L., Hodges, K. and Hagemann, S. 2004a. Sensitivity of large-scale atmospheric analyses to humidity observations and its impact on the global water cycle and tropical and extratropical weather systems in ERA-40. *Tellus* **56A**, 202–217.

Bengtsson, L., Hodges, K. and Hagemann, S. 2004b. Sensitivity of the ERA-40 reanalysis to the observing system: determination of the global atmospheric circulation from reduced observations. *Tellus* **56A**, 456–471.

Hagemann, S., Machenhauer, B., Jones, R., Christensen, O. B., Deque, M. and co-authors. 2004. Evaluation of water and energy budgets in regional climate models applied over Europe. *Climate Dyn.* **23**, 547–567.

Hólm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F. and co-authors. 2002. Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. ECMWF Technical Memo 383, ECMWF, Shinfield Park, Reading, RG2 9AX, UK

Jacob, D. 2001. A note to the simulation of the annual and interannual variability of the water budget over the Baltic Sea drainage basin. *Meteorol. Atmos. Phys.* **77**, 61–73.

Kamineni, R., Krishnamurti, T. N., Ferrare, R. A., Ismail, S. and Browell, E. V. 2003. Impact of high-resolution water vapour cross-sectional data on hurricane forecasting. *Geophys. Res. Lett.* **30**, 1234 (doi: 10.1029/2002GL016741).

Krishnamurti, T., Xue, N., Bedi, J. H. S., Ingles, K. and Oosterhof, D. 1991. Physical initialization for numerical weather prediction over the tropics. *Tellus* **43AB**, 53–81.

Lorenz, E. N. 1982. Atmospheric predictability experiments with a large numerical model. *Tellus* **34**, 505–513.

Mahfouf, J.-F., Bauer, P. and Marécal, V. 2003. The assimilation of SSM/I and TMI rainfall rates in the ECMWF 4DVar System. ECMWF Technical Memo 429, ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Phillips, N. A. 1973. Principles of large-scale numerical weather prediction. In: *Dynamic Meteorology* (ed. P. Morel). Reidel, Dortrecht.

Simmons, A. J. and Gibson, J. K. 2000. The ERA-40 Project Plan. ERA-40 Project Report Series No. 1 ECMWF, Shinfield Park, Reading, UK, 63pp.

Simmons, A. J. and Hollingsworth, A. 2002. Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.* **128**, 647–677.

Smagorinsky, J., Miyakoda, K. and Strickler, R. F. 1970. The relative importance of variables in initial conditions for dynamical weather prediction. *Tellus* **22**, 141–157.

Temperton, C. 1976. Dynamic initialization for barotropic and multilevel models. *Q. J. R. Meteorol. Soc.* **102**, 297–311.

Wang, X. L. and Zwiers, F. W. 1999. Interannual variability of precipitation in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *J. Climate* **12**(5), 1322–1335.

White, P. 2000. IFS Documentation Part III: Dynamics and Numerical Procedures (CY21R4). Meteorological Bulletin M1.6/4, ECMWF, Shinfield Park, Reading, UK.