

# Sensitivity of large-scale atmospheric analyses to humidity observations and its impact on the global water cycle and tropical and extratropical weather systems in ERA40

By L. BENGTSSON<sup>1,2</sup>, K. I. HODGES<sup>1\*</sup> and S. HAGEMANN<sup>2</sup>, <sup>1</sup>*Environmental Systems Science Centre, University of Reading, Harry Pitt Building, 3 Earley Gate, Whiteknights, PO Box 238, Reading, UK;*  
<sup>2</sup>*Max-Planck-Institut für Meteorologie, Bundesstrasse 55, D-20146, Hamburg, Germany*

(Manuscript received 2 June 2003; in final form 10 November 2003)

## ABSTRACT

Reanalysis data obtained from data assimilation are increasingly used for diagnostic studies of the general circulation of the atmosphere, for the validation of modelling experiments and for estimating energy and water fluxes between the Earth surface and the atmosphere. Because fluxes are not specifically observed, but determined by the data assimilation system, they are not only influenced by the utilized observations but also by model physics and dynamics and by the assimilation method. In order to better understand the relative importance of humidity observations for the determination of the hydrological cycle, in this paper we describe an assimilation experiment using the ERA40 reanalysis system where all humidity data have been excluded from the observational data base. The surprising result is that the model, driven by the time evolution of wind, temperature and surface pressure, is able to almost completely reconstitute the large-scale hydrological cycle of the control assimilation without the use of any humidity data. In addition, analysis of the individual weather systems in the extratropics and tropics using an objective feature tracking analysis indicates that the humidity data have very little impact on these systems. We include a discussion of these results and possible consequences for the way moisture information is assimilated, as well as the potential consequences for the design of observing systems for climate monitoring. It is further suggested, with support from a simple assimilation study with another model, that model physics and dynamics play a decisive role for the hydrological cycle, stressing the need to better understand these aspects of model parametrization.

## 1. Introduction

Short-term climate studies of the Earth's atmosphere (~20–40 yr) are increasingly being carried out with data sets produced by advanced data assimilation methods which make use of operational Numerical Weather Prediction (NWP) techniques. In order to ensure consistency in the assimilation with respect to the model and assimilation method, such studies are mostly undertaken using reanalyses (Bengtsson and Shukla, 1988). The approach is to run a 'frozen' version of an operational model and data assimilation system in a successive mode (Kalnay et al., 1996; Gibson et al., 1997; Kistler et al., 2001) generating a sequence of comprehensive meteorological fields for an extended period of time. Some of the available fields are directly analysed by the system including surface pressure, temperature, wind and

humidity; other fields consist of derived quantities such as fluxes of water, heat and momentum. These derived quantities consequently depend on the type of model used and are therefore not uniquely determined by the observations. These quantities also depend on different aspects of the assimilation and generally cannot be obtained directly from the initial state. Instead they are normally calculated from model estimates used in the assimilation, such as a six-hourly forecast integrated from the preceding analysis step.

Reanalyses are being used in a multitude of different investigations, including attempts to determine the Earth's hydrological cycle. Due to the absence of rainfall measurements over the oceans and uncertainties in the calculation of evaporation over both land and sea, there are considerable uncertainties in the hydrological cycle. Comparing the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis moisture convergence data with estimated run-off from the Global Precipitation Climatology Project

---

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

(GPCP) the total land annual water budget (moisture convergence over land = runoff) can presently at best be closed to within 10% (Roads, 2002). However, much larger dissimilarities occur over limited regions, for shorter periods and individual processes. This underestimates the error of individual processes that have some cancellation. For example, Adler et al. (2001) showed a precipitation intercomparison which included data from *in situ* networks, satellite observations, and results from numerical modelling. Although the data sets are qualitatively similar in that they all capture the tropical maxima, the subtropical minima and the mid-latitude maxima, there are in certain areas considerable quantitative differences between the observational products. The peak value in the Inter-Tropical Convergence Zone (ITCZ) at 8°N, for example, varies from 1300 to 3200 mm in the annual total.

Several studies to determine the hydrological cycle from NWP data sets have been carried out in recent years; see, for example, Roads (2002) and references therein, and Trenberth and Guillemot (1998). The hydrological cycle (including sources and sinks) is obtained during the data assimilation by an optimized use of the time evolution of the model and available observations, which directly (via humidity data) or indirectly (via the model as driven by winds, temperature and pressure data) determine the water cycle. At the surface, some implicit analysis adjustment occurs for snow correction as well as for surface moisture, which uses either observed precipitation or the model calculated precipitation. Because of these implicit adjustments one may think that reanalyses cannot be used to study the hydrological budget, or, for that matter, budget calculations of energy and water fluxes in general.

Another important question to address is which atmospheric observations are crucial for the determination of the hydrological cycle and which are only of secondary influence. This important question is related to the overall issue of what atmospheric parameters should and can be observed (with present capabilities) and what parameters or quantities are not feasible to observe (because of their very fine structure and very high variability), but should preferably be calculated from models in an assimilating mode.

Here we investigate the relative importance of current atmospheric humidity observations as used in the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr reanalysis (ERA40) system with respect to the indirect forcing of the model dynamics. This entails establishing the degree to which the atmospheric humidity observations influence the different components of the hydrological cycle. This will highlight possible weaknesses in the assimilation system used in ERA40 in making use of the current humidity observations and the fact that there are problems in the current ERA40 system with the simulation of the hydrological cycle associated with the humidity analysis. This should be of interest to users of the ERA40 data and perhaps the wider NWP community. Although our study is dependent on the ERA40 system, we believe this study high-

lights the need for an improved understanding of the relative importance of different observations and a thorough insight into the use of atmospheric assimilation. This will then help in the setting of realistic priorities for the implementation and use of future atmospheric observing systems of relevance to the monitoring of the hydrological cycle. Here we have reasons to assume that current atmospheric humidity data may be of limited use in present assimilation schemes because they have to be consistent with the divergence field in order to be correctly assimilated. This is generally not the case because of insufficient wind (and temperature) observations of sufficiently high quality. The study will also highlight the choice of model in the data assimilation, because in areas where the model is inadequately forced by consistent observations then the model dynamics and physical parametrizations will essentially determine the results. This study will also show that the humidity observations have very little impact on the individual weather systems indicating that the dynamical observations are of more importance; this is particularly the case in the extratropics.

Although the results, and hence conclusions, presented in this paper are dependent on the ERA40 system and the way this system makes use of the available humidity data, we believe that many of the conclusions are of more general applicability although they may be modified to some extent with the introduction of new assimilation techniques and new observations.

The paper continues as follows. In Section 2 we describe the data assimilation and the experimental approach. In Section 3 we present the global results and the way humidity data may influence the global water cycle including differences between land and ocean areas. In Section 4 we explore how humidity data may influence the representation of tropical and extratropical cyclones. In Section 5 the results are discussed and put into context, and finally conclusions are in Section 6.

## 2. The data-assimilation and experimental approach

The assimilation system used for our experiments is the one used in the ERA40 project (Simmons and Gibson, 2000). In the ERA40 reanalysis, past data for the period 1958–2001 are re-assimilated using a frozen version of the ECMWF forecasting system. The version used is based on a three-dimensional variational (3DVar) data assimilation system and a model with a horizontal spectral resolution of T159 and with 60 hybrid levels in the vertical; advanced physical parametrizations are also used.

The main experiment reported here has been set up in the following way. The ERA40 reanalysis system was rerun for two limited periods of time, December–February 1990/91 (DJF9091) and June–August 1991 (JJA91), with all direct humidity observations removed. This means that in this experiment atmospheric humidity is only determined indirectly, as it is forced by the model, which in turn is driven by observations of surface

Table 1. Precipitation over land, P (land), and ocean, P (ocean) for the period December 1990 through February 1991. Units are  $10^{12}$  m<sup>3</sup> of water (mm d<sup>-1</sup> in brackets)

	CMAP	GPCP	ERA40 (control)	ERA40 (no-moisture)
P (land)	23.7 (1.81)	27.1 (2.07)	27.8 (2.15)	28.0 (2.16)
P (ocean)	97.0 (2.96)	92.3 (2.81)	112.5 (3.48)	102.2 (3.16)

pressure, wind and temperature. Humidity observations from surface data and radiosondes are excluded. All Special Sensor Microwave/Imager (SSM/I) data are excluded, as well as channels 11 and 12 for the High-Resolution Infrared Spectrometer (HIRS) data. A possible route for moisture information to enter the system is via the control against supersaturation in the hydrostatic check, which can create some humidity increments. Another route is via indirect changes to other HIRS channels (6, 7 and 8) which may affect humidity. These may take place when these channels are strongly biased with respect to the background state. In such a situation, the 3DVar system is unable to adjust temperature any further, and in order to satisfy the radiance observations, adjustment with respect to humidity may occur (Andersson and Simmons, ECMWF, private communication). This is particularly apparent after the Pinatubo volcanic eruption during the summer of 1991 where the volcanic aerosols impact on the HIRS temperature channels.

In the following we will call the standard ERA40 the ‘control’ experiment, and that without humidity observations the ‘no-moisture’ experiment. For intercomparison and validation we have selected the two well-evaluated precipitation data sets, CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997) and Global Precipitation Climatology Project (GPCP; Rudolf et al., 1996; Huffman et al. 1997), which are available as monthly averages since 1979. CMAP is a combination of satellite data and gauge measurements from more than 6500 land stations (Rudolf et al., 2000). GPCP is similar to CMAP but the precipitation data from the land-based stations have been corrected for a systematic undercatch of the rain gauges, especially when snowfall, in combination with strong winds, occurs. Consequently, winter precipitation at high latitude land areas is higher in GPCP than in CMAP (see Table 1). Precipitation over extratropical oceans is also higher in GPCP. Over tropical land areas the two data sets agree well, but over tropical oceans precipitation from CMAP is larger than from GPCP. Globally averaged, they differ by less than 2% during this period. Over other periods, the differences are typically larger with a long-term average of some 8%.

### 3. Global results

The estimated global precipitation, for land and ocean areas separately, from the two empirical estimates together with ERA40 control and the experiment, are summarized in Table 1 for the DJF9091 period. The different estimates agree reasonably well

over land, but over the oceans the ERA40 control value is some 20% higher than the empirical estimates. While it may be expected that both CMAP and GPCP may underestimate precipitation over the oceans (Trenberth et al. 2001), the ERA40 control precipitation is most likely overestimated and is in fact higher than the corresponding ocean evaporation ( $112.5 \times 10^{12}$  versus  $111.6 \times 10^{12}$  m<sup>3</sup> water for the same period). This may be related partly to adjustment processes in the early part of the integration, and may be different when ensembles of predictions longer than 6 h are used. However, this is a slow process as has recently been demonstrated by Hólm et al. (2002), the ERA40 system requiring some 3–4 d of integration to spin down the overly high initial precipitation. Another cause of the severe imbalance, to be discussed later, is the way satellite moisture data are used in ERA40.

In general, the precipitation from the ERA40 control (Table 1, Figs. 1a and b) is higher, over both tropical land and oceans than the estimates from GPCP and CMAP, respectively, in particular over the ITCZ region, apart from some regions mentioned below. For middle and high latitudes, the ERA40 control results fall between the two observational estimates, the estimates from CMAP being higher than those of GPCP. This behaviour is further highlighted in the zonally averaged precipitation, Figs. 2a and b, for land and ocean respectively, where the peak in the precipitation between 5°S and 5°N is considerably stronger for the ERA40 control than in both CMAP and GPCP, both for land and ocean. For the tropical ocean, the empirical data sets are based on indirect assessment of satellite data such as outgoing long-wave radiation (OLR) and are dependent on the algorithm used to estimate the rainfall so that there is some uncertainty in these estimates. Over the land the differences between the ERA40 control and the empirical estimates are mainly due to lower model generated precipitation over central Africa, the Amazon basin and Borneo where rather few *in situ* rainfall observations exist and to larger model generated precipitation over the Andes. Here, numerical handling of the flow over the mountains may be a contributing factor to very high precipitation. The underestimation of modelled precipitation over central South America is quite pronounced and most likely incorrect, because several *in situ* measurements from the GPCP are higher than the results from the assimilation runs. The comparatively large extratropical difference in the zonally averaged precipitation between 45°S and 55°S over the land comes from the high orographically induced precipitation over southern Chile, although these zonal means will be based on only a few grid points. On the other hand, the

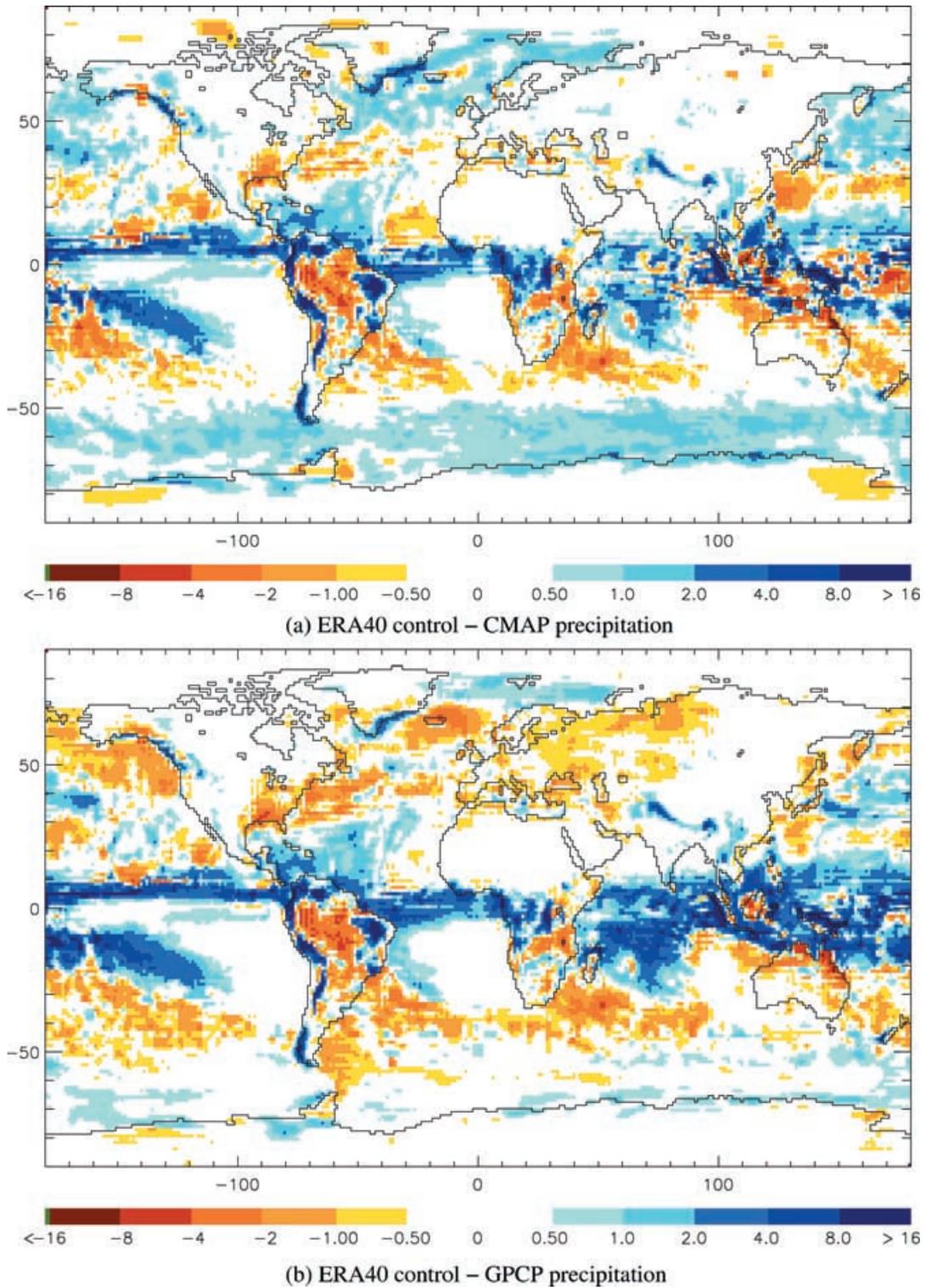


Fig. 1. Precipitation difference for the winter (DJF) 1990/1991: (a) ERA40 – CMAP; (b) ERA40 – GPCP. Units in mm d<sup>-1</sup>.

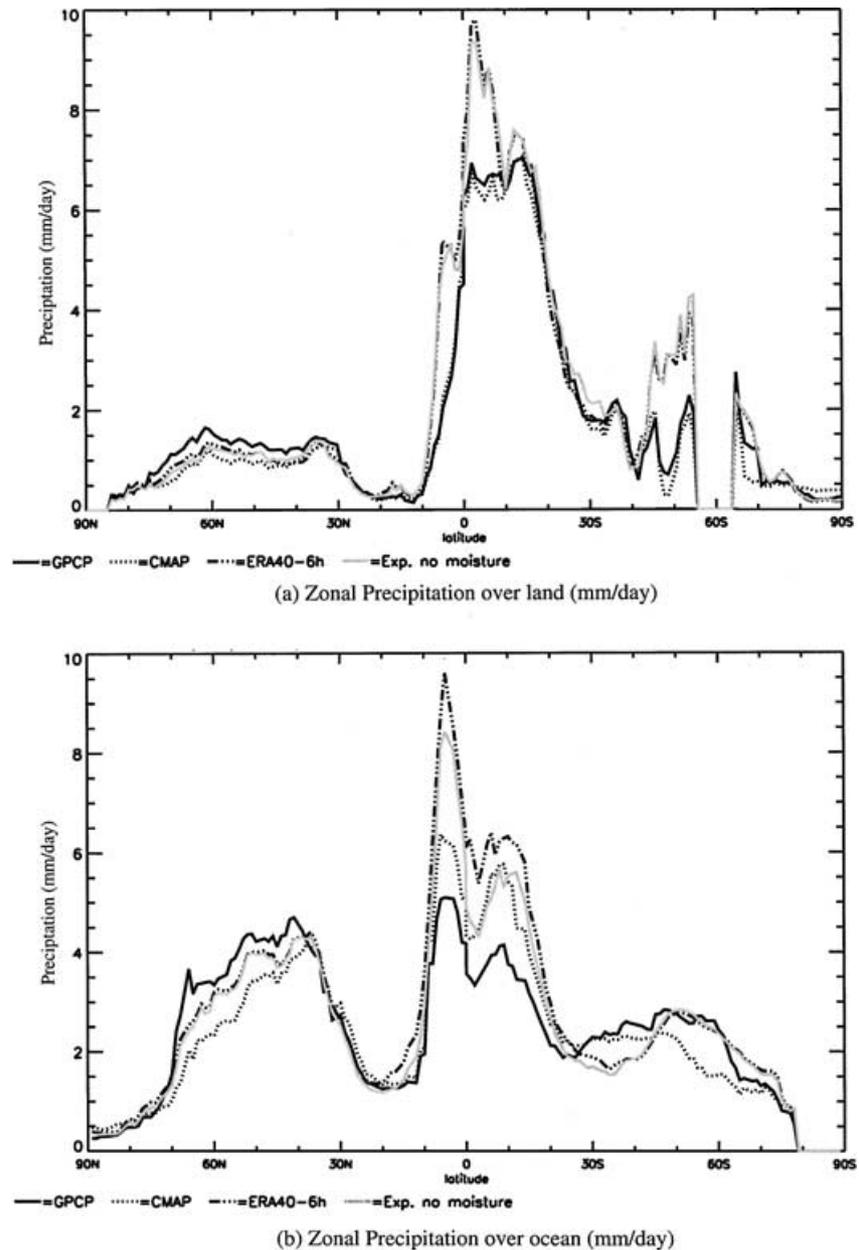


Fig. 2. Zonally averaged precipitation (a) over land for GPCP (full line), CMAP (dotted line), ERA40 control (dash-dotted line), ERA40, no-moisture (grey line) and (b) for the oceans. Unit  $\text{mm d}^{-1}$ .

precipitation appears to be underestimated in both CMAP and GPCP in this region.

The precipitation over land from the ERA40 no-moisture experiment (Table 1, Fig. 2a) is practically identical to the control run, thus showing the same difference relative to CMAP and GPCP. On the other hand, over the ocean regions the precipitation is 10% less than in the ERA40 control run. Polewards of  $30^{\circ}\text{N}$  and  $30^{\circ}\text{S}$  the calculated precipitation by both assimilation runs is practically identical (Fig. 2b). These results suggest that the differences from CMAP and GPCP for the ERA40 control and no-moisture experiment are not related to whether atmospheric humidity observations are included or not.

Evaporation over land is virtually the same in the no-moisture and control runs, while over the ocean the assimilation without moisture is 3.5% higher; the difference is essentially confined to the tropics (Fig. 3). The slightly higher evaporation in the no-moisture run is consistent with a drier atmosphere.

The global water balance for land and ocean separately is summarized in Table 2. As a result of the high ocean precipitation in the ERA40 control run, the global water cycle is unbalanced. This is because the precipitation–evaporation (P–E) is close to zero over the oceans, being slightly negative in December, slightly positive in January and close to zero in February. For land areas P–E is positive, as of course it should be, amounting on average

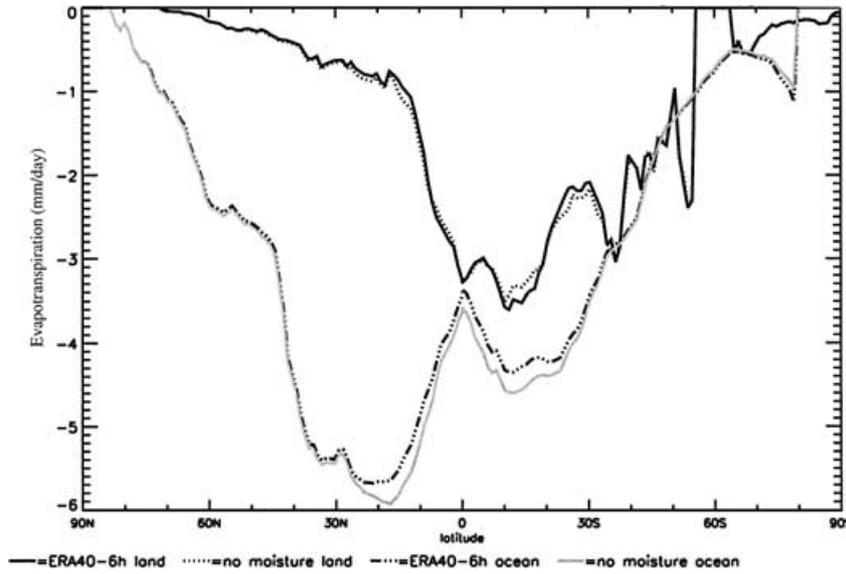


Fig. 3. Zonally averaged evaporation over land and ocean. ERA40 control over land (full line), ERA40 control over ocean (dash-dotted line), ERA40 no-moisture over land (dotted line) and over ocean (grey line).

Table 2. Global water balance for land and ocean and per month. Units as in Table 1

	Total	Dec.	Jan.	Feb.
Land				
P-E ERA40 (control)	12.8	4.3	4.9	3.6
P-E ERA40 (no-moisture)	12.8	4.0	5.1	3.7
Ocean				
P-E ERA40 (control)	0.9	-0.6	1.6	-0.1
P-E ERA40 (no-moisture)	-13.4	-5.1	-4.1	-4.2

to  $4.3 \times 10^{12}$  m<sup>3</sup> water per month in both control and the no-moisture assimilation. For the no-moisture experiment P-E over the ocean is now much more realistic and results in the water cycle between land and ocean balancing rather well, having a net water imbalance of  $-0.6 \times 10^{12}$  m<sup>3</sup> water over the whole winter season (a negative value indicates a land-ocean loss), compared to a net global water imbalance of  $+13.7 \times 10^{12}$  m<sup>3</sup> water for the ERA40 control run, or the same amount as the complete water transport from ocean to land. These results strongly suggest that the ocean precipitation in the ERA40 control run is too high.

Although the main interest in this study is the hydrological cycle we have also investigated the analysed humidity field and the OLR in order to better understand how well the experiment can reconstitute atmospheric humidity. Figure 4 shows the zonal cross-sections for precipitable water content and OLR for the control and no-moisture experiments, respectively. Minor differences can be found in the precipitable water where the control has slightly higher values essentially confined to the equatorial

region. The globally integrated values differ by less than 2% (3–4% in the tropics) and are well within the accuracy of individual observing systems. The high level of similarity between the experiment and the control can be seen in Fig. 5, showing the daily calculated precipitable water at three grid points for February 1991: North Atlantic (55N, 20W); South Indian Ocean (23S, 60E); South Atlantic (60S, 0W).

The zonal structure of OLR is also very similar in the two experiments, with slightly higher values in the tropics for the control experiment, suggesting slightly drier conditions in the upper troposphere at least outside the ITCZ. This seems to indicate a stronger tropical meridional circulation in the control experiment. Globally averaged values differ by  $1.6 \text{ W m}^{-2}$ , and again are well within the limits of the available estimates (Kiehl and Trenberth, 1997).

A comparison between the ERA40 control and no-moisture experiment has also been explored in the context of the general circulation. Figures 6a and b show the difference in the mean geopotential height (in metres) at the 500- and 50-hPa levels ( $Z_{500}$  and  $Z_{50}$ ). Differences are typically very small, mostly less than 1 m over the globe, but with some regions having slightly higher differences but no more than 5 m, around Antarctica, and over Siberia at the 50-hPa level. Why these larger values are attained in specific regions and why they are related to the humidity data is unclear. A more detailed exploration of the difference between the two experiments is provided by the root mean square (rms) difference between various fields from the two experiments. Figures 6c and d show the rms difference for  $Z_{500}$  and the relative vorticity at 500 hPa ( $\xi_{500}$ ), respectively, normalized by the standard deviation of the relevant field from the no-moisture experiment.  $Z_{500}$  shows that the instantaneous differences are very small throughout the globe and practically zero in the extratropics. Differences are higher over the tropical

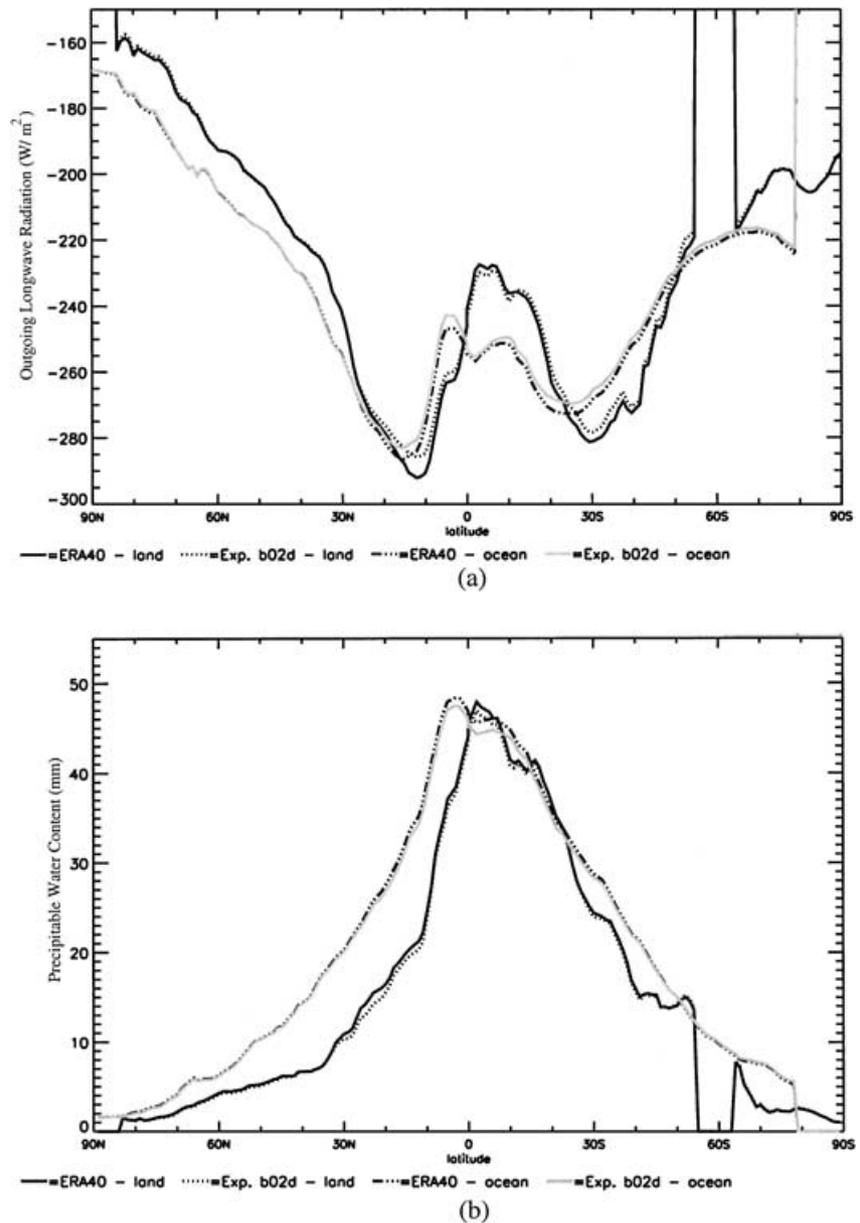


Fig. 4. (a) OLR ( $\text{W m}^{-2}$ ) and (b) precipitable water content (mm), for ERA40 control and the no-moisture experiment over land and ocean for the winter (DJF) 1990/1991.

regions, in particular the land regions of Africa and South America but these are much less than unity.  $\xi_{500}$  also shows small differences in the extratropics although these are higher than in the  $Z_{500}$  field, consistent with  $\xi_{500}$  focusing on smaller scale variability. In fact, this field now shows the differences in the SH extratropics to be slightly higher than in the Northern Hemisphere (NH). This is consistent with the results that are discussed next on the weather systems. The largest differences, however, occur through the tropics with values often greater than unity, in particular throughout most of the regions of deep convection of the ITCZ and in the same regions previously discussed for  $Z_{500}$ . Although these differences can be large, they appear not to be related to organized weather systems (see next section) but

are more likely related to random differences associated with the convection. These results on the general circulation are consistent with the results in the next section.

#### 4. Tropical and extratropical cyclones

An analysis of the intensity and trajectory of tropical and extratropical cyclones, which are likely to be influenced by the moisture data, has been performed. This is based on the methodology described in Hoskins and Hodges (2002) for the extratropical cyclones and Thorncroft and Hodges (2001) for the tropical cyclones. For the extratropical activity, the tracking has been performed on the mean sea level pressure (MSLP) field

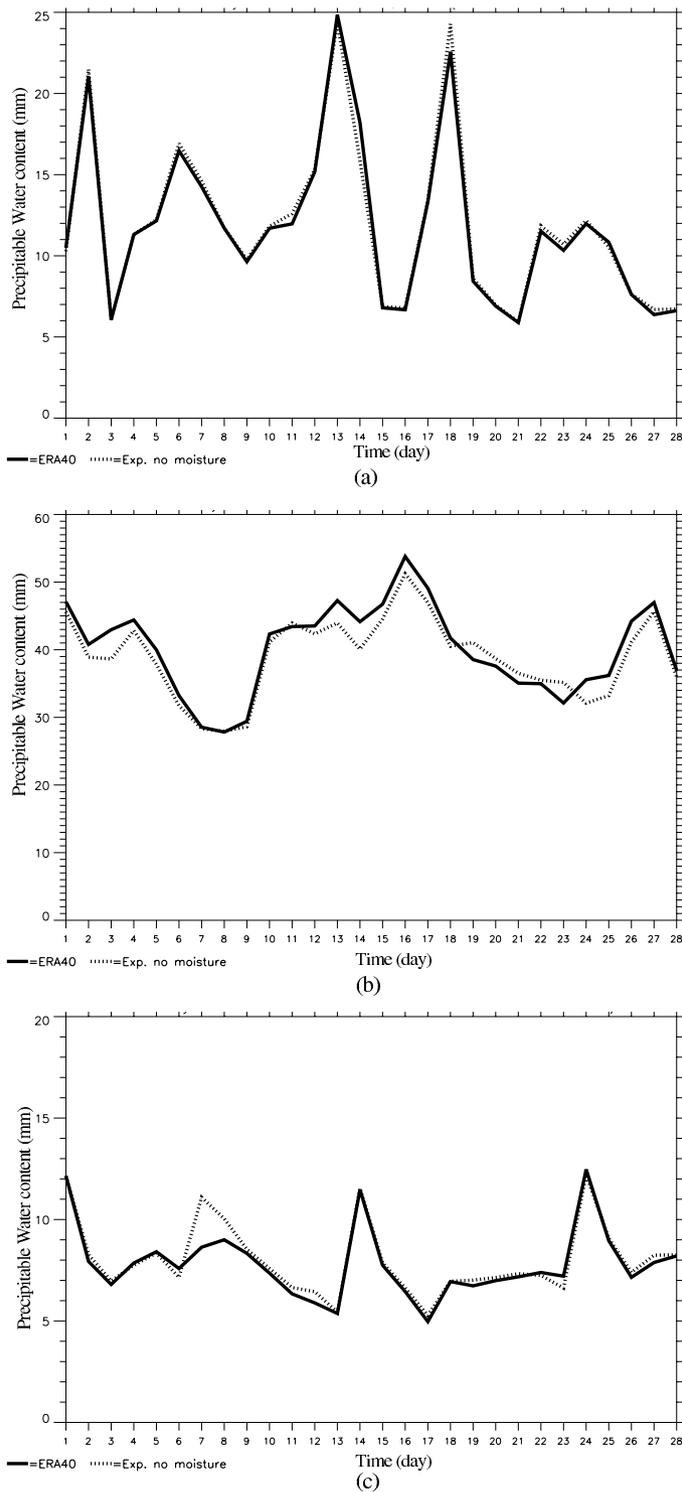


Fig. 5. Analysed precipitable water for the experiment and control (ERA40) at three grid points for February 1991: (a) 55N, 20W; (b) 23S, 60E; (c) 60S, 0W. Values are in kg water m<sup>-2</sup>.

as well as the relative vorticity on the 850-, 500- and 250-hPa ( $\xi_{850}$ ,  $\xi_{500}$ ,  $\xi_{250}$ ) levels. Note that for the extratropical activity the fields have the planetary scales removed before identification and tracking are performed as described by Hoskins and

Hodges (2002). This makes the systems easier to identify, in particular for MSLP. For the tropical cyclone activity the tracking was performed on the  $\xi_{850}$  field only. For both the extratropical and tropical analysis, only those systems that last at least 2 d and

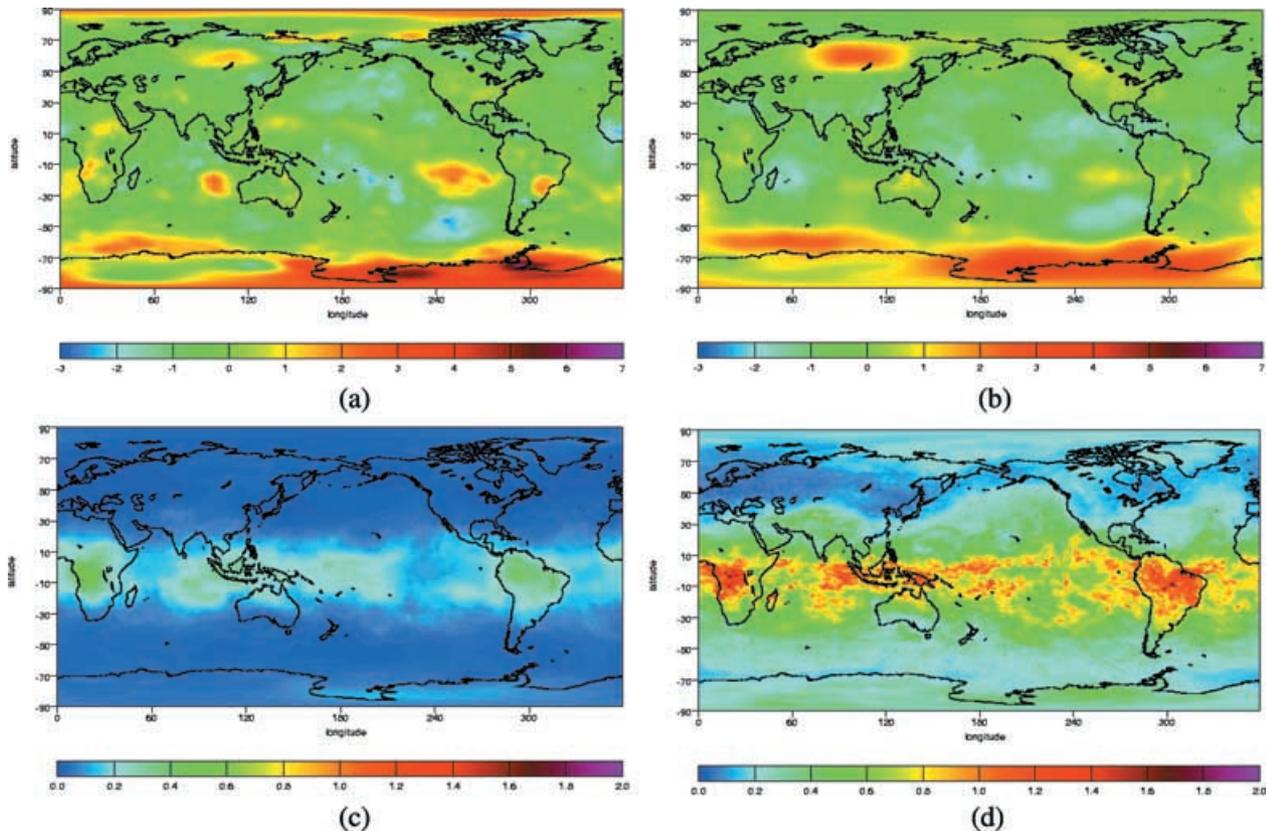


Fig. 6. (a) Difference in metres of the mean  $Z_{500}$  (no-moisture – control). (b) Difference in metres of the mean  $Z_{50}$  (no-moisture – control). (c) The rms difference between the no-moisture and control experiments for  $Z_{500}$ , normalized by the standard deviation of the no-moisture experiment. (d) Same as (c) but for  $\xi_{500}$ .

travel further than 1000 km are retained. Because results for only one winter are available, the generation of spatial statistics such as those produced in Hoskins and Hodges (2002) (which were based on results for 20 NH winters) for extratropical cyclones will not be statistically meaningful and are not considered. Instead we have used the approach of Hodges et al. (2003) for the direct comparison of track ensembles. This compares two track ensembles track by track to produce statistics. This approach is used for the comparison of the extratropical track ensembles for the MSLP and relative vorticity fields. The matching parameters for what constitutes a good match are an overlap in time of at least 60% of the points and a mean geodesic separation, computed for those points that match, of less than  $1.0^\circ$ ; see Hodges et al. (2003) for further details. In general, the results for MSLP focus on the larger end of the spatial synoptic scale range, whilst the relative vorticity focuses on the smaller spatial scales and results in many more systems being identified. The summary statistics are shown in Table 3.

For the NH, Table 3 shows the agreement between the track ensembles for the ERA40 control and no-moisture experiment. For MSLP this shows that the agreement between the tracks is very good with the number of tracks that match being greater

than 90% and with the number of tracks that match with greater than 95% of their points being greater than 80%. As was seen in Hodges et al. (2003), the tracks that do not match, or that match with fewer than 95% of their points, tend to correspond to the weakest systems which are more sensitive to the available observations and the way these are assimilated. This can be seen in Fig. 7a, which shows the distributions, in terms of the mean track intensities (averaged along a track) for NH MSLP, for the tracks that match and those that do not match. The fact that for the two ensembles the distributions for the tracks that do match are very similar indicates that there is very little difference in the system intensities on a point-by-point basis; a separate point-by-point comparison confirms this. This situation is also reflected in the  $\xi_{850}$  field (Fig. 7c) where there are now many more systems identified, reflecting the smaller spatial scale nature of these systems. This also shows a good correspondence between the two track ensembles with  $\sim 85\%$  matching well and more than 60% matching for greater than 95% of their points. As we go to higher levels where there are fewer observations this view persists, although there is some degradation in the percentage of systems that match for greater than 95% of their points. The distributions of those systems that match for  $\xi_{850,500,250}$  between the ERA40

*Table 3.* Summary statistics for extratropical cyclone matching in the NH and SH (percentages of totals are in brackets). The keyword ‘ERA40 total’ indicates the total number of systems identified for the ERA40 control and no-moisture experiment; ‘Match’ indicates the number of systems that match between the two track ensembles; ‘No Match’ denotes the number of systems that do not match; ‘Match >95%’ are the number of systems that match for greater than 95% of their points

	MSLP	$\xi_{850}$	$\xi_{500}$	$\xi_{250}$
NH				
ERA40 total (control)	165	439	564	433
ERA40 total (no-moisture)	163	440	548	432
Match	154 (94)	379 (86)	461 (82.9)	353 (82.6)
No match ERA40 (control)	11 (6.7)	60 (13.7)	103 (18.3)	80 (18.5)
No match ERA40 (no-moisture)	9 (5.5)	61 (13.8)	87 (15.9)	79 (18.3)
Match >95%	138 (84)	278 (63.2)	269 (48.4)	235 (54.3)
SH				
ERA40 total (control)	145	359	456	427
ERA40 total (no-moisture)	153	368	453	441
Match	134 (90)	275 (75.6)	335 (73.7)	341 (78.6)
No match ERA40 (control)	11 (7.6)	84 (23.4)	121 (26.5)	86 (20.1)
No match ERA40 (no-moisture)	19 (12.4)	93 (25.3)	118 (26.0)	100 (22.7)
Match >95%	105 (70.5)	149 (41)	174 (38.3)	170 (39.2)

control and the no-moisture experiment, as for MSLP, show good agreement indicating that there is little difference in the individual system intensities between the two track ensembles; as with MSLP a point-by-point comparison confirms this. These results indicate that the current water vapour observations and the way they are assimilated in ERA40 have a minimal effect on the representation of the synoptic weather systems in the extratropics, indicating the dominance of the dynamics over the water vapour field at mid to high latitudes. This situation might change with the use of a different assimilation system such as 4DVar where the humidity observations will have a stronger feedback on to the dynamical fields (Hólm et al. 2002).

In the Southern Hemisphere (SH) the results indicate a similar picture, although with slightly fewer systems as a percentage of the total providing a good match. The slightly poorer results, in terms of the number of systems that match well between the ERA40 control and the no-moisture experiment for all fields, may reflect the fact that there are relatively few ground-based observations of winds and temperature so that the assimilation is more reliant on satellite observations.

In terms of tropical cyclones, the tropical activity for this period is confined to the SH. However, for this SH summer period

there are relatively few tropical cyclones. We have identified several of the tropical cyclones in this period in both the control and no-moisture experiment and compared them with the best track data from the Data Support Section of the Scientific Computing Division at NCAR. These are shown in Fig. 8 for tropical cyclones Joy, Chris, Daphne and Bella, and indicate that both the ERA40 control and no-moisture experiment have both captured these tropical cyclones quite well. The main differences in the cyclone tracks occur at the beginning and end of the storm life cycles when the storms are quite weak and are more sensitive to the observations used. In terms of the maximum intensities, Table 4 shows the maximum attained intensity together with the date this occurred. These results show that, whilst three of the storms are marginally less intense in the no-moisture experiment, with one more intense these differences are quite small, typically less than  $1.0 \times 10^{-5} \text{ s}^{-1}$ . The times at which the maximum intensity is attained are nearly identical for the ERA40 control and no-moisture experiment for all the cyclones. However, comparing with the observed times at which the cyclones reached their maximum intensity in terms of wind speed there are some differences. This is only a small sample study of tropical cyclones so no general conclusions should really be considered.

The results for tropical cyclone Bella highlight some of the problems in tracking tropical cyclones when they are weak disturbances and are more sensitive to the assimilation. The problems are highlighted by the labels 1, 2 and 3. These indicate a break in the track due to the generation of a multiple centre which we have fixed manually; this occurs twice for the ERA40 control (green, labels 1 and 2) and once for the ERA40 no-moisture experiment (red, label 3). Ultimately, this problem will be fixed by using a more objective method of merging tracks in the tracking algorithm.

## 5. Discussion

There are two important results from this study which need to be highlighted. The first point to make is the very limited contribution from humidity observations in general in ERA40. This does not mean that moisture observations are unimportant per se, but, as we have shown here, a comprehensive data assimilation system is able to reconstitute the moisture field to a considerable degree from the dynamics of the large-scale model and from the sources and sinks of water vapour in the model. We have explored the individual maps and the representation of synoptic weather systems in both the tropics and extratropics. As demonstrated above there are hardly any noticeable differences in the weather systems identified in surface pressure fields or in the different tropospheric vorticity fields consistent with the general circulation statistics also discussed above. Where we do see differences between the control and no-moisture experiment is in the global hydrological cycle and precipitation over the tropical oceans where it in fact appears that the no-moisture experiment

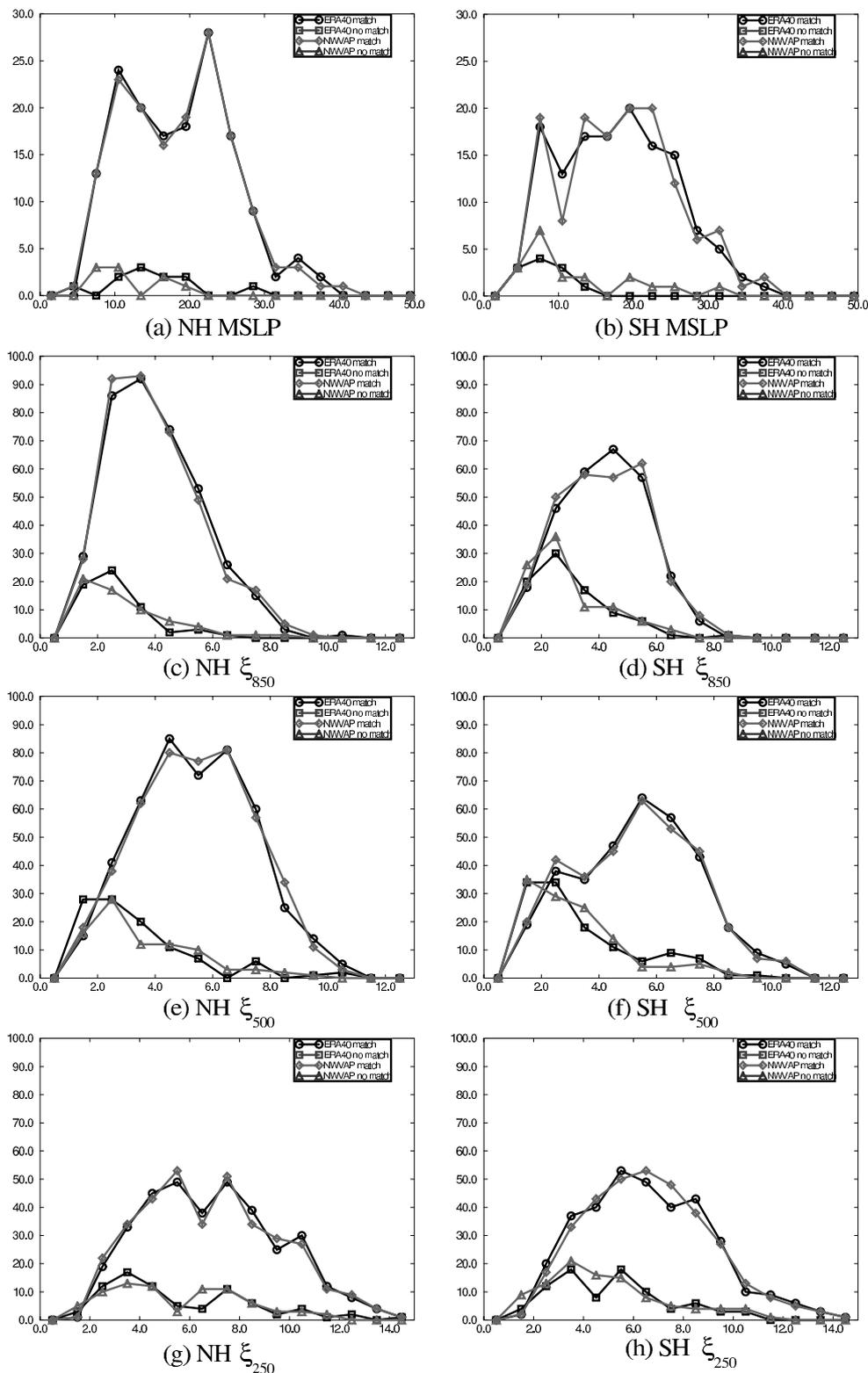


Fig. 7. Tracking statistics of extratropical depressions for the NH and SH with respect to surface pressure and relative vorticity at three different levels (850, 500, 200 hPa). Number of weather systems as a function of mean intensity, units of hPa for MSLP and  $\times 10^{-5} \text{ s}^{-1}$  for vorticity relative to the background field removed. Matching and non-matching systems separately indicated.

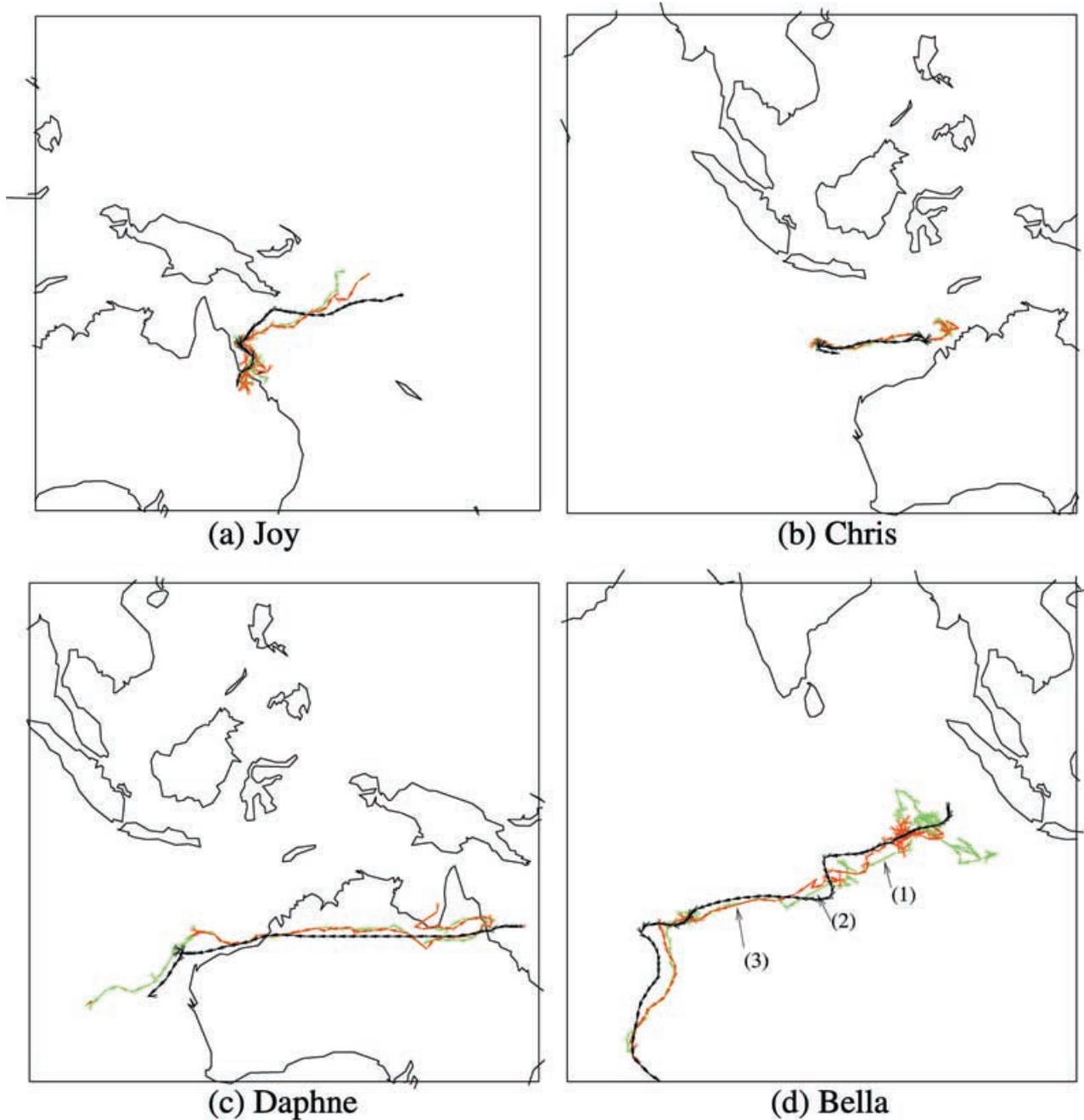


Fig. 8. Tracks for four tropical cyclones identified during the period. Best track data obtained from the Data Support Section at NCAR (black), ERA40 control (green) and from the no-moisture experiment (red): (a) Joy; (b) Chris; (c) Daphne; (d) Bella.

is more credible than the control assimilation, because the global water balance is practically balanced in the experiment but not so in the control assimilation. We have explored this further in a separate experiment for December 1990 whereby we removed all the SSM/I and HIRS data in the ERA40 control to test the individual effect of these observations on the atmospheric humidity. This experiment resulted in a reduction of the overly high ocean

precipitation by  $3 \times 10^{12} \text{ m}^3$  water ( $\sim 10\%$ ) and increasing ocean evaporation by  $1.3 \times 10^{12} \text{ m}^3$  water ( $\sim 3\%$ ), resulting in close agreement with the experiment without moisture observations. It therefore seems that it is the way moisture data from the satellite observing systems are used that is the main contributing factor to the high ocean precipitation and imbalance of the global hydrological cycle.

*Table 4.* Maximum intensities and the dates at which they are attained for the four identified tropical cyclones. Intensities are in units  $10^{-5} \text{ s}^{-1}$  and dates are in the format YYMMDDHH

	ERA40 (control)	ERA40 (no-moisture)
Joy	-14.2 (90122500)	-13.7 (90122418)
Chris	-13.4 (91021800)	-12.5 (91021800)
Daphne	-12.6 (91022100)	-11.5 (91022018)
Bella	-12.5 (91020112)	-13.2 (91020112)

These findings are supported by Hólm et al. (2002) that the excessive tropical precipitation in the ERA40 assimilation system is due to the tropical circulation in the assimilation cycles being too intense, which leads to a tendency for the descent regions to be too dry. The clear sky observations of humidity from the satellites that become assimilated in the descent regions then feed into the precipitation ascent region, generating too much precipitation, and maintaining the overly strong circulation. This is consistent with the differences in OLR and precipitable water content as shown in Fig. 4.

The fact that a comprehensive model driven by observed boundary conditions and atmospheric dynamics through the time evolution of surface pressure and the vertical profiles of horizontal winds and temperatures is a suitable way to calculate atmospheric fluxes has long been recognized (Charney et al. 1969). The general message from this finding is that dynamical variables are fundamentally important in weather prediction and in the reconstitution of the general circulation of the atmosphere by means of a comprehensive forecasting system. In general, observations of pressure, wind and temperature are difficult to observe because of the cost and technical difficulties. In fact, cloud and moisture are generally easier to observe, at least from space-based observing systems. Yet, if we use observations of moisture in data assimilation it is necessary to also adjust simultaneously pressure, wind and temperature in a way that is consistent with the moisture field, so that the divergence pattern set up by the dynamical variables does not destroy the presumed correctly observed moisture field.

This study also highlights the key role of the atmospheric model in the reconstitution of the hydrological cycle. The actual model used in the data assimilation probably influences the hydrological cycle more than the moisture observations. In order to estimate the influence of the choice of model on the hydrological cycle, we have dynamically adjusted the European Centre/Hamburg Model (ECHAM4, version 4.5; Roeckner et al. 1996) towards the ERA40 analysed fields of surface pressure, temperature and wind by means of a so-called ‘nudging’ technique (Jeuken et al. 1996). The ECHAM4 model is successively adjusted towards the ERA40 analyses at every time-step of the integration (interpolated from every 6 h), in such a way that the difference in surface pressure, wind and temperature after a short period of adjustment stays well within the observational

*Table 5.* The hydrological cycle integrated over all land and ocean areas, respectively, for the period December 1990 to February 1991. Units are  $10^{12} \text{ m}^3$  water. For further information see text

	ERA40 (control)	ERA40 (no-moisture)	ECHAM4 (nudged)	ECHAM4 (free run)
P (land)	27.8	28.0	21.3	24.1
P (ocean)	112.5	102.2	95.7	94.5
E (land)	15.0	15.2	12.5	14.2
E (ocean)	111.6	115.6	104.9	105.4
P-E (land)	12.8	12.8	8.8	9.9
P-E (ocean)	0.9	-13.4	-9.2	-10.9
P-E (l + o)	13.4	-0.6	-0.4	-1.0

accuracies of these variables. Forced in such a way, the result is very similar to a continuous set of analyses of the same model using observations of surface pressure, temperature and wind. The result can be considered more or less as a no-moisture assimilation with the ECHAM model. The results, together with the two ERA40 assimilation experiments, and the free run with ECHAM4 (only using the same sea surface temperature) are summarized in Table 5.

We first note that the ECHAM4 nudged and free runs generate a very similar hydrological cycle, albeit a weaker one over land for the nudged version. Over ocean areas the difference is about 1% for both precipitation and evaporation. The reduced similarity over land is probably related to inherent inconsistency in the ‘nudged’ run, because for example there is no feedback between the land surface conditions and the atmospheric fields.

A second observation is the more intense hydrological cycle in the ERA40 control and no-moisture runs compared to ECHAM. In the ERA40 no-moisture run, which is based on the same observations as the nudged ECHAM run, the precipitation over land is 20% higher and over ocean is ~10% higher. All four experiments, except the ERA40 control, balance the global water cycle as well as could be expected in view of the short integration time.

The energy balance for the ERA40 assimilation runs has also been calculated. The differences are small and well within the uncertainty limits. The global surface thermal radiation and latent heat flux are 2% and 3% larger, respectively, in the no-moisture experiment while the sensible heat flux is 2% smaller. The largest differences occur over ocean areas. As a result the ocean net warming amounts to  $23 \text{ W m}^{-2}$  in the control run and  $19 \text{ W m}^{-2}$  in the no-moisture run. A net ocean warming is expected due to the season. We have also compared these flux components from the ERA40 control and no-moisture experiment with a corresponding calculation with the ECHAM4 model, both in a free and ‘nudged’ mode. The different flux components are smaller for ECHAM than for both the ERA40 control and no-moisture experiment by some 10–20%, although the net ocean warming is  $16 \text{ W m}^{-2}$  and thus closer to the no-moisture run.

In concluding this discussion, it is important to also clarify that we do not want to suggest that satellite observed moisture fields are unimportant in the simulation of the general circulation of the atmosphere or for the representation of weather systems, but that currently observed moisture data are of insufficient detail and are assimilated in ERA40 in a way that is inconsistent with the fundamentally important dynamical variables. For example, the HIRS weighting functions for channels 11 and 12 are very broad in the vertical, peaking at 700 and 500 hPa, respectively. This gives very poor vertical resolution of the moisture fields. In addition to poor resolution in the vertical as well as horizontally, the HIRS satellite observed water vapour is limited to cloud-free regions so that in the tropical ITCZ region, for example, the water vapour field is poorly observed by this instrument. For SSM/I, vertical and horizontal resolutions are also poor. Recent work has shown that the assimilation of high-resolution vertical profiles of water vapour provided by active instruments such as airborne lidar (only in cloud-free regions) can have a significant impact on the representation of weather systems and in particular tropical cyclone forecasts (Kamineni et al., 2003). The difficulty with this type of data is its lack of global coverage so that it is limited to particular regions such as the Atlantic coast of North America. New passive instruments may provide better observations of the atmospheric water vapour such as the new Atmospheric Infrared Sounder (AIRS) on the recently launched *AQUA* satellite, although resolution and accuracy are still poor compared to the lidar observations. Also the introduction of the assimilation of data from cloudy regions such as rainfall rates, from the Tropical Rainfall Measuring Mission (TRMM), and/or cloud and rain affected radiances using 4DVar will also hopefully lead to a significant improvement in the simulation of moist processes and the hydrological cycle. Improvements in the data assimilation systems will also be required to handle these data, which can vary rapidly in space and time, as well as improvements in the parametrization of moist processes in the model. A good discussion of all these issues can be found in Hölm et al. (2002) including the introduction of a new formulation for the assimilation of humidity data, which provides a stronger coupling between the humidity data and the dynamical variables. However, these improvements are likely to have an impact in NWP rather than reanalyses in the near future. In particular, the introduction of ever-increasing new observing systems adds to the temporal data inhomogeneity, which can often lead to spurious climate trends that should ideally be avoided in reanalyses.

## 6. Concluding remarks

We have shown that the ERA40 assimilation system is able to generate a credible global hydrological cycle without the use of humidity observations, and that there are only small differences in the dynamical fields and the hydrological cycle at high latitudes from the ERA40 control assimilation using all available

observations. The differences between the two assimilation cycles are well within the bounds of present empirical estimates. In fact, some aspects of the no-moisture assimilation are more realistic than in the control assimilation, which has a major error in the global water cycle, because the oceans do not provide a net source of water for the atmosphere. We believe that this deficiency, as supported by a limited experiment, is related to the way SSM/I and HIRS data are used in the ERA40 control assimilation.

We have restricted this study to the winter period 1990/91 because we believe the result is representative enough. However, an identical study has also been completed for the summer period 1991, with very similar results, in particular for the water balance. Zonal cross-sections of the integrated water vapour and OLR are in all respects consistent with the winter results.

In addition to the investigation of the extratropical and tropical cyclones, we have compared daily changes in precipitable water, precipitation and evaporation at individual grid points. These are practically identical for extratropical regions but differ slightly in the tropics (not shown). Systematic differences are consistent with the zonal average results.

Determination of the hydrological cycle using the ECHAM4 climate model at comparable resolution using a nudging technique provides a different estimate of the hydrological cycle. Only surface pressure, temperature and wind fields from ERA40 were used for the nudging. The result is a weaker hydrological cycle of some 10% than the ERA40 no-moisture run. The hydrological cycle of the nudged ECHAM4 run is rather similar to the ECHAM4 free run, which is only constrained by the sea surface temperature. The result indicates a strong model dependence, and that observed humidity in particular has little influence in determining the global hydrological cycle from data assimilated fields at present.

The objective of the present study was to obtain better insight into the way the ERA40 reanalyses are able to reproduce the global water cycle and the relative importance of moisture observations in this respect and their influence on extratropical and tropical weather systems. We believe this study adds to this understanding and will be of value for the users of the ERA40 data set. Will these results also be of importance for NWP? We believe this may be the case as well, at least until the improvements being explored for the assimilation of humidity data and the introduction of new observations are demonstrated. The very small differences between the two assimilations in the extratropics suggest that the error growth will be rather similar and any significant differences will be difficult to demonstrate against the background of growing unpredictable noise. The differences in the tropics are detectable and could presumably influence tropical weather forecasts. The importance here though is to ensure a better consistency between model dynamics and the assimilation of moisture observations. Some inconsistencies in the assimilation system may be more detrimental than omitting the moisture observations, although higher resolution and more

accurate observations of the moisture field together with the new assimilation methodologies are likely to be required before the assimilation of water vapour will be of benefit to reanalyses and NWP.

The results of this study raises a number of fundamental questions and issues, as follows.

What is the best data assimilation approach that makes the best use of the information content available in observations of humidity, and in atmospheric hydrological information in general? Clearly this is a very active area of development in the operational centres such as ECMWF.

Secondly, the result of this study suggests that pressure, wind and temperature data are perhaps the most important data for the determination of the hydrological cycle when integrated into an advanced data assimilation system. This may be important for the setting of priorities of future observing systems concerned with large-scale weather and climate prediction and the determination of the hydrological cycle of the Earth as opposed to NWP where priorities are not necessarily the same.

Thirdly, because the choice of model appears to be crucial it will be necessary to identify those aspects of the model that are most important in the determination of surface fluxes. Clearly, several reanalysis exercises are called for. Scientists concerned with diagnostic studies, or who wish to have estimates of fluxes for ocean and land surface modelling, are advised to undertake calculations using more than one data set due to the uncertainties in these quantities.

The results in this study are of further interest when trying to assess how climate may have changed over the last century. The reanalysis data now cover more than 50 yr, during which time the global observing systems have undergone substantial changes. It is our intention to undertake similar studies as described here by reducing the present observing system towards a system typical of the pre-satellite era. Whilst the radiosonde system has undergone substantial changes both with respect to networks and sounding equipment, it will nevertheless be possible to estimate how our knowledge of the general circulation of the atmosphere is related to changes in the observing systems. Such investigations have started and will be reported in a future study.

## Acknowledgments

The authors would like to thank ECMWF and especially the ERA40 research team for making the ERA40 system available to us, and the computer support staff for their help in operating the system. The experiment has been undertaken as an ECMWF special project study.

## References

Adler, R. F., Kidd, C., Petty, G., Morrissey, M. and Goodman, H. M. 2001. Intercomparison of global precipitation products: the third pre-

- cipitation intercomparison Project (PIP-3). *Bull. Am. Meteor. Soc.* **82**, 1377–1396.
- Bengtsson, L. and Shukla, J. 1988. Integration of space and in situ observations to study global climate change. *Bull. Am. Meteor. Soc.* **69**, 1130–1143.
- Charney, J., Halem, M. and Jastrow, R. 1969. Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.* **26**, 1160–1163.
- Gibson, J. K., Källberg, P., Uppala, S., Hernandez, A., Nomura, A. and Serrano, E. 1997. Era description. In: *ECMWF Reanalysis Project Report Series 1*. ECMWF, Reading, UK.
- Hodges, K. I., Hoskins, B. J., Boyle, J. and Thorncroft, C. 2003. A comparison between recent re-analysis data sets using objective feature tracking: storm tracks and tropical easterly waves. *Mon. Wea. Rev.* **131**, 2012–2037.
- Hólm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A. and Thépaut, J.-N. 2002. *Assimilation and Modelling of the Hydrological Cycle: ECMWF's Status and Plans*, ECMWF Technical Report 383. ECMWF, Reading, UK.
- Hoskins, B. J. and Hodges, K. I. 2002. New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.* **59**, 1041–1061.
- Huffman, G. J., Adler, R. F., Arkin, A., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., Joyce, R. J., McNab, A., Rudolf, B., Schneider, U. and Xie, P. 1997. The Global Precipitation Climatology Project (GPCP) combined precipitation data set. *Bull. Am. Meteor. Soc.* **78**, 5–20.
- Jeuken, A. B. M., Siegmund, P. C., Heijboer, L., Feichter, J. and Bengtsson, L. 1996. On the potential of assimilating meteorological analyses into a global climate model for the purpose of model validation. *J. Geophys. Res.* **101**, 16939–16950.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Leetmaa, A., Reynolds, R. and Jenne, R. 1996. The NCEP/NCAR reanalysis project. *Bull. Am. Meteor. Soc.* **77**, 437–471.
- Kamini, 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.
- Kiehl, J. T. and Trenberth, K. E. 1997. Earth's annual global mean energy budget. *Bull. Am. Meteor. Soc.* **78**, 197–208.
- Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H. and Fiorino, M. 2001. The NCEP-NCAR 50-year reanalysis: monthly means cd-rom and documentation. *Bull. Am. Meteor. Soc.* **82**, 247–268.
- Roads, J. 2002. Closing the water cycle. *GEWEX Newsletter* **12**, (1). International GEWEX Project Office, MA, USA, 8 pp.
- Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U. and Schulzweida, U. 1996. *The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate*, Rep. 218, Max-Planck-Inst. for Meteorol., Hamburg, Germany.
- Rudolf, B., Hauschild, H., Rütth, W. and Schneider, U. 1996. Comparison of rain gauge analyses, satellite-based precipitation estimates and forecast model results. *Adv. Space. Res.* **7**, 53–62.
- Rudolf, B., Gruber, A., Adler, R., Huffman, G., Janowiak, J. and Xie, P. 2000. GPCP precipitation analyses based on observations as a basis of NWP and climate model verification. In: *Proceedings*

- of the 2nd WCRP International Conference on Reanalyses, Reading, UK, 23–27 August 1999, WCRP-109, WMO/TD-No. 985, 197–200.
- Simmons, A. J. and Gibson, J. K. 2000. *The ERA40 Project Plan, ERA40 Project Report Series No. 1*. ECMWF, Reading, UK, 63 pp.
- Thorncroft, C. and Hodges, K. I. 2001. African easterly wave variability and its relationship to atlantic tropical cyclone activity. *J. Climate* **14**, 1166–1179.
- Trenberth, K. E. and Guillemot, C. J. 1998. Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. *Climate Dyn.* **14**, 213–231.
- Trenberth, K. E., Stepaniak, D. P., Hurrell, J. W. and Fiorino, M. 2001. Quality of reanalyses in the tropics. *J. Climate* **14**, 1499–1510
- Xie, P. and Arkin, P. 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull. Am. Meteor. Soc.* **78**, 2539–2558