Validation of precipitation from ERA40 and an ECHAM4.5 simulation nudged with ERA40 data

Stefan Hagemann, Klaus Arpe, Lennart Bengtsson and Ingo Kirchner

Max Planck Institute for Meteorology
Hamburg, Germany

Abstract

The European Centre for Medium-Range Weather Forecasts (ECMWF) is currently preparing a new 40 year reanalysis dataset (ERA40). For the years available at the time of this study (1958, 1987-92), the ERA40 precipitation is compared to the previous 15 years ECMWF reanalysis (ERA15) and several observational datasets. The ERA40 precipitation has changed in several respects compared to ERA15 and the on-going validation yielded that the ERA40 precipitation has several deficiencies. The global water budget is not only unbalanced but also precipitation minus evaporation over the ocean is negative in the long term mean. This seems to be related to an overestimation of precipitation over the ocean, especially in the tropics. Although the spin-up of the hydrological cycle is reduced in ERA40 compared to ERA15, it is still a problem as it increases the global imbalance of the water budget.

In order to derive more realistic precipitation fields from ERA40, a simulation with the GCM ECHAM4.5 (*Roeckner et al.*, 1996) was conducted where the dynamics were nudged to the ERA40 data. Results of this simulation suggest that the overestimated ERA40 precipitation may also have influenced the ERA40 circulation.

1. Introduction

The European Centre for Medium-Range Weather Forecasts (ECMWF) is currently preparing a new 40 year reanalysis dataset (ERA40). The dataset is planned to be constructed in three streams to be started at different times: stream 1 for the years 1987-2001, stream 2 for the years 1957-1972 and stream 3 for the years 1972-1988. The final ERA40 dataset is expected to be a major dataset for climate research. Within the ERA40 project, the MPI (Max Planck Institute for Meteorology) has the task to perform a validation of the hydrological cycle.

In this study we focus on the validation of the ERA40 precipitation which is described in Sect. 2. Here, we will compare the ERA40 precipitation to the previous 15 years ECMWF reanalysis (ERA15; *Gibson et al.*, 1997) and several observational datasets. If not mentioned otherwise the 6 hour forecasts are considered in the following.

In order to derive more realistic precipitation fields from ERA40, a simulation with the GCM ECHAM4.5 (*Roeckner et al.*, 1996) was conducted where dynamical values were nudged (see, e.g., *Jeuken et al.*, 1996) towards the ERA40 data. Results of this simulation are considered in Sect. 3.

2. Precipitation of ERA40

For stream 1, the ERA40 system was changed several times during the production of the ERA40 data for the years 1987-88. Thus, only the years 1989-92 are considered for an overview of the precipitation characteristics if not mentioned otherwise. With regard to the validation of the ERA40 precipitation over the ocean it has to be mentioned that the precipitation observations over the ocean are very uncertain since almost no direct measurements exists. In the precipitation datasets of HOAPS (Hamburg Ocean-Atmosphere

Parameters and fluxes from Satellite data; *Graßl et al.*, 2000), GPCP (Global Precipitation Climatology Project; *Huffman et al.*, 1997) and CMAP (CPC Merged Analysis of Precipitation; *Xie and Arkin*, 1997), the precipitation values are derived from satellite measurements of infrared or microwave emissions which inherently include algorithms that are also some kind of model assumptions. During this study an error was found in the HOAPS precipitation data before 1992, so that these data should only be used for validation from 1992 onwards.

First the global water budget is discussed in Sect. 2.1. Then Sect. 2.2 shows a general comparison of the ERA40 precipitation patterns to ERA15 and observations, and in Sect. 2.3 the precipitation variability in time is considered. Hereafter, the following sections deal with special problems of the ERA40 precipitation such as the severe precipitation bias in the first realization of stream 2 (Sect. 2.4), the apparent changes in the global precipitation behaviour over the ocean (Sect. 2.5), and the spin-up during the forecast and the diurnal cycle (Sect. 2.6).

2.1. The global water budget

In long term means, the global water budget should be closed, i.e. the convergence (positive amount of precipitation minus evaporation (P-E)) of moisture over land should equal the divergence (negative P-E) of moisture over the ocean. Table 1 compares the 4 year (1989-92) global water budget over land and sea of ERA40 and ERA15 with climatological values (BR) of *Baumgartner and Reichel* (1975) and precipitation data of GPCC (Global Precipitation Climatology Centre; *Rudolf et al.*, 1996), GPCP, CMAP and NCEP reanalysis data (*Kalnay et al.*, 1996). The values for ERA40 yield that its global water budget is not closed, and, in addition, there is convergence of moisture over the ocean instead of divergence which is quite unrealistic. Although the budget in ERA15 is not closed either, the P-E values over the ocean for this period agree well with BR. Below it will be seen that the positive values of P-E over the ocean result mainly from too much precipitation over the tropical oceans from autumn 1991 onwards.

Data field at T106	ERA40	ERA15	GPCC	GPCP	CMAP	NCEP	BR
Precipitation over land	117	113	98	106	99	113	111
Evaporation over land	75	83	-	-	-	96	71
Precipitation over ocean	463	394	-	373	393	376	385
Evaporation over ocean	448	433	-	-	-	421	424
Total runoff	51	45	-	-	-	..	40
P-E over land	42	30	-	-	-	17	40
P-E over ocean	15	-39	-	-	-	-45	-39

Table 1. *Global water balance over land and ocean for the years 1989-92 in 10^{15} kg/a. BR designates climatological estimates according to Baumgartner and Reichel (1975).*

Over land P-E should be positive at each grid point. Figure 1 shows four year means of P-E. The negative annual mean values of P-E over land for ERA15, exceeding at places 1 mm/d, are significantly reduced in ERA40 but the bias is still not negligible in very few regions.

In ERA40, several major observing systems for water vapour are used, not only the radiosondes (where we know the humidity observations have deficiencies) but also HIRS and SSM/I radiances which also have their own biases and errors. Also SYNOP observations are used which may not be representative for the free atmosphere. The variational assimilation tries to adjust the model atmosphere to fit the observations, taking both the model's own error structures and the error structures of the observations into account combining the

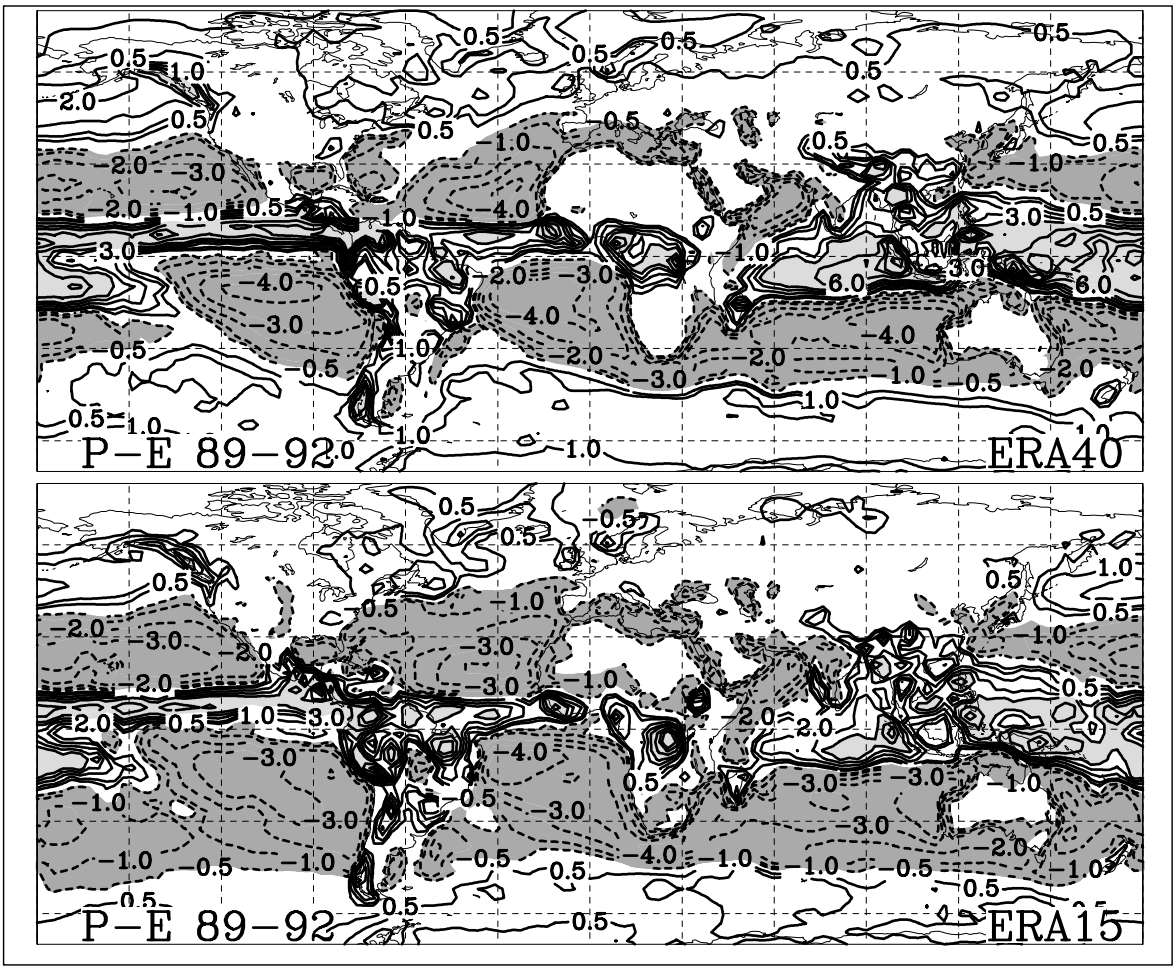

Fig. 1. *Mean distribution of precipitation minus evaporation in the years 1989-92 for ERA15 and ERA40. Contour lines are at +/- 0.5, 1, 2, 3, 4, 6, 8, 10 mm/d. The shading is done for > 4 mm/d and < -0.5 mm/d.*

different data. The result is an analysis which is not necessarily in balance with the model's own 'moist climate'.

The article by *Källberg* (same issue) explains how the ERA40 model is dumping net water out of the tropical atmosphere during the 6 h forecast period to the next analysis cycle and how the observations from satellite (presumably SSM/I or HIRS) are adding new water at each analysis cycle. This 'artificial' source of water causes too large precipitation over the tropical oceans where ERA40 produces more precipitation than ERA15, CMAP and GPCP data (see Table 1).

2.2. Validation of the precipitation patterns

In Fig. 2, the precipitation differences to GPCP analysis data in the winter (DJF) are shown for ERA15 and ERA40. In the extra tropics, the ERA40 precipitation is generally larger than in ERA15 which seems to be an improvement if compared to GPCP data, especially over the north western Atlantic and Europe where the negative biases in ERA15 are clearly reduced. Largest differences between ERA40 and ERA15 occur of course in the tropics where there are also the largest amounts of precipitation. ERA40 shows large increases of precipitation over the tropical oceans compared to ERA15, in particular over the tropical Indian ocean this increase is exceptionally strong. The comparison with GPCP suggests that this increase is an error. However, there were no geostationary satellite observations available over the Indian ocean which is the most important

input for the GPCP analysis over the other tropical oceanic areas and therefore the estimates of GPCP are less reliable for the Indian tropical ocean.

Another area of interest in analyses and model validation has been the SPCZ (South Pacific Convergence Zone). Many models tend to simulate a too zonal direction of the SPCZ, more like a double ITCZ. From the difference plots in Fig. 2 it is hard to judge if both reanalyses are different in this respect because ERA40 produces much stronger precipitation amounts than ERA15 or GPCP in the tropics and this difference in the amplitudes dominates the difference plots but plots of the precipitation amounts (not shown) suggest that ERA40 is slightly superior to ERA15 in the positioning of the SPCZ.

Large differences between the different precipitation estimates in Fig. 2 can be found over tropical South America. However, the uncertainties of the analyses in this area are very large and it is hard to judge which of both reanalyses is more realistic although it seems that the biases to GPCP data are slightly reduced in ERA40. The same applies to the tropical Africa. The high precipitation amounts along the Andes in both reanalyses are probably erroneous. .

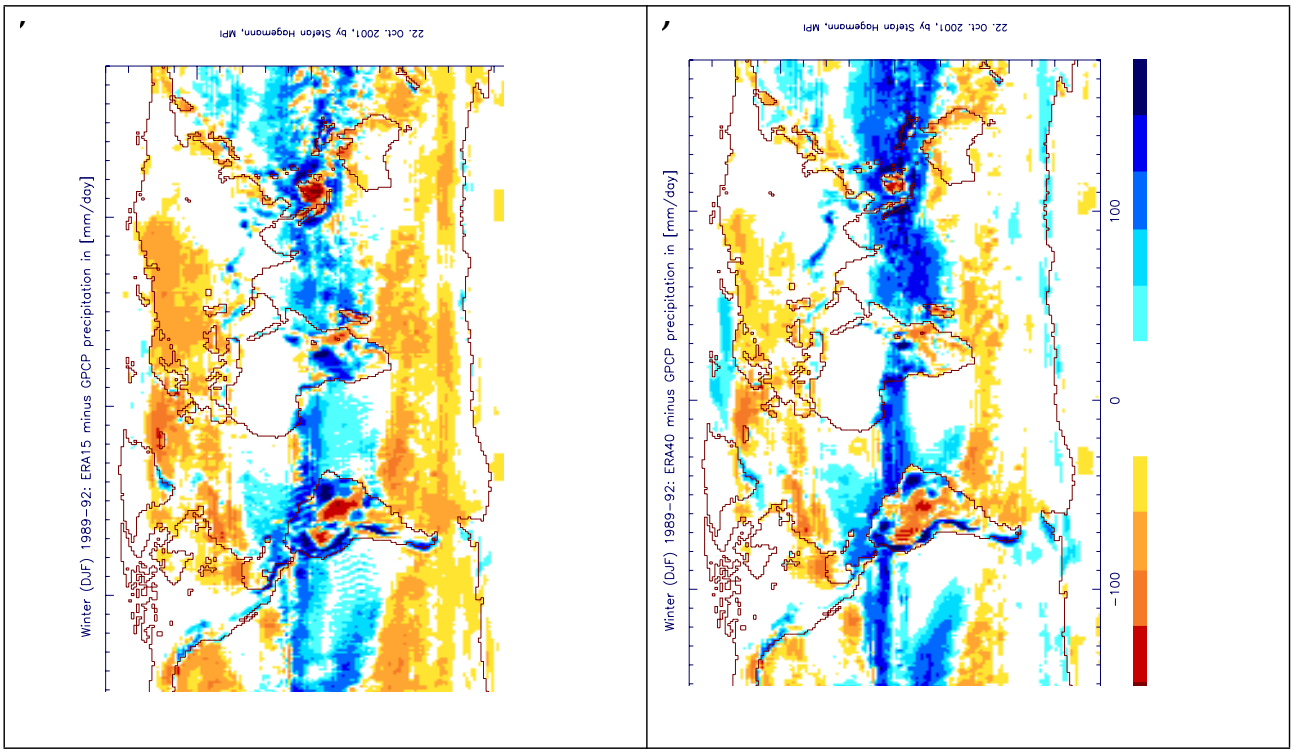

Fig. 2. *Precipitation difference in the winter (DJF) 1989-92 of* ***a)*** *ERA15 and GPCP data, and* ***b)*** *ERA40 and GPCP data at T106 degree resolution in mm/day*

During the summer (JJA) shown in Fig. 3 one finds similarities to the winter season, e.g. much more precipitation in the extra tropics, especially in the winter hemisphere which seem to be more realistic although the observations are not very reliable in this area. In addition, the wet bias over the Himalaya region is increased in ERA40. However, we do not know how much precipitation is really falling along the Himalayan mountains where several major Asian rivers have their main sources. The problems over the equatorial South America and Africa as well as over the tropical Indian ocean have already been mentioned above for the winter. A new feature is a shift of the ITCZ in ERA40 towards the north in the eastern Pacific. See, e.g., the dipole pattern in the difference plots between ERA40 and GPCP, which is not present in the ERA15 versus GPCP plot. In this respect ERA40 is probably worse than ERA15 as the GPCP data should be able to give a correct position of the ITCZ.

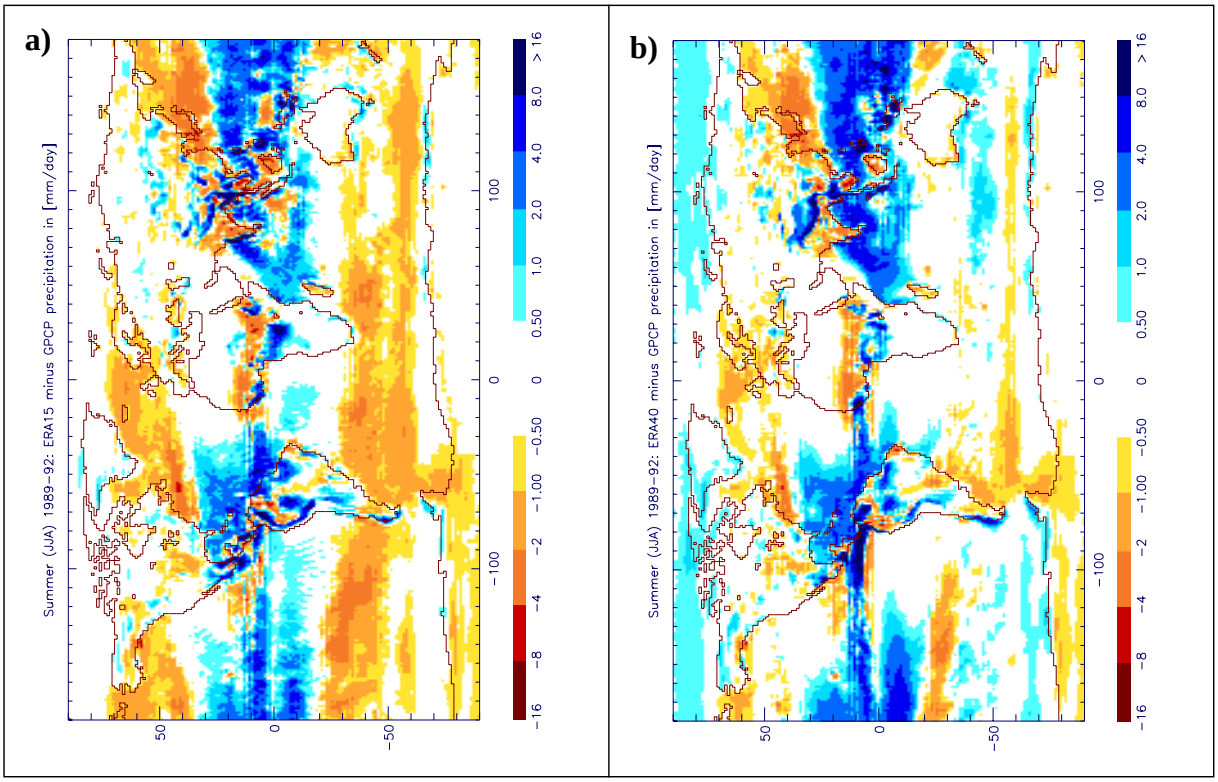

Fig. 3. *Precipitation difference in the summer (JJA) 1989-92 of a) ERA15 and GPCP data, and b) ERA40 and GPCP data at T106 degree resolution in mm/day.*

2.3. Precipitation variability in time

Generally the annual cycles of precipitation look very similar in both reanalyses and observational datasets when e.g. investigating maps of seasonal mean deviations from their annual means (not shown) but there are some areas where differences become obvious. Annual cycles of area mean precipitation are shown in Figure 4 for several areas.

Over India, ERA40 largely underestimates the precipitation in summer time. From Fig. 3 it can be recognized that the precipitation is generally lower in both reanalyses compared to GPCP all over India except in ERA15 around the western Ghats, which compensates for the lower values in other areas for the all India average in Figure 4.

For the two Atlantic areas ERA40 shows stronger annual cycles than the other datasets. For the equatorial western Pacific area, the whole annual cycle of ERA40 precipitation is shifted upwards compared to the other datasets. According to these few examples, which exhibit areas of larger differences in the annual cycle between ERA40 and GPCP/CMAP analysis, ERA40 seems to be inferior to ERA15.

In order to find areas where the variability of monthly mean values differ strongest between different datasets, maps of anomaly correlation coefficients between different datasets are shown in Figure 5. The variability of the monthly means has been reduced by subtracting the annual cycles, as this variability has already been discussed above. The polar regions have been excluded because the CMAP and GPCP data are not reliable in these areas, especially over the oceans. This can already be seen in the lowest panel showing the correlation between CMAP and GPCP. Over the oceans poleward of about 45˚ the variability of CMAP and GPCP hardly correlates (less than 40%). Low correlations can also be found in areas of low precipitation, e.g. off the coast

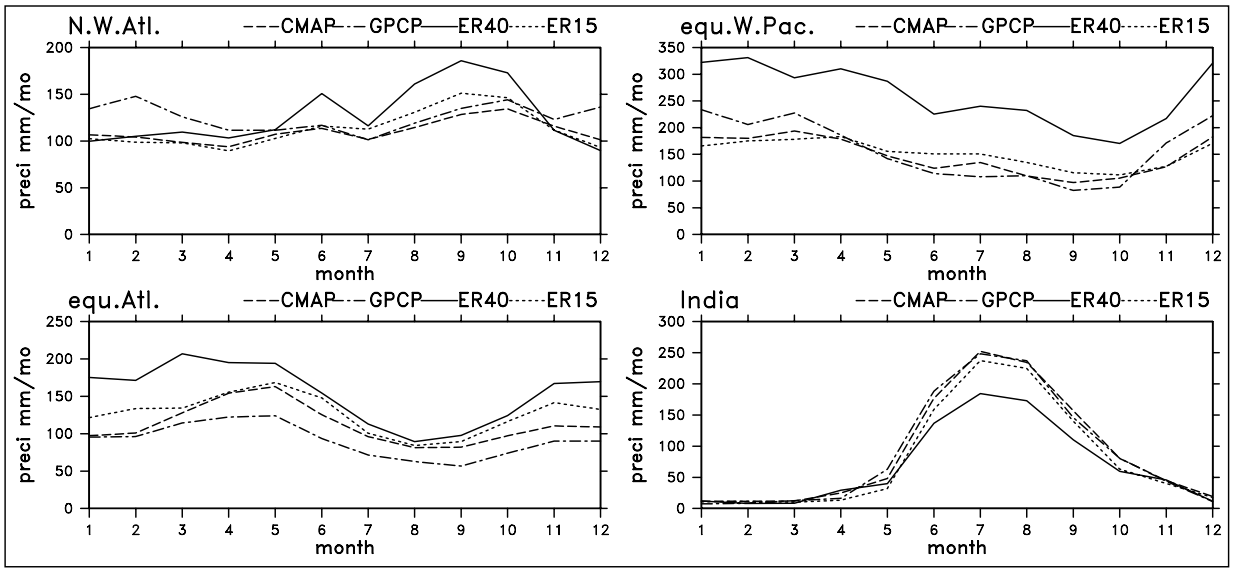

Fig. 4. *Annual cycle of precipitation for several areas where ERA40 strongly differs from precipitation data of CMAP, GPCP, and ERA15. The following areas are considered:*

N.W.Atl.	*north west Atlantic*	*60˚-80˚W, 20˚-40˚N, sea only*
equ.Atl.	*equatorial Atlantic*	*5˚-45˚W, 6˚S-6˚N, sea only*
equ.W.Pac.	*equatorial west Pacific*	*160˚-220˚E, 6˚S-2˚N, sea only*
India		*60˚-90˚E, 10˚-25˚N, land only*

of Peru, partly because we approaching there the division zero by zero and partly because low precipitation amounts are difficult to measure.

ERA15 correlates much better with CMAP than ERA40 while the difference between both reanalyses is much less dramatic when using GPCP as the corresponding truth but also there ERA40 seems to be inferior. This is demonstrated by area means of precipitation for northern Europe and eastern USA in Figure 6, two areas where observational data are reliable. Many of the month by month variation in the analysis for northern Europe are not followed by ERA40 but by ERA15, e.g. during the winter 1990/91.

For the eastern USA it is more that ERA40 underestimates many extreme values. The correlations and RMS differences between CMAP and the two reanalyses are:

	correlation		RMS	
	N. Europe	E. USA	N. Europe	E. USA
ERA40	0.83	0.88	5.7	12.5
ERA15	0.85	0.90	5.1	10.7

Here, it is obvious that ERA15 agrees better with CMAP than ERA40.

The low correlations in Figure 5 over the tropical oceans are mainly related to a northwards shift in the ERA40 precipitation over the ocean which is discussed in Sect. 2.5. The correlations between ERA15 and ERA40 are high (>85%) mainly over the continents where a good observational data basis exists while they are low in the tropics suggesting that the forcing from observational data in the tropics is low.

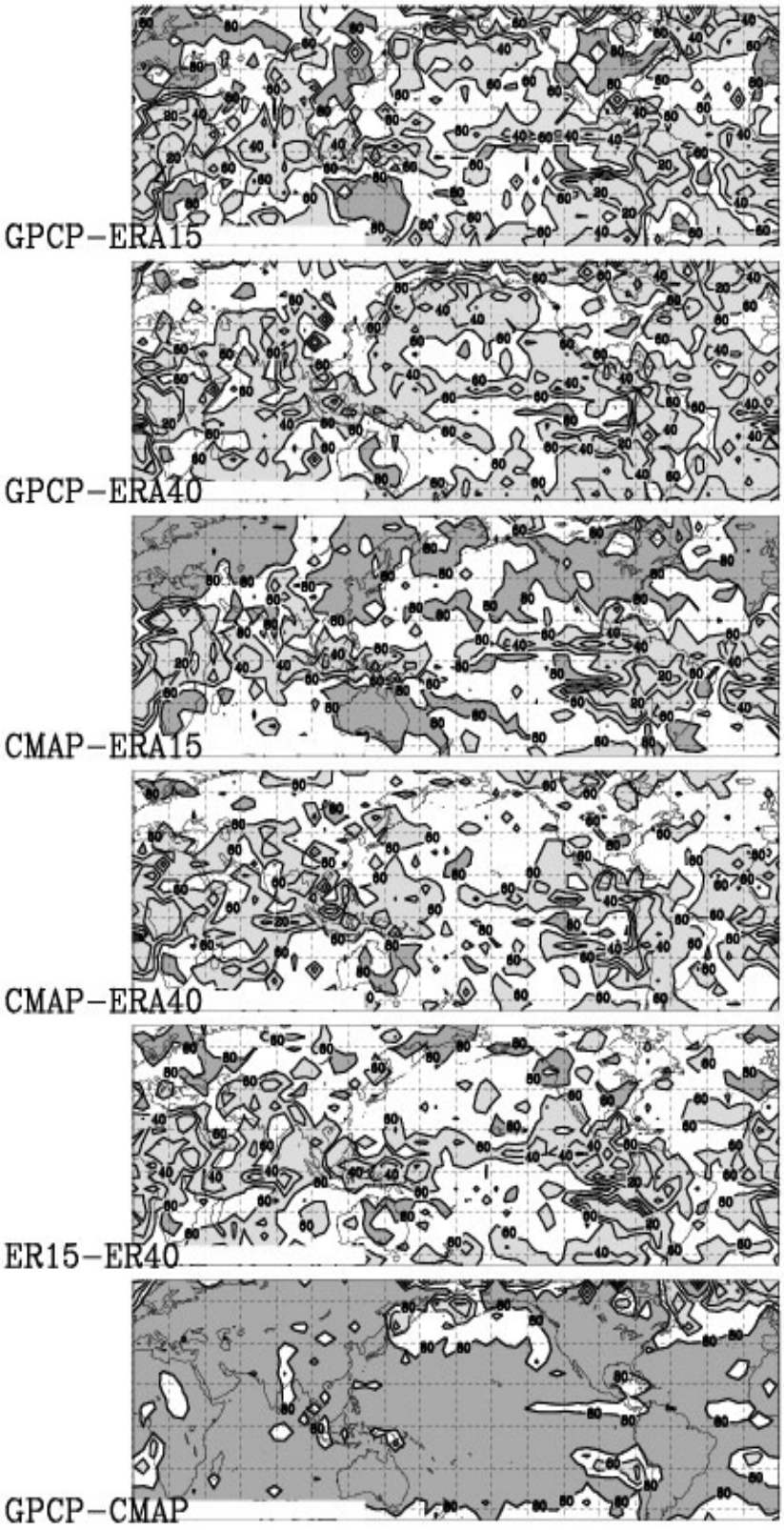

Fig. 5. *Correlations at each grid point between precipitation anomalies (deviation from the mean annual cycle) of different precipitation data. Contour lines are plotted at 20, 40, 60, 80, 100%. Shading for >80% and <60%.*

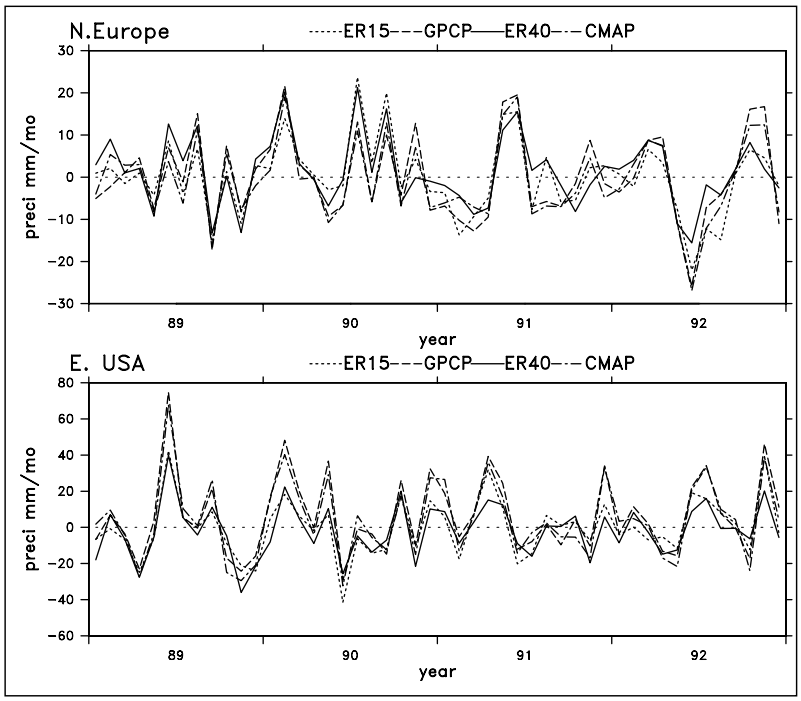

Fig. 6. *Time series of precipitation anomalies (deviation from annual mean cycle) of different datasets averaged for land grid points: Northern Europe (10˚-40˚E, 50˚-70˚N) and Eastern USA (80˚-100˚W, 30-40˚N).*

2.4. Precipitation of 1958 from stream 2

Figure 7a shows the ERA40 annual precipitation for 1958. Apparent features of the precipitation patterns are the comparatively low values over Brazil and the Congo catchment and the comparatively large values over India, the Sahel zone and the areas surrounding the Congo catchment. CRU precipitation data (*New et al.*, 2000) for the same year (Figure 7b) as well as the GPCC climatology for the years 1961-90 (Figure 7c) prove that the simulated ERA40 patterns are quite unrealistic in these regions. Especially over Africa, the observed patterns (large values in the Congo catchments, lower values in the surrounding areas including the Sahel zone) are reverted. Note that the GPCC climatology is also shown as the CRU precipitation data may be not that reliable (*Rudolf*, personal communication, 2001) but other observations for the year 1958 are currently not available at MPI.

Induced by this comparison, ECMWF found out that the erroneous precipitation could be traced back to corrupted temperature and dew points in the main SYNOP dataset for the period provided by NCEP and NCAR. NCEP and NCAR have identified the source of the corruption and produced a corrected dataset, so that the stream 2 could be restarted from the beginning at the time when the present report was written (*Simmons*, personal communication, 2001). Hereafter a data error was also found for stream 3 which had to be redone as well so that stream 3 is not considered in the present study.

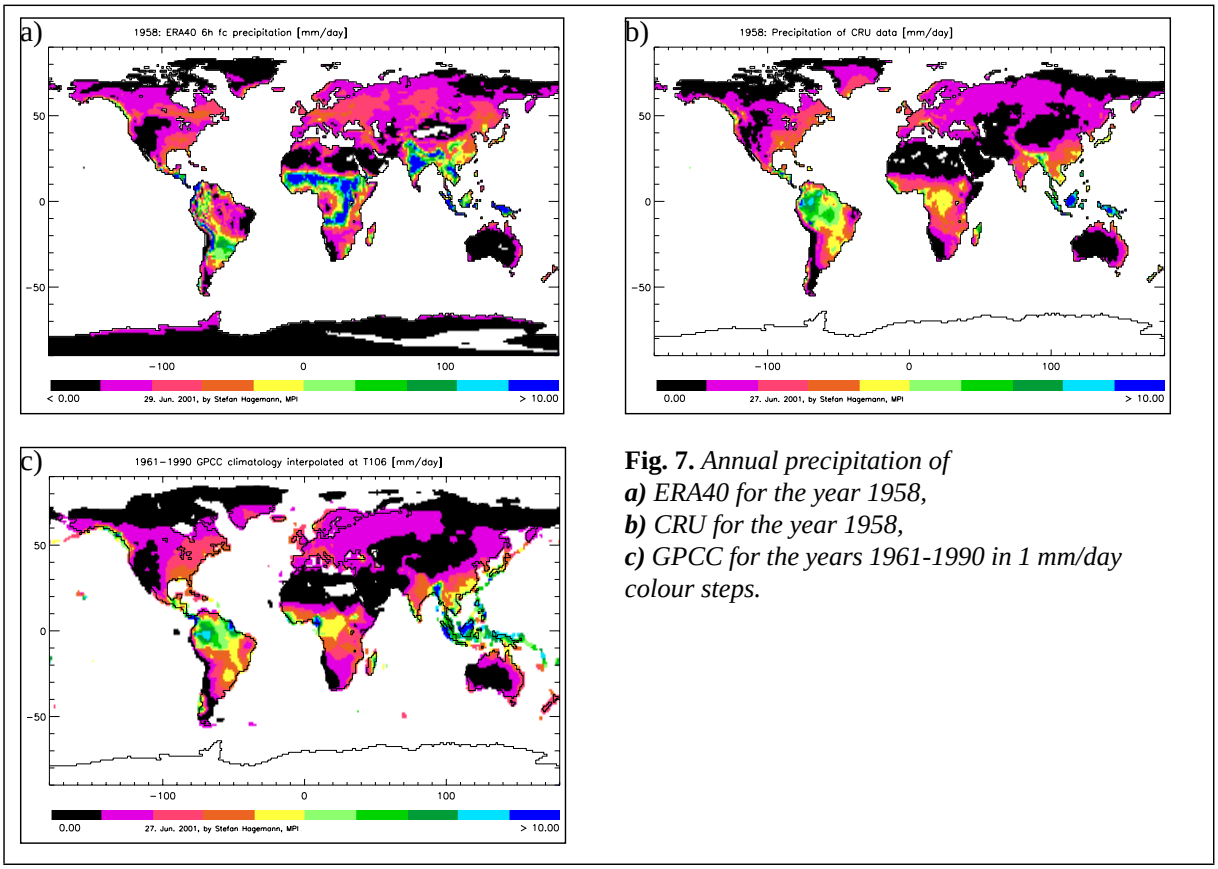

Fig. 7. *Annual precipitation of*
a) *ERA40 for the year 1958,*
b) *CRU for the year 1958,*
c) *GPCC for the years 1961-1990 in 1 mm/day colour steps.*

2.5. Major changes in the global precipitation time series over the ocean

It can be obtained from the monthly time series of global ocean precipitation (Figure 8, upper panel) that two major shifts characterize the ERA40 precipitation over the ocean. In September 1987, the first shift in the simulated precipitation behaviour occurs. This shift coincides with two changes that were made to the ERA40 system: the halving of the model time step and the inclusion of SSM/I data into the assimilation. It was not evaluated which of the two system changes has mainly influenced the shift in the precipitation.

In July 1991, a second shift can be seen in the ocean precipitation curve. None of the precipitation estimates from satellite (GPCP and CMAP) show such an increase nor does the reanalyses by NCEP. Consequently, we assume that ERA40 is erroneous in this respect. This shift is especially large over the tropical Indian ocean (not shown). The shift in the simulated precipitation behaviour coincides with the month of the Pinatubo eruption. This may lead to the conclusion that the aerosols, which were injected into the atmosphere by the volcanic eruption, have interfered with the satellite measurements thereby introducing biases in the data that were assimilated by the ERA40 data assimilation system. From our results it can be concluded that there is an error in the ERA40 system with shows up as excessive ocean precipitation, so we argue that this is related to an inconsistent use of data (SSM/I or HIRS) during the Pinatubo eruption. This conclusion is supported by the fact that no shift occurs in the ERA40 precipitation over land (Figure 8, lower panel) where no SSM/I data are used (see contribution by *Källberg* in same issue). The interference of the Pinatubo aerosols with SSM/I observations have at least lead to erroneous SST data in the ERA40 dataset as pointed out by *Reynolds* (personal communication, 2001).

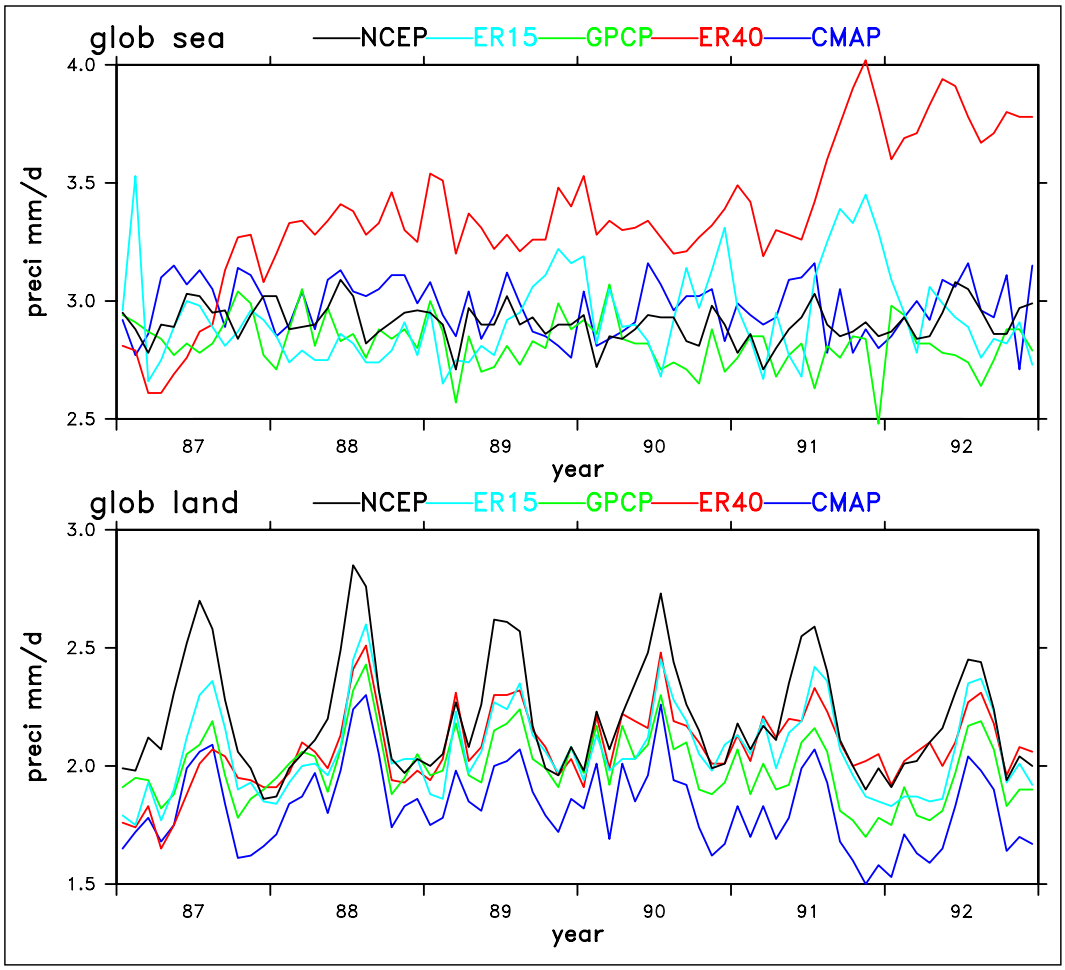

Fig. 8. *Global sum of precipitation over the ocean and land for the years 1987-92 in mm/day*

It has to be mentioned that the ERA15 precipitation over the ocean peaks also in the second half of 1991 which is not observed by GPCP and CMAP. Here, the peak may also be caused by the Pinatubo eruption but contrary to ERA40, ERA15 recovers from this event which might be related to the fact that no SSM/I or HIRS data were assimilated. On the other hand, similar, but lower, peaks can be seen in the two preceding years so that the reason for the peak may be totally different.

A similar jump in the precipitation had been found by *Stendel and Arpe* (1997) for the area around Lake Victoria for the year 1988 in ERA15 from which the analysis scheme did not recover for a few years. This jump was never understood but ERA40 recovered from a similar increase in this area very quickly which can be seen in Fig. 2 and Fig. 3 by strongly reduced precipitation amounts in ERA40 compared to ERA15 over southern Africa. A possible cause for such unrealistic jumps in time series of precipitation could be that erroneous observations did pass the different data checking procedures and were finally accepted by the analysis scheme which then lead to a sudden increase. In the following months the bias correction scheme made sure that the high values continued in the analysis even if the erroneous observations were no longer in the input data. *Trenberth et al.* (2001) suggest such an impact from bias corrections of satellite observations.

2.6. Spin-up during the forecast and the diurnal cycle

In order to show the spin-up of the hydrological cycle in ERA40, the global water balance of the year 1989 is considered in Table 2 for the 6h, 24h, 36h and the last 12 h of the 36 forecasts (24-36h). Note that with the

year 1989 a time period is chosen before the large shift in the precipitation regime of ERA40 has occurred (see Sect. 2.5).

Data field	6h	24h	36h	24-36h	GPCP	CMAP
Precipitation P over land	117	124	122	118	109	102
Evaporation E over land	75	75	75	75	-	-
Precipitation P over ocean	442	465	466	468	375	388
Evaporation E over ocean	452	459	462	468	-	-
Total runoff	52	52	53	55	-	-
P-E over land	42	49	47	43	-	-
P-E over ocean	-10	6	4	0	-	-

Table 2. *Global water balance over land and ocean for the year 1989 in 10^{15} kg/a. For ERA40, the 6h, 24h, 36h and the last 12 h of the 36 forecasts (24-36h) are considered.*

Table 2 shows that there is a large spin-up of the precipitation over the ocean during the first 24 hours of the forecast. Also the evaporation over the ocean and the precipitation over land increase during this time span, but to a lesser extent. In the 6h forecast, the water balance is not closed as P-E over land does not equal E-P over the ocean. This imbalance gets worse due to the spin-up of the hydrological cycle over the ocean.

For a closer investigation of the spin-up in the early forecasts, the diurnal cycle has to be considered at the same time. *Stendel and Arpe* (1997) have shown that the spin-up has a local dependency on the phase of the diurnal cycle, e.g. the spin-up over Australia is very pronounced when the forecasts starts at 12GMT, i.e. local evening. This dependency was found to be on local time, i.e. more spin-up in areas where the initial analysis refers to late afternoon to midnight. Therefore both aspects are investigated here together. Large signals where found in the following areas:

Winter (DJF)

Atl.ITCZ (50˚-25˚W, 2˚-5˚N, sea only)	NE.Atl. (10˚W-10˚E, 60˚-70˚N, sea only)
S.eq.Indonesia (100˚-120˚E, 2˚-9˚S)	E. USA (100˚-60˚W, 30˚-45˚N, land only)
Japan+ (125˚-155˚E, 25˚-45˚N)	SPCZ (120˚-150˚W, 10˚-30˚S, sea only)
S.W.Atl. (60˚-10˚W, 30˚-50˚S, sea only)	N.Brazil (50˚-75˚W, 15˚-0˚S)

Summer (JJA)

N.Atlantic (20˚-80˚W, 25˚-60˚N, sea only)	E.Pac.ITCZ (90˚-120˚W, 2˚-6˚N, sea only)
N.Brazil (50˚-80˚W, 6˚S-6˚N)	E. USA (100˚-60˚W, 30˚-45˚N, land only)
N.subtrop.Atl (45˚-80˚W, 25˚-40˚N, sea only)	S.of Japan (120˚-145˚E, 10˚-40˚N, sea only)
ctrl.Trop.Pac. (130˚-180˚W, 10˚S-5˚N)	India (60˚-80˚E, 10˚-25˚N, land only)

Figure 9 shows the precipitation averaged for the areas given above over two diurnal cycles in the 6 hour forecasts of ERA40 and ERA15 available four times per day and the values of the 36 hour forecasts starting at 00GMT and 12GMT. The value for the forecast range 00-06 GMT is plotted at 03GMT, for the range 06-12GMT at 09GMT, etc. Two diurnal cycles are shown so that the 36 hour forecasts can be fully displayed.

Common in areas with a spin-up and without a real diurnal cycle is a semi-diurnal cycle with lower precipitation in the 6h forecasts starting at 00GMT and 12GMT, i.e. analysis times when there are conventional observational data available. At the other analysis times there are not enough conventional observations to disturb the models own climatology. Therefore we assume that this semi-diurnal is an artifact.

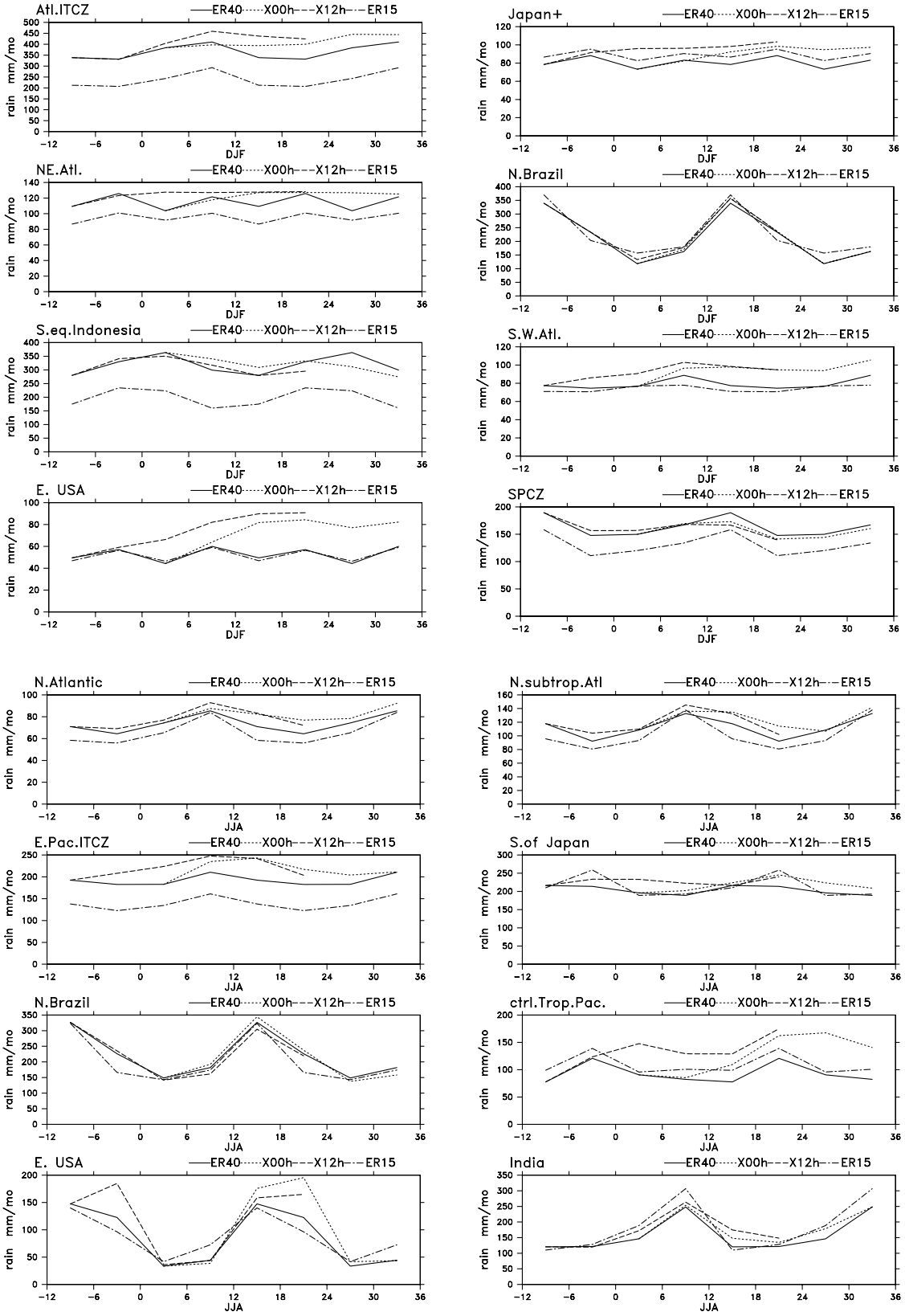

Fig. 9. *Figure 9 Diurnal cycle of precipitation averaged over a selection of areas during the winter (DJF) and the summer (JJA). The six hour forecasts of ERA40 and ERA15 available four times per day (ER40, ER15) and the values of the ERA40 36 hour forecast at 00GMT and 12GMT (X00h, X12h) are shown.*

Areas with such a feature are: NE.Atl., E. USA and Japan+. Areas where there is a spin-up only at one analysis time, e.g. S.eq.Indonesia and ctrl.Trop.Pac., show a diurnal cycle which is probably erroneous as well.

In areas with very strong diurnal cycles, such as over the Congo and Brazil, the spin-up is only a minor feature. *Stendel and Arpe* (1997) have shown that in ERA15 the diurnal cycle has almost everywhere over continents its maximum at local noon which is not realistic, as *Dai* (2001) reports about maximum convective precipitation over continents in the late afternoon. Table 9 shows that this has hardly changed with ERA40, only for some areas there is a hint that ERA40 has the maximum a little later than ERA15, e.g. for E.USA and N.Brazil in the summer (JJA). The precipitation maximum seems to occur before the maximum of 2m temperature, however, this is difficult to show as 2m temperature is archived at the synoptic hours, 00, 06, 12, 18 GMT, while the precipitation are means between these synoptic hours.

Dai (2001) reports maximum precipitation over northern oceans around 0600 local time and for tropical oceans from midnight to 0400 local time. In this respect the reanalyses agree better with observations, e.g. Atl.ITCZ, SPCZ, E.Pac ITCZ and S.W.Atl. Over the western and central tropical Pacific the maximum occurs at 12 - 18GMT (not shown) which means midnight local time.

Large differences in the spin-up over Australia, which where reported by *Stendel and Arpe* (1997) for ERA15, do not occur anymore. There is still a dependency of the spin-up on the phase of the diurnal cycle. In ERA15 there was more spin-up when the analysis starts in the dark part of the day and less during day-time. In ERA40 the spin-up difference is much smaller.

A prominent feature in the diurnal cycle is a marked difference between land and sea, which is especially strong for the 6h forecasts from 00Z and 12Z in the tropical Atlantic and neighbouring continents. ERA40 seems to produce a very large scale land-sea circulation which was not produced in earlier analyses. This land-sea contrast agrees with the observational study of *Dai* (2001) but we believe that it is too strong.

3. An ECHAM4.5 simulation nudged with ERA40 data

As the ERA40 precipitation has several deficiencies, we tried to calculate the precipitation fields by using the ECHAM4.5 model at T106 resolution nudged with ERA40 dynamical atmospheric data (vorticity, divergence, temperature, surface pressure). This nudged simulation was conducted from 1 January 1989 to 28 February 1991. In order to account for spin-up, especially with regard to soil moisture initialization, only results for 1990/91 are considered.

As we wanted to judge whether precipitation differences to observations are related to ECHAM4.5 model weaknesses or the nudged ERA40 data, a free ECHAM4.5 simulation was also conducted. Here, ERA40 SST was used as lower ocean boundary condition for the period 1 November 1989 to 28 February 1991, and fields from the nudging simulation were taken as initial fields at the start of the simulation.

Both the nudging simulation and the free run have a closed global water budget listed for the year 1990 in Table 3. The small differences between P-E over land and the runoff simulated by ECHAM4.5 are related to rounding and grib data conversion.

Data field at T6106	ERA40	nudged	free run	GPCC	GPCP	CMAP
Precipitation P over land	117	91	96	101	109	102
Evaporation E over land	76	67	69	-	-	-
Precipitation P over ocean	440	390	388	-	374	394
Evaporation E over ocean	454	416	418	-	-	-
Total runoff	51	23	26	-	-	-
P-E over land	41	24	27	-	-	-
P-E over ocean	-14	-26	-30	-	-	-

Table 3. *Global water balance over land and ocean for the year 1990 in 10^{15} kg/a. ERA40 data are compared to values of the ECHAM4.5 simulation nudged with ERA40 data and the ECHAM4.5 free run.*

In the following, precipitation differences to GPCP data were considered. In the summer 1990 (Figure 10) and in the winter 1990/91 (Table 11), the nudged ECHAM4.5 has a reduced wet bias over the tropical oceans but exhibits a largely increased dry bias over South America and Central Africa. In the summer 1991, the nudged simulation (Figure 10b) is generally drier than ERA40 (Figure 10a) so that the dry biases over northern high latitudinal land occur more widespread than in ERA40. Only over the high latitudinal oceans the nudged simulation is slightly wetter. In the winter 1991, the nudged simulation (Figure 11b) is also drier than ERA40 (Figure 11a) except for the high northern latitudes where it is wetter so that the European ERA40 dry bias is reduced and a wet bias occurs over North America in the nudged simulation.

In the summer 1990, the free run (Figure 10c) is closer to GPCP data than the nudged simulation (Figure 10b) over tropical land (not so dry), but worse over the high northern latitudes land (distribution). It is wetter over the southern subtropical oceans (seems to be worse) and drier over the northern high latitudes ocean. In the winter 1991, the latter is vice versa. The free run (Figure 11c) is wetter over the northern high latitudinal oceans and drier over the southern subtropical ocean. Over the tropical ocean, the maximum precipitation at the equator seems to be displaced in the free run, since the observations show a local minimum at the equator. Over tropical land, the precipitation is apparently wetter and more realistic in the free run, although the precipitation is still underestimated. Except for the Pacific warm pool, the free run seems to be closer to GPCP data than the nudged simulation in the tropical oceans, too.

As there is too much precipitation over the tropical oceans in ERA40, this is working as a heat source which enhances the Hadley circulation and the Walker circulation in the atmosphere. Thus, more subsidence over land will occur which may lead to drier conditions resulting in less precipitation. In ERA40, this effect is seen in the winter over Brazil and in the summer over the Sahel zone. The effect may be partially counteracted by the data assimilation where moisture is added to the atmosphere at each analysis cycle (cf. Sect. 2.1). In the nudged ECHAM4.5 simulation, where the enhanced circulation enters the simulation via the nudging, this causes an intensification of the model‘s dry bias over Brazil and central Africa.

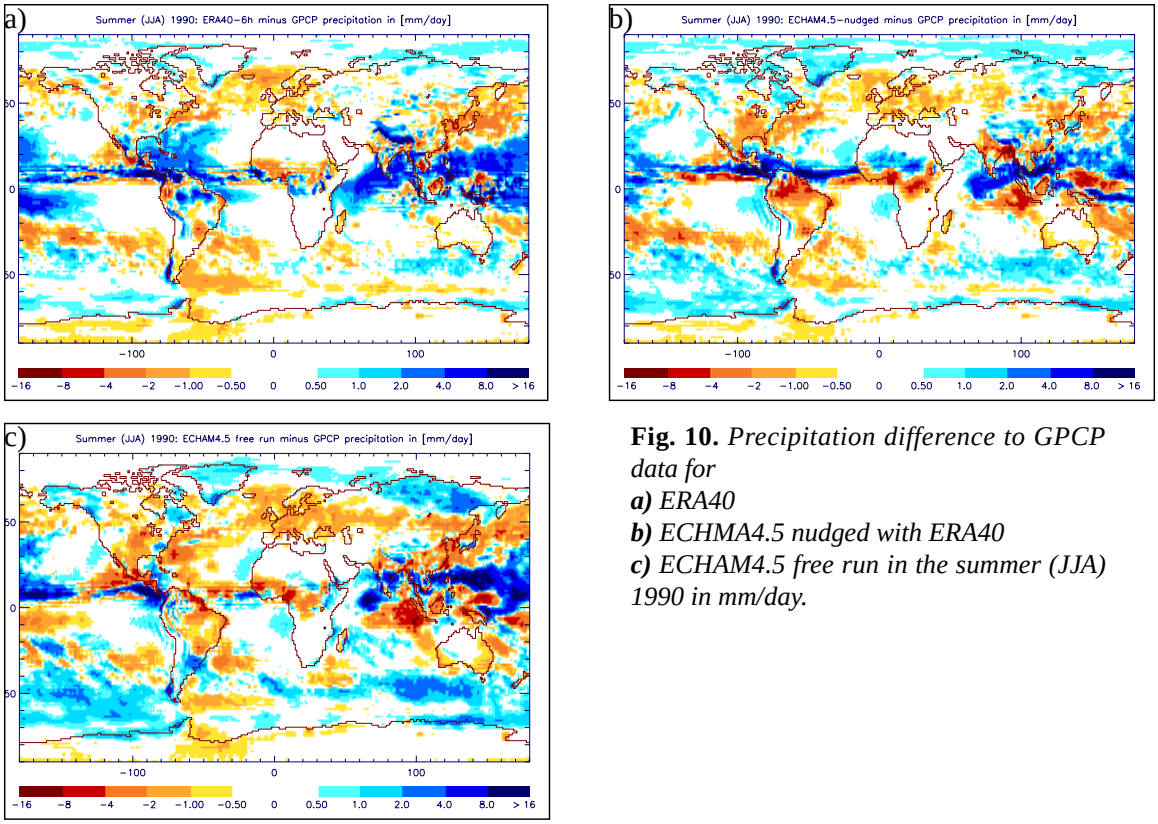

Fig. 10. *Precipitation difference to GPCP data for*
a) *ERA40*
b) *ECHMA4.5 nudged with ERA40*
c) *ECHAM4.5 free run in the summer (JJA) 1990 in mm/day.*

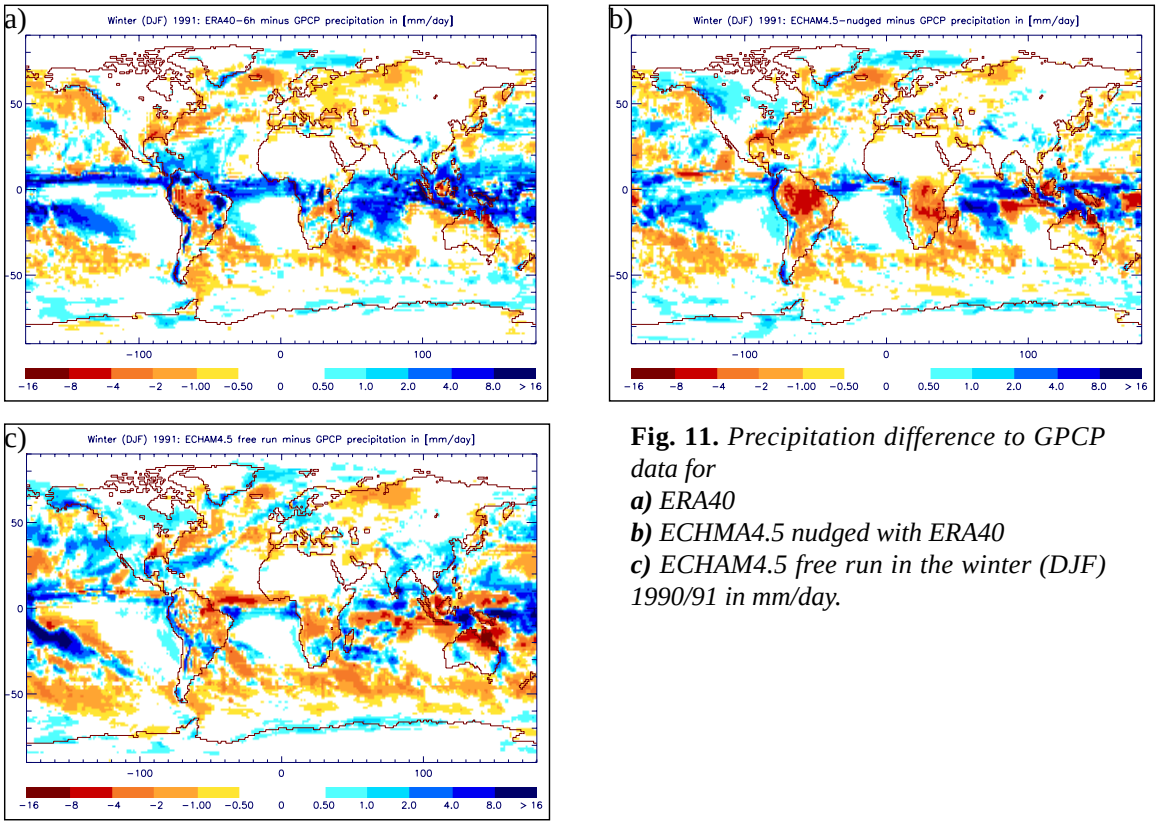

Fig. 11. *Precipitation difference to GPCP data for*
a) *ERA40*
b) *ECHMA4.5 nudged with ERA40*
c) *ECHAM4.5 free run in the winter (DJF) 1990/91 in mm/day.*

4. Summary and future plans

The results of the precipitation validation suggests errors in the ERA40 system. This is clearly the case for the unrealistic distribution of precipitation in 1958 in stream 2 which was caused by a corrupted SYNOP dataset that entered the data assimilation. In stream 1, the deficiencies of the precipitation are less obvious but they are partially be related to an error in the bias correction of the SSM/I data which ECMWF has found (*Kållberg*, personal communication, 2001). After the correction of this error it will be investigated how the hydrological cycle in ERA40 will change and presumably improve. Although the hydrological cycle of the current ERA40 data has improved in some aspects compared to ERA15, such as the increase of precipitation in the extra-tropics and the large reduction of grid points with negative P-E over land, there seems to be no real advantage of using the current ERA40 data (that includes the bias correction error) instead of ERA15 due to the large errors in precipitation.

In a dynamical adjusting or nudging process (*Jeuken et al.*, 1996) we forced ECHAM4.5 with the current ERA40 data. The nudged simulation has a closed global water balance but the precipitation difference to GPCP data is still comparatively large. Although the wet bias over the ocean is reduced, the adjusted precipitation with the ECHAM4.5 model also has deficiencies, in particular over South America and Africa. In the future it has to be investigated how the nudging simulation will behave when ERA40 data are used that do not contain the SSM/I data bias correction error. As the ECHAM4.5 free run is generally closer to GPCP data than the nudged simulation over the tropics, a special nudging experiment is planned using ECHAM4.5 with a free running model atmosphere in the tropics. This means a model setup with no nudging over the tropics between 10˚N and 10˚S and an intermediate zone of 10˚ to the north and south, respectively.

In ERA40, such as in ERA15, only little surface runoff occurs so that almost the whole total runoff comprises of drainage from the soil. This is unrealistic for most parts of the world so that these fields can not be used for a global discharge simulation. When a longer time series of ERA40 data will be available (at least 6 years), the hydrological discharge (HD) model (*Hagemann and Dümenil*, 1998; *Hagemann*, 1998) will be applied to ERA40 to compute discharges which will be compared to observations (*Dümenil Gates et al.*, 2000). The HD model requires daily time series of surface runoff and drainage from the soil as input fields. As mentioned above, appropriate fields are not directly available from ERA40. To generate consistent input values, time series of runoff and drainage will be calculated from the re-analysis data of precipitation and 2 m temperature using a simplified land surface (SL) scheme (*Hagemann and Dümenil Gates*, 2001). The SL scheme was recently improved by the intro-duction of a new surface runoff parameterization based on the ARNO scheme (*Dümenil and Todini*, 1992) and a high resolution land surface parameter dataset of *Hagemann et al.* (1999).

Compared to CRU (Climate Research Unit) temperature data (*New et al.*, 2000), the large northern hemisphere winter cold bias of ERA15 is replaced by a warm bias over northern Asia and North America. An indication of this warm bias can be already seen in ERA15 in the northern eurasian coastal regions. Therefore it is suggested that the bias is also included in the ERA15 data but it is overlaid by the severe cold bias. For the southern hemisphere, the winter cold bias of ERA15 is eliminated in ERA40. The elimination of the cold bias is mainly related to the inclusion of soil water freezing (*Viterbo et al.*, 1999) and a change in the snow albedo over forest-covered areas (*Viterbo and Betts*, 1999). It will be investigated how the warm bias may influence the hydrological cycle.

References

Baumgartner, A., and E. Reichel, 1975: The world water balance, *Elsevier*, Amsterdam

Dai, A., 2001: Global precipitation and thunderstorm frequencies. Part II: Diurnal variations. *J. Climate*, **14**, 1112-1128

Dümenil, L., and E. Todini, 1992: A rainfall-runoff scheme for use in the Hamburg climate model, in *Advances in Theoretical Hydrology — A Tribute to James Dooge*, edited by J.P. Kane, 129-157, Elsevier Sci., New York

Dümenil Gates, L., S. Hagemann, and C. Golz, 2000: Observed historical discharge data from major rivers for climate model validation, *Rep. 307*, Max-Planck-Inst. for Meteorol,, Hamburg, Germany

Gibson, J.K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura and E. Serrano, 1997: Era description, *ECMWF Reanal. Proj. Rep. Ser.* **1**, Eur. Cent. for Medium-Range Weather Forecasting, Geneva

Graßl, H., V. Jost, R. Kumar, J. Schulz, P. Bauer and P. Schlüssel, 2000: The Hamburg ocean-atmosphere parameters and fluxes from satellite data (HOAPS): a climatological atlas of satellite-derived air-sea-interaction parameters over the oceans, *Rep. 312*, Max-Planck-Inst, für Meteorol,, Hamburg, Germany

Hagemann, S., 1998: Entwicklung einer Parameterisierung des lateralen Abflusses für Landflächen auf der globalen Skala, *Examensarb. 52*, Max-Planck-Inst. for Meteorol., Hamburg, Germany

Hagemann, S., and L. Dümenil, 1998: A parameterization of the lateral water flow for the global scale, *Clim. Dyn.*, **14**, 17-31

Hagemann, S., and L. Dümenil Gates, 2001: Validation of the hydrological cycle of ECMWF and NCEP reanalyses using the MPI hydrological discharge model, *J. Geophys. Res.*, **106**, 1503-1510

Hagemann, S., M. Botzet, L. Dümenil and B. Machenhauer, 1999: Derivation of global GCM boundary conditions from 1 km land use satellite data, *Rep. 289*, Max-Planck-Inst, für Meteorol,, Hamburg, Germany

Huffman, G.J., R.F. Adler, A. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, R.J. Joyce, A. McNab, B. Rudolf, U. Schneider and P. Xie, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation data set, *Bull. Amer. Meteor. Soc.*, **78**, 5-20

Jeuken, A.B.M., P. C. Siegmund, L. Heijboer, J. Feichter and L. Bengtsson, 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., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, A. Leetmaa, R. Reynolds, and R. Jenne, 1996: The NCEP/NCAR Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471

New, M, M. Hulme and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate, *J. Climate*, **13**, 2217-2238

Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese and U. Schulzweida, 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., H. Hauschild, W. Rüth, and U. Schneider, 1996: Comparison of rain gauge analyses, satellite-based precipitation estimates and forecast model results, *Adv. Space. Res.*, **7**, 53-62

Stendel, M. and K. Arpe, 1997: Evaluation of the hydrological cycle in reanalyses and observations. *ECMWF Reanal. Proj. Rep. Ser.* **6**, 53p.

Trenberth, K.E., D.P. Stepaniak, J.W. Hurrell and M. Fiorino, 2001: Quality of reanalyses in the tropics. *J. Climate*, **14**, 1499-1510

Viterbo, P., and A.K. Betts, 1999: Impact on ECMWF forecasts of changes to the albedo of the boreal forests in the presence of snow, *J. Geophys. Res., 104*, 27,803-27,810

Viterbo, P., A.C.M. Beljaars, J.-F. Mahfouf, and J. Teixeira, 1999: The representation of soil moisture freezing and its impact on the stable boundary layer, *Q. J. R. Meteorol. Soc., 125*, 2401-2426

Xie, P., and P. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs, *Bull. Amer. Meteor. Soc*, **78**, 2539-2558