Meteorol. Atmos. Phys. 77, 19–43 (2001)

Meteorology and Atmospheric Physics

© Springer-Verlag 2001
Printed in Austria

[1] Max-Planck-Institute for Meteorology, Hamburg, Germany
[2] Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[3] Finnish Meteorological Institute, Helsinki, Finland
[4] Rossby Centre, Norrköping, Sweden
[5] German Weather Service, Offenbach, Germany
[6] Institute for Marine Research, University of Kiel, Germany
[7] University of Veterinary Medicine, Vienna, Austria
[8] Met. Office, Bracknell, United Kingdom
[9] Royal Netherlands Meteorological Institute, De Bilt, The Netherlands
[10] Danish Meteorological Institute, Copenhagen, Denmark
[11] GKSS Research Centre, Geesthacht, Germany
[12] Onsala University of Technology, Onsala, Sweden

A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period

D. Jacob[1], B. J. J. M. Van den Hurk[9], U. Andræ[2], G. Elgered[12], C. Fortelius[3], L. P. Graham[4], S. D. Jackson[8], U. Karstens[1], Chr. Köpken[5], R. Lindau[6], R. Podzun[1], B. Rockel[11], F. Rubel[7], B. H. Sass[10], R. N. B. Smith[8], and **X. Yang[10]**

With 13 Figures

Received September 8, 2000
Revised April 3, 2001

Summary

A comparison of 8 regional atmospheric model systems was carried out for a three-month late summer/early autumn period in 1995 over the Baltic Sea and its catchment area. All models were configured on a common grid using similar surface and lateral boundary conditions, and ran in either data assimilation mode (short term forecasts plus data assimilation), forecast mode (short term forecasts initialised daily with analyses) or climate mode (no re-initialisation of model interior during entire simulation period). Model results presented in this paper were generally post processed as daily averaged quantities, separate for land and sea areas when relevant. Post processed output was compared against available analyses or observations of cloud cover, precipitation, vertically integrated atmospheric specific humidity, runoff, surface radiation and near surface synoptic observations.

The definition of a common grid and lateral forcing resulted in a high degree of agreement among the participating model results for most cases. Models operated in climate mode generally displayed slightly larger deviations from the observations than the data assimilation or forecast mode integration, but in all cases synoptic events were well captured. Correspondence to near surface synoptic quantities was good. Significant disagreement between model results was shown in particular for cloud cover and the radiative properties, average precipitation and runoff. Problems with choosing appropriate initial soil moisture conditions from a common initial soil moisture field resulted in a wide range of evaporation and sensible heat flux values during the first few weeks of the simulations, but better agreement was shown at later times.

1. Introduction

The major objective of BALTEX (the BALtic Sea EXperiment) is to explore and quantify the energy and water cycles of the Baltic region (Bengtsson, 1995). One way to contribute to this objective is to reduce the uncertainties in our understanding by means of model exercises, aiming for the

quantification of the variety of processes which determine the spatial and temporal variability of the water and energy cycles. During the last few years, model simulations of atmospheric phenomena on time scales from days to decades have been carried out and validated against observations within the EU-funded projects NEWBALTIC I and II (Numerical Studies of the Energy and Water Cycle of the Baltic Region; Bengtsson, 1998 and 2000).

Christensen et al. (1997) describe a model inter-comparison with 7 limited area (regional) atmospheric models. Such an inter-comparison proved to be a very suitable exercise to assess the current state of the art of regional climate modelling (see also Gates et al., 1999). Their study covered entire Europe and results reflected model behaviour in a wide range of climatological conditions. They found considerable differences between land surface hydrological processes related to surface hydrological parameterisations. In most models summer precipitation was biased low due to a too rapid soil dry down, giving rise to a positive feedback of reduced evaporation and reduced precipitation over continental areas.

In this paper we describe a similar inter-comparison experiment focused on a smaller climatological range and hydrologically well defined area: the Baltic Sea catchment area. The inter-comparison exercise, carried out in the context of NEWBALTIC, was designed to span a three month period from August–October 1995 and to cover the $1.8 \times 10^6\,\mathrm{km}^2$ catchment area of the Baltic Sea. The aim of this comparison effort is to reveal differences in current state-of-the-art atmospheric models used for general Numerical Weather Prediction (NWP) and global or regional climate simulations. Where possible, observations are used as validation material. However, observational data do not cover all components of the hydrological and energetic cycles. For some relevant quantities the model inter-comparison will disclose uncertainties in the model generated values. This provides an estimate of the accuracy of these components in state-of-the-art atmospheric models. A total of 12 groups participated in this effort by either submitting one or more of the 8 participating model configurations or in the context of validation.

Initially, the comparison was organised by grouping results from the different models taken as they were used within the participating groups. The results of this effort appeared to be rather incomparable due to large differences in resolutions and boundary conditions. Therefore a unique effort has been made here: with two exceptions, all participating groups agreed on a common vertical and horizontal grid spacing, and identical atmospheric initial and boundary conditions. Furthermore, a validation strategy was designed and has been applied in the project. This includes agreement about variables to be validated, the observations and statistical methods used for the validation, and the layout of figures, tables etc.

Results of the comparison are grouped in quantities that could be validated by means of some sort of independent observations (vertically integrated specific humidity, near surface quantities, surface radiation, cloud cover, precipitation), and quantities where only a comparison of model output could be produced (surface turbulent fluxes, runoff). In the following section we will describe the set-up of the inter-comparison and the participating models in more detail. Section 3 presents the results for all quantities

2. Experiment design and model description

The simulations have been performed on an area covering well the Baltic Sea catchment. The simulation domain covered an area of approximately $7 \times 10^6\,\mathrm{km}^2$ except the model systems operated in so-called assimilation mode (DMI-CSE and DWD-BM; see below), all models used a horizontal resolution of 1/6° on a rotated latitude-longitude grid (Fig. 1) and a identical vertical layer structure of 24 layers up to 10 hPa. Model simulations were initialised on 1st of August 1995 and covered the three-months PIDCAP period (Pilot study of Intense Data Collection and Analyses of Precipitation) up to 31st of October 1995. During this period data coverage over the Baltic Sea catchment was enhanced. The period encompasses late summer with weather regimes dominated by convective systems, and some strong synoptic depressions passing the Baltex area roughly between 20 August and 10 September 1995.

Lateral boundary conditions were interpolated from HIRLAM reanalyses at 0.4° resolution (Sass and Yang, 1997). All models were initialised similarly at 1 August 1995, 00:00 UTC. Integra-

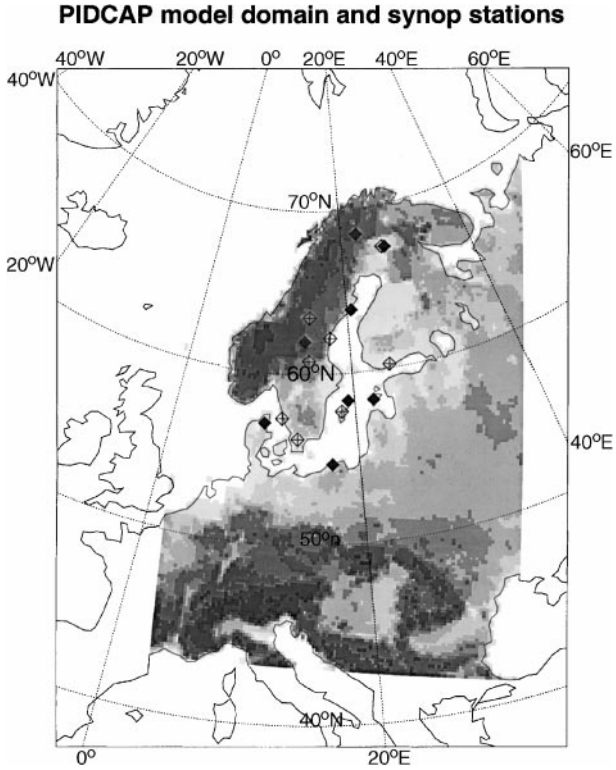

Fig. 1. Outline of the simulation domain. Shown is the area covered by the atmospheric models, and in grey shades the height of the surface geopotential. Also shown are the locations where Synop (solid symbols) and GPS observations (open symbols) were used to validate model results

tion lasted until 1 November 1995. The models were operated either in forecast mode (consecutive 30 hour forecasts including 6 hours spin-up time), in assimilation mode (intermittent data assimilation with 6 hourly forecasts and analyses) or in climate mode (initialised once and continuously forced at the lateral boundaries). Initial soil prognostic variables were derived from the DMI analyses with model specific scaling procedures to transform the HIRLAM variables into the model-specific quantities. Output from these runs was processed as synchronised daily averages between 6 am and 6 am the next day (see Table 1).

Table 1 lists the participating models and the mode they were operated in. In Table 2, a summary of the main components of the participating models is given. More details on the model configuration can be found in the NEWBALTIC I final report (Bengtsson, 1998).

The common HIRLAM analyses (DMI-CSE) were prepared by the Danish Meteorological Institute (DMI). The analyses with the limited area model HIRLAM (Källen, 1996) were computed in delayed-mode at 6-hours intervals using ECMWF analyses as lateral forcing. Such a delayed-mode data assimilation, which is essentially a reanalysis at regional scale, uses a more complete observational data set with longer cut-off time than normal operational use. As surface forcing, operational ECMWF-analyses of sea surface

Table 1. List of participating models, their acronyms and operation mode

Acronym	Institute	Model	Mode	Time aggregation	Hor. resol.	Vert. resol. (levels)	Lateral boundary condition
Forecast and analysis mode integrations							
DMI-CSE	DMI	HIRLAM	assimilation	6–12 hr	0.4°	24	ECMWF analysis
DWD-BM	DWD	BM	assimilation	0–6 hr	1/6°	30	DWD EM analysis
REMO-GKSS	GKSS	REMO+ GKSS cloud	forecast	6–30 hr	1/6°	24	DMI CSE reanalysis
Climate mode integrations							
RACMO	KNMI	HIRLAM+ ECHAM4 phys.	climate	Daily 6am–6am	1/6°	24	DMI CSE reanalysis
REMO-DWD	MPI	REMO+DWD phys.	climate	Daily 6am–6am	1/6°	24	DMI CSE reanalysis
REMO-EC4	MPI	REMO+ ECHAM4 phys.	climate	Daily 6am–6am	1/6°	24	DMI CSE reanalysis
SMHI	SMHI	HIRLAM	climate	Daily 6am–6am	1/6°	24	DMI CSE reanalysis
UKMO	UKMO	Unified Model	climate	Daily 6am–6am	1/6°	24	DMI CSE reanalysis

Table 2. Summary of physical parameterisations and key references of participating models

Acronym	Description	Key References
DMI-CSE	HIRLAM delayed-mode reanalyses; Sundqvist et al. (1989) cloud scheme with Kuo-convection; Non-local first order turbulence by Holtslag and Boville (1993); Radiation by Savijarvi (1990) and Sass et al. (1994); Early ECMWF land surface scheme with climatological deepest layer (Sommeria, 1985);	Källen (1996), Sass and Yang (1997)
DWD-BM	Baltex version of Europe Model. OI analysis following Loennberg and Shaw (1987); Vertical turbulence of Louis (1979) and Mellor and Yamada (1974); hourly radiation of Ritter and Geleyn (1992). Cloud microphysics; mass flux scheme of Tiedtke (1989); 3-layer soil model of Jacobsen and Heise (1982) with climatological deepest layer	Majewski (1991)
REMO-GKSS	REMO model, a modified version of Europe Model (DWD). DWD physics (see DWD-BM); cloud scheme after Doms (priv. comm.)	Majewski (1991)
RACMO	HIRLAM dynamics with ECHAM4 physics; surface scheme of Viterbo and Beljaars (1995); vertical diffusion using TKE-budget (Brinkop and Roeckner (1996), with exchange coefficients from Mahrt (1987) for stable stratification; shortwave and longwave radiation by Morcrette (1989); Convection using mass flux (Tiedtke, 1989), with deep convection adjustments due to Nordeng (1994); Condensation following Sundqvist (1978); gravity wave drag from Palmer et al. (1986)	Christensen et al. (1996)
REMO-DWD	REMO model, a modified version of Europe Model (DWD), with DWD physics (see DWD-BM)	Majewski (1991), Jacob and Podzun (1997)
REMO-EC4	REMO with ECHAM4 physics: land surface from Dümenil and Todini (1992); vertical diffusion using TKE-budget (Brinkop and Roeckner (1996); shortwave and longwave radiation by Morcrette (1989); Convection using mass flux (Tiedtke, 1989), with deep convection adjustments due to Nordeng (1994); Condensation following Sundqvist (1978); gravity wave drag from Palmer et al. (1986)	Majewski (1991) Roeckner et al. (1996) Jacob (2001)
SMHI	HIRLAM as DMI-CSE, but with local first-order turbulence according to Louis (1979).	Källen (1996)
UKMO	Unified Model; surface subgrid hydrology (Dolman and Gregory, 1992) and MOSES land surface parameterisation (Cox et al., 1999); drag from unresolved orography by Milton and Wilson (1996); first-order turbulent mixing (Smith, 1990); Mass-flux convection scheme by Gregory and Rowntree (1991); large scale cloud and precipitation by Smith (1990); radiation following Slingo and Wilderspin (1986) and Slingo (1989)	Cullen (1993)

temperature (SST) and sea ice cover data were combined with weekly manual analyses of SST and sea ice cover over the Norwegian Sea and North Sea (prepared by the Norwegian Meteorological Institute), and with SMHI-analyses over the Baltic Sea which are produced about once every 4 days. These SST and sea ice cover fields are used in all the model runs presented in this paper, with the exception of the runs made with the Baltic Model operated by the German Weather Service (DWD-BM, see below).

As analysis scheme, the Optimum Interpolation technique is used, which features a multivariate three dimensional interpolation of observed deviations from first guess fields. Only conventional observation data are used in the analysis. The 6-hourly data assimilation cycling for the PIDCAP period is performed on a 146×146 horizontal grid with 0.4° spacing, and 24 vertical levels. At each cycle, a 12 hour forecast is made, but only results after 6 hour integration are used, to account for typical spin-up time that modelled moisture quantities need. In constructing daily precipitation time series, accumulated precipitation from +6–12 hour from each cycle is extracted.

Also the Baltic Model (DWD-BM), run at the German Weather Service (DWD), was run in assimilation mode. The DWD-BM is a version

of the hydrostatic limited area Europa-Modell (EM) operational at DWD until December 1999 (Majewski, 1991). For the purpose of creating an improved analysis for the inter-comparison period, an updated model orography based on a 1 km data set (GTOPO30) has been used and the vertical resolution was increased to 30 levels to achieve a better representation of the boundary layer (10 layers below 1500 m). Lateral boundaries are provided at hourly intervals from the operational Europe Model (EM). The analysis is a three-dimensional multivariate optimal interpolation (Loennberg and Shaw, 1987) carried out in an intermittent 6-hour cycling. Conventional observations (from the synoptic surface network, radiosondes and aeroplanes) as well as satellite data (satellite derived wind vectors and geopotential thickness, when available) are used. Snow depth is analysed every cycle. The sea-surface temperature is updated once daily. The SST analysis from the National Centre for Environmental Prediction (NECP, former NMC, USA) serves as a background field in an additional successive corrections analysis using data of the 1 degrees SST analyses of the SMHI as additional observations in the Baltic Sea. The values used in the inter-comparison are taken from the forecasts of the assimilation cycle started from each analysis after a nonlinear normal mode initialisation. Daily values are thus composed of four 6 hour forecasts and certain variables (like precipitation) may therefore contain spin-up effects.

The REMO-GKSS simulations were carried out by the GKSS Research Centre using a combination of the REMO limited area model (Jacob and Podzun, 1997) and the DWD physical parameterisation schemes (Majewski, 1991). The cloud scheme in the DWD package was replaced by a cloud parameterisation where cloud liquid water and ice are both prognostic variables (G. Doms, DWD, private communication). For the model inter-comparison consecutive 30 h forecasts were performed with hourly output of the main variables. Simulations were restarted daily from analyses.

RACMO (Christensen et al., 1996) is set up as a combination of the HIRLAM dynamical structure and global climate model ECHAM4 physical parameterisation schemes (Roeckner et al., 1996). In the version operated by the Royal Netherlands Meteorological Service (KNMI) only the land surface scheme is different from the default ECHAM4 package: the ECMWF scheme by Viterbo and Beljaars (1995) is implemented. This is a 4-layer prognostic soil scheme, with a free drainage and zero heat flux bottom boundary condition. Initial soil water content is derived from the DMI-CSE analyses by keeping the relative soil saturation similar. Owing to differences in the formulation of the dependence of runoff and evaporation on soil moisture content, this does not necessarily produce similar evaporation rates as in the DMI-CSE model. A more common soil initialisation procedure would be to apply a soil moisture scaling preserving the relative soil moisture difference to the model- and season specific climatology (Robock et al., 1998), but this procedure was not applied here. Exchange coefficients for vertical turbulence under stable conditions is parameterised following Mahrt (1987), which generally enhances vertical exchange in the lowest levels compared to the original formulation by Holtslag and De Bruin (1988).

The model denoted by the acronym REMO-DWD run by the Max-Planck Institute in Hamburg, is the REMO limited area model (Jacob and Podzun, 1997), combined with the physical package of the DWD operational forecast model. This model version is identical to REMO-GKSS apart from parts of the cloud parameterisation scheme as described above.

REMO-EC4 denotes the same dynamical host model REMO, but with the ECHAM4 physical package (Jacob, 2001). As to the physical parameterisations it is very much comparable to RACMO, apart from the use of the default ECHAM4 land surface scheme by Dümenil and Todini (1992). Also, vertical turbulence is not modified from the ECHAM4 scheme described by Roeckner et al. (1996).

The model labelled SMHI operated by the Swedish Meteorological and Hydrological Institute (SMHI) is similar to the HIRLAM model (Källen, 1996) used to create the DMI-CSE forcing analyses. The only difference is in the turbulence parameterisation, which is a local scheme according to Louis (1979).

The model labelled UKMO denotes a limited area configuration of the United Kingdom Met Office Unified Model (Cullen, 1993). The version used for this inter-comparison study was that operational in 1997 except that a newly developed

parameterisation of land-surface processes (Cox et al., 1999) was included. The surface hydrology scheme is described by Dolman and Gregory (1992); the treatment of stratiform clouds and boundary layer turbulent mixing by Smith (1990); the convection scheme by Gregory and Rowntree (1991) and the orographic surface roughness by Milton and Wilson (1996). (Note that the gravity wave drag scheme was not switched on, as is operational practice for high resolution configurations of the Unified Model.)

The main ingredients of the physical parameterisation used in the different models can be found in Table 2.

3. Results of the model integration

The model results are compared with respect to a number of quantities: precipitation, vertically integrated specific humidity, total cloud cover, surface radiation, surface sensible heat flux, evaporation, total runoff, and near surface weather parameters. In some cases model output and validation material are presented for selected observation locations. In other cases, area integrated quantities are presented.

3.1 Precipitation

More than 4000 daily rain gauge observations from 14 countries have been collected by the BALTEX Meteorological Data Centre (BMDC; Lehmann et al., 1999). This data have been analysed using the Precipitation Correction and Analysis (PCA) model developed by the Working Group Biometeorology at the University of Veterinary Medicine Vienna. The PCA model consists of a dynamic correction module for the reduction of systematic measurement errors (Rubel and Hantel, 1999) and a geostatistical module to estimate spatial averaged precipitation values representative for the grid used in this model inter-comparison study. The daily precipitation fields for PIDCAP have been published in the PIDCAP Ground Truth Precipitation Atlas (Rubel, 1998). The precipitation analysis was carried out for all grid points within the Baltic Sea catchment area. Average precipitation values were calculated for the land and sea grid points separately.

Time series of model output, averaged over the Baltic Sea region and over the BALTEX land area

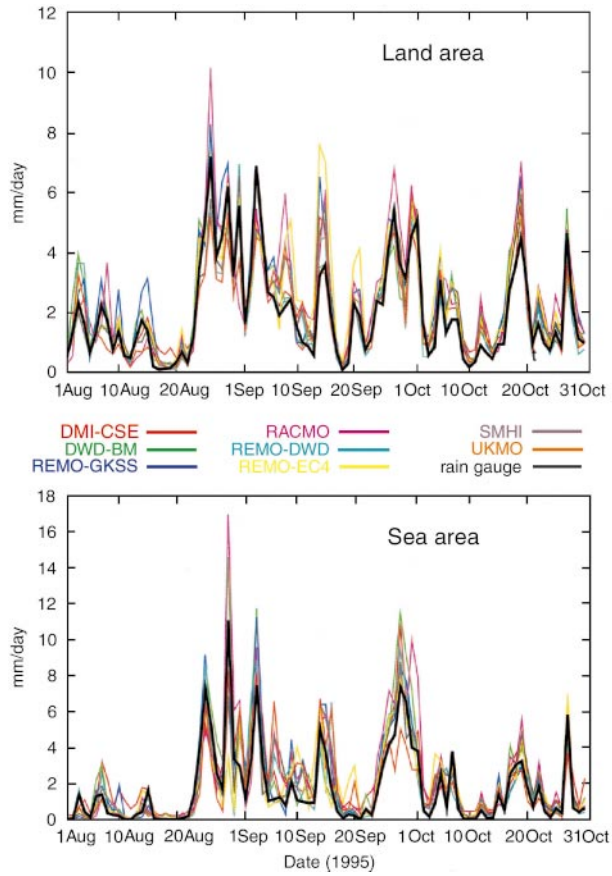

Fig. 2. Total precipitation (mm/day) over land surfaces (top panel) and sea surfaces (bottom panel) within the Baltic Sea drainage basin during PIDCAP

separately, were compared to the analysed gauge data. Figure 2 presents the temporal evolution of the daily precipitation for the PIDCAP period, averaged over the entire BALTEX area excluding the Baltic sea. The time series plotted are from models as well as from the rain gauge analysis. In Table 3 the statistics of simulated time series of daily precipitation in comparison to the corresponding rain-gauge data analyses are presented. The time series in Fig. 2 reveal several PIDCAP episodes with intensive rainfalls over the BALTEX area, and representation of these events by models as well as the rain-gauge data analysis seem to agree generally well. The averaged model bias relative to the rain-gauge data, as shown in Table 3, indicates that simulated precipitation matches closely with the averaged measurement value. Overall the two independent short range forecasts extracted from the reanalysis cycles in DMI-CSE and DWD-BM, seem to agree best with the rain-gauge analysis in spite of being at quite

Table 3. Statistics of comparison of modelled and analysed daily precipitation (mm/day) over the Baltic sea catchment land area, averaged over the entire PIDCAP simulation period

	Bias (model vs analysis)		RMS (model vs analysis)		Correlation (model vs analysis)
	mm	%	mm	%	
DMI-CSE	0.14	7	0.58	30	0.94
DWD-BM	0.23	11	0.59	29	0.95
REMO-GKSS	0.46	23	0.82	41	0.93
RACMO	0.48	24	1.05	52	0.88
REMO-DWD	−0.05	−3	0.66	33	0.92
REMO-EC4	0.45	22	1.01	50	0.86
SMHI	0.16	8	0.69	34	0.91
UKMO	−0.07	−3	0.75	37	0.89

different horizontal and vertical resolutions. The time series from climate mode runs also show generally good correspondence with the observation analysis. In particular, the time series from climate mode runs with HIRLAM at SMHI and the REMO-DWD are close to the analysis data. In fact, the statistical properties in terms of bias, RMS and correlation relative to rain gauge data from several climate mode runs are quite similar to those from short range forecasts with frequent restart. Besides this, higher total precipitation amounts are seen for simulations by REMO-GKSS (23%) and the climate mode runs in RACMO and REMO-EC4 (+24% and +22%). The correspondence between the latter two was likely due to the equivalent parameterisation schemes used by these models.

Statistics for the precipitation time series in individual PIDCAP months have also been examined. These statistics (not shown here) indicate that despite different integration modes and configuration, the simulated time series by various models for August and October months agree well among each other, and they all have a good correspondence to the analysed time series, with correlation levels ranging between 0.90 and 0.98. Worse agreement is seen for September. For that month, the climate mode runs tend to show relatively low correlation levels compared to analysis and runs with assimilation modes. In particular, RACMO and REMO-EC4 show excessive amounts of precipitation, which are significantly larger than the rain-gauge data and the other models. The results from runs with assimilation and forecast modes still agree relatively well with the analysis based on rain-gauge data. A likely cause of the large differences in precipitation simulation from models and measurement data could be that the precipitation pattern for September month reflects more dominance of convective events, thereby governed more significantly by local and regional scale effects. The difference among model systems, especially those in physical parameterisations, may thus cause significantly different results in simulation of rainfall events. As for all the numerical models, an extended time integration suffers from risk of inaccurate representation of current synoptic situations within a model domain. In that respect, a short range forecast at assimilation or forecast modes, with a frequent restart from analyses, has clearly an advantage over an extended range climate run, which is only restricted by periodic analyses at the lateral boundaries.

In Fig. 2 (lower panel) and Table 4, comparisons for precipitation time series averaged over the Baltic Sea are presented. The figure again indicates many episodes of intensive precipitation during PIDCAP over the Baltic Sea. Similar to that for land surface, the overall correlation between analysis and model data from assimilation and forecast runs is relatively high, although the models give a higher precipitation estimate. One should, however, take into account that the precipitation analysis relies on a rather limited amount of observations over sea and is itself less reliable. But it can be noted that climate mode runs tend to feature lower agreement with the analysis than the assimilation or forecast runs. Table 4 shows that the scatter between modelled precipitation time series and analysis is considerably larger than over land, measured in RMS values. An examination of

Table 4. As for Table 3 for the sea area

	Bias (model vs analysis)		RMS (model vs analysis)		Correlation (model vs analysis)
	mm	%	mm	%	
DMI-CSE	0.49	29	0.95	56	0.92
DWD-BM	0.56	33	1.20	71	0.95
REMO-GKSS	0.35	21	0.98	58	0.92
RACMO	0.61	36	1.48	87	0.89
REMO-DWD	0.11	6	1.37	81	0.80
REMO-EC4	0.13	8	1.32	78	0.79
SMHI	0.25	15	1.10	65	0.87
UKMO	0.08	5	1.26	74	0.81

the cross-correlation between modelled time series over sea also indicates a rather low agreement level among the climate mode runs, indicating stronger model dependence over the sea area, where the physical parameterisations of individual models play a dominant role.

3.2 Vertically integrated specific humidity

For the validation of model results with integrated specific humidity (IWV) the observed IWV derived from 8 GPS-stations has been used. The GPS data have been processed as part of the Swedish and Finnish network at the Onsala Space Observatory, Sweden (Emardson et al., 1998), and are available with a time resolution of 10 minutes, thus providing quasi-continuous data coverage. For compatibility with the 6-hourly model output interval, the GPS data are averaged to the same time interval. Model output is calculated for the nearest grid point, regardless whether that point is treated as land or as sea for that location by the models. Additionally, at 3 of the 8 GPS-stations IWV-values are available derived from nearby radiosondes (at a distance of less than 40 km).

Generally there is a good agreement between observations and GPS data (Fig. 3). The high variability of the observed IWV is simulated very well by all models. The strong variations in IWV are mostly linked to the variability of the large scale flow, and are well captured by all models.

For the models which are integrated in forecast and assimilation modes the correlation often exceeds 0.90. For all models integrated in climate mode the correlation are slightly lower. This is mainly caused by one particular weather situation occurring in the southern part of Scandinavia during 8–20 September, in which the models in climate mode do not reproduce the development of a cyclone as well as the models integrated in forecast mode. Small differences in the modelled large scale flow lead to strong deviations of

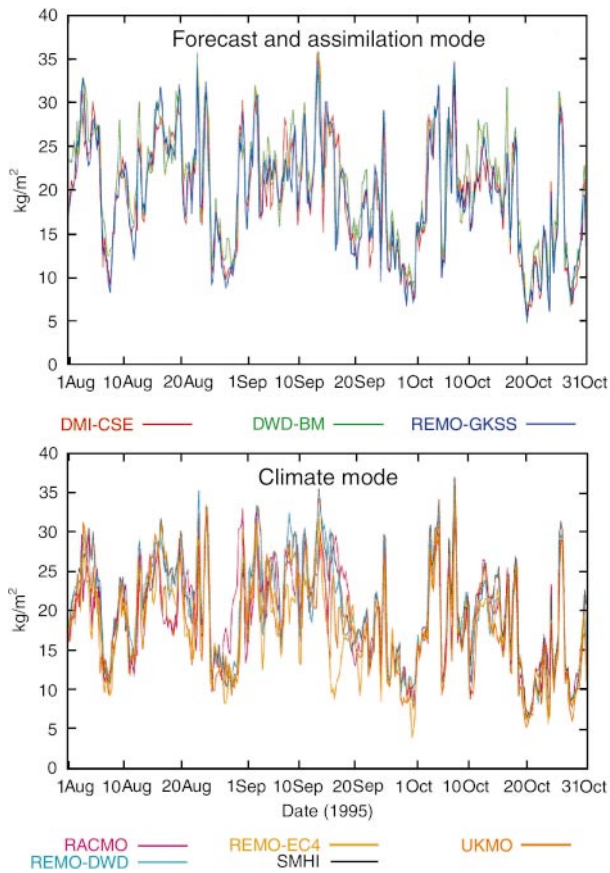

Fig. 3. Time series of the mean 6 h vertically integrated specific humidity (kg/m^2) at the station Haessleholm (Sweden) from GPS observations and models in assimilation and forecast mode (upper panel) and observations and models in climate model (lower panel)

Table 5. 3-month average vertically integrated specific humidity (IWV; kg/m^2) for 8 locations within the simulation domain, as measured using GPS (bold faced) and simulated by the models shown. Displayed bias and correlation coefficients are with reference to the GPS data. RACMO correlation coefficients were not computed. Also shown are estimates from radiosonde profiles at three locations. For this comparison, GPS-data were averaged slightly differently to match the temporal resolution of the radiosonde data

Model/Observation	mean	bias	correlation	mean	bias	correlation	mean	bias	correlation
	Haessleholm			Leksand			Metsaehovi		
GPS	**18.7**	**–**	**–**	**14.6**	**–**	**–**	**15.7**	**–**	**–**
DWD-BM	20.5	1.8	0.95	18.9	4.4	0.91	19.5	3.8	0.91
DMI-CSE	19.2	0.5	0.93	15.3	0.7	0.94	17.7	2.0	0.90
REMO-GKSS	19.2	0.5	0.93	16.0	1.4	0.94	17.7	2.0	0.88
UKMO	17.9	−0.8	0.85	16.3	1.7	0.92	17.0	1.4	0.82
RACMO	19.4	0.7		16.4	1.8		18.0	2.3	
SMHI	20.0	1.3	0.89	16.7	2.1	0.92	18.3	2.6	0.85
REMO-DWD	19.7	1.0	0.90	16.5	1.9	0.92	17.8	2.1	0.82
REMO-EC4	17.9	−0.8	0.84	15.2	0.7	0.83	16.8	1.1	0.81
	Oestersund			Visby					
GPS	**13.6**	**–**	**–**	**17.2**	**–**	**–**			
DWD-BM	16.6	3.0	0.95	20.9	3.7	0.93			
DMI-CSE	13.6	0.0	0.94	18.4	1.2	0.90			
REMO-GKSS	14.8	1.1	0.94	18.5	1.4	0.90			
UKMO	14.2	0.6	0.91	18.1	1.0	0.79			
RACMO	14.3	0.7		18.6	1.4				
SMHI	15.7	2.1	0.92	19.7	2.5	0.80			
REMO-DWD	15.6	2.0	0.91	19.0	1.8	0.82			
REMO-EC4	14.3	0.6	0.91	17.4	0.3	0.70			
	Onsala			Sodankylae			Sundsvall		
GPS	**16.9**	**–**	**–**	**12.8**	**–**	**–**	**15.9**	**–**	**–**
DWD-BM	21.5	4.6	0.96	14.5	1.7	0.93	18.9	2.9	0.94
DMI-CSE	19.4	2.5	0.95	13.7	0.9	0.93	15.5	−0.5	0.95
REMO-GKSS	19.6	2.7	0.92	14.0	1.2	0.92	16.3	0.3	0.92
UKMO	19.1	2.2	0.84	13.1	0.3	0.90	16.2	0.3	0.91
RACMO	19.3	2.4		14.6	1.8		16.7	0.8	
SMHI	20.5	3.6	0.85	15.0	2.2	0.89	17.6	1.7	0.93
REMO-DWD	20.1	3.2	0.85	15.0	2.2	0.88	17.1	1.1	0.94
REMO-EC4	18.4	1.5	0.82	13.9	1.1	0.87	16.0	0.0	0.88
GPS (Radiosonde)	**16.7**			**13.9**			**15.1**		
Radiosonde (302)	18.5	1.8	0.93	14.4	0.5	0.91	16.5	1.4	0.94

the IWV compared to GPS data (Lenderink and Van Meijgaard, 2001). Yang et al. (1999) find satisfactory correspondence between the model (HIRLAM) simulated IWV and GPS observations during PIDCAP. This indicates an encouraging prospect for the future use of GPS observations in the atmospheric data assimilation of NWP.

The bias of the simulated data compared to GPS varies between −0.8 and 4.6 kg/m^2 (Table 5). It should be noted, though, that the current GPS data set is underestimating the amount of atmospheric specific humidity by approx. 1 kg/m^2 compared to water vapour radiometer data for the PIDCAP period and earlier (Emardson et al., 1998). Taking this into account, most models are within ±1 kg/m^2. Higher positive biases occur for RACMO, but also for the models which are based on DWD-physical parameterisations, in particular DWD-BM. This is probably due to the fact that DWD-BM is the only model using EM analyses as lateral forcing, which have the same positive water vapour bias (not shown). The reason for this positive bias is not fully understood, but it may well be related to the way surface cloud cover

observations are used in the analysis, in order to enhance low-level cloudiness.

In Table 5, a comparison between GPS and radiosondes is also shown. GPS data shown here are averaged corresponding to the temporal resolution of the radiosonde data, which differs from the 6-hour model output interval. All models which are based on DWD-physical parameterisation schemes (DWD-BM, REMO-DWD, REMO-GKSS in this order) overestimate the IWV, which can be seen clearly in Table 5 and Fig. 3.

3.3 Total cloud cover

The total cloud cover from the models is compared to a cloud analysis based on satellite imagery and conventional observations. This analysis merges the visible and infrared data from the 12 UTC METEOSAT image with direct observations from radiosondes (humidity) and the surface synoptic network (cloud cover and cloud height; Rosenow et al., 1992). It should be kept in mind, however, that the cloud analysis itself may also not represent the true cloudiness. All observations, be it by satellite from space or observations from the ground have their deficiencies in different situations. The analysis, combining VIS and IR information with direct observations from the surface network, partly alleviates this problem by exploiting the strengths of the different systems. But still, the analysed cloud cover values should not be regarded as absolute truth but as an estimate.

The total cloud cover output from the models is generally produced as instantaneous cloud cover at 12 UTC. An exception to this is the output from REMO-EC4, which consists of hourly averages from 11 to 12 UTC. Cloud cover output was not available for the UKMO and DMI-CSE models. The aim of the comparison is to verify the overall cloud amount produced by the different models and parameterisation schemes. Therefore, statistics of the mean total cloud cover for the land and sea areas of the Baltic sea catchment area are compiled.

Figure 4 displays time series of the 12 UTC mean cloud cover over land for analysed cloud cover (labelled as ‘‘analysis’’) and for the different models. The observed cloud cover varies mostly between about 45–80% in August 1995 and 50–90% in September and October 1995. Large

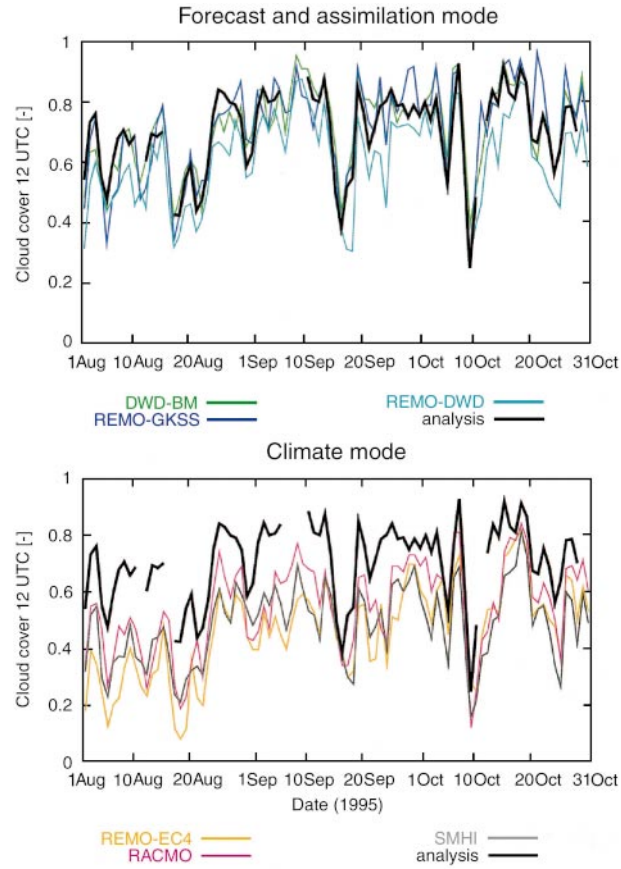

Fig. 4. Timeseries of cloud cover at 12 UTC each day during PIDCAP. On each panel, the black line ‘‘analysis’’ represents the analysed cloud cover. On the upper panel the modelled cloud cover from REMO-GKSS, DWD-BM and REMO-DWD, on the lower panel from REMO-EC4, RACMO and SMHI are displayed additionally

variations caused by synoptic systems are visible, strong events occurring especially in the late summer and early autumn where occasional peaks may be as low as 25% or as high as 93% for the whole land area. All models show the variations linked to synoptic events. They are reproduced most closely by DWD-BM and REMO-GKSS (being short term forecasts), but also the models running in climate mode reproduce clearly the major and even many smaller events. This indicates that the boundaries, where analysed fields are specified, exercise a strong control over the complete model domain. The overall value of cloudiness, however, shows marked differences between the models. The results from the DWD-BM as well as from REMO-GKSS and also, to a somewhat lesser extent, REMO-DWD are in very good agreement with the observations, while

Table 6. Monthly mean values of total cloud cover at 12:00 UTC of the Baltic region over land. Results from UKMO and DMI-CSE were not available

Model	August	September	October
REMO-GKSS	0.61	0.76	0.77
DWD-BM	0.61	0.77	0.75
REMO-DWD	0.52	0.66	0.64
REMO-EC4	0.32	0.49	0.57
RACMO	0.44	0.58	0.61
SMHI	0.41	0.53	0.51
Analysis	0.63	0.74	0.72

Table 7. As Table 6 but for the sea area

Model	August	September	October
REMO-GKSS	0.51	0.74	0.73
DWD-BM	0.48	0.71	0.70
REMO-DWD	0.36	0.63	0.61
REMO-EC4	0.23	0.48	0.50
RACMO	0.38	0.58	0.59
SMHI	0.21	0.39	0.38
Analysis	0.50	0.69	0.68

REMO-EC4, RACMO and SMHI have significantly lower mean total cloud cover. This is especially the case during August and September, slightly less in October 1995.

The mean total cloud cover values for the three months are summarised in Table 6 and Table 7 for land and sea, respectively. Figure 5 shows the frequency distribution of occurrence of different total cloud cover levels (averages for the whole domain, i.e. land and sea). The analysed values for land vary between 63% on average in August and 74% in September.

DWD-BM and REMO-GKSS have very similar average cloud cover, although the latter produces slightly higher values due to many more situations with very high cloud cover (see Fig. 5). Both models have identical dynamics and similar physics, apart from REMO-GKSS using separate prognostic equations for temperature, water vapour, cloud water and ice, whereas DWD-BM uses only two prognostic equations for total heat and total water (assuming equilibrium at saturation). An independent evaluation (Zhang et al., 2000) revealed an overestimation of high clouds in the REMO-GKSS scheme. This is supported by the current comparison, where the cloud analysis suggests a slightly lower level of cloudiness during September and October than produced by REMO-GKSS (about 2–5%).

REMO-DWD, using the same physical package as DWD-BM but different dynamics, shows roughly 10% less cloudiness than DWD-BM on average and lies about 6–8% below the cloud analysis. The models REMO-EC4 and RACMO use essentially the same physics. The difference in mean cloud cover, about 5–10% higher values in RACMO, may be caused either by different dynamics or slight changes of parameters in the cloud scheme. A sensitivity study was carried out with REMO-EC4 addressing the value of the relative humidity threshold which defines the onset of condensation. Air which is unsaturated on average within a $1/6^\circ \times 1/6^\circ$ grid box can lead to partial condensation due to subgrid variability. Using a profile for this threshold (80% in the upper third of the model domain) rather than requiring saturation at all heights showed an increase of 10% in total cloud cover in REMO-EC4 (results not shown). Note that the underestimated cloud cover in RACMO and REMO-EC4 provides a compensation of the surface radiation underestimation (next section). SMHI yields the lowest cloud cover and has virtually no situations with cloud cover exceeding 68% on average, contrary to most other models and the analysis.

The mean total cloud cover simulated by the different models varies by as much as 20–35% (different ranges are found depending on month and land or sea area). Overall, the models tend to produce lower cloud cover than is analysed. The deviations model minus analysis range from -20% to $+5\%$ for the monthly averages. It is interesting to note that two model pairs using essentially the same physics with different dynamical formulations produce about 10% differences in average total cloud cover (REMO-EC4 versus RACMO; REMO-DWD versus DWD-BM).

3.4 Surface radiation

Modelled daily mean values of downwelling short and long-wave radiation at the surface were compared to measurements from the four stations Schleswig (Germany), Lund, Borlänge, and Luleå (Sweden), obtained from the BALTEX meteorological data centre (Bengtsson, 1995). Figure 6

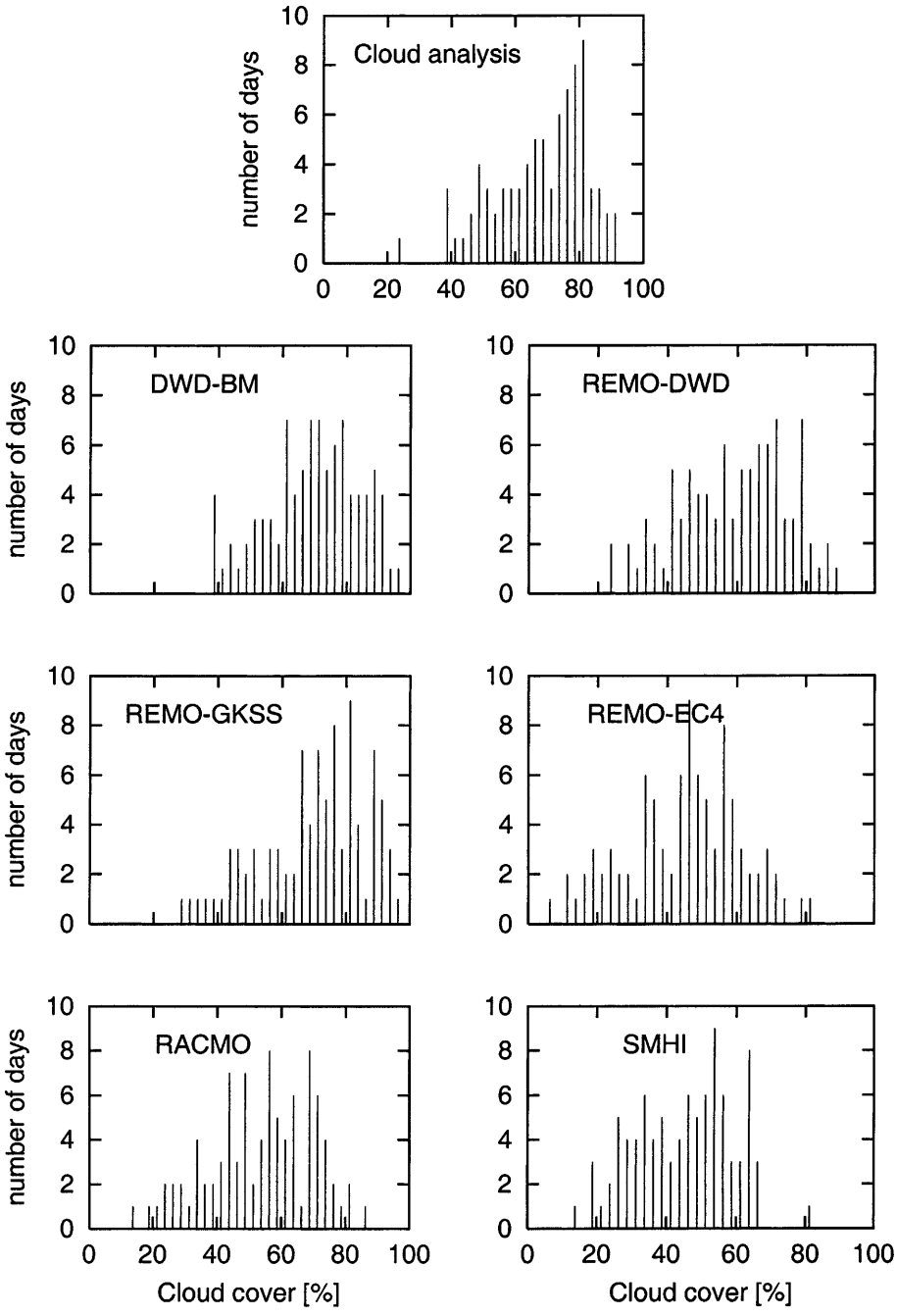

Fig. 5. Histograms of the number of days on which a certain average cloud cover occurred (binned in 2.5% steps) for the cloud analysis (top) and the different models (panels below)

shows typical time series from the station Lund in southern Sweden.

Short-wave fluxes were available from the models DWD-BM, REMO-GKSS, REMO-EC4, REMO-DWD, SMHI and RACMO. All these models reproduce the observed seasonal trend, and a fair proportion of the synoptic variability, although episodic discrepancies on the order of 100 W/m^2 are not uncommon (Fig. 6, upper panel). Such discrepancies are much too large to be explained by differences in the parameterised clear sky fluxes. They must rather reflect differences in the cloud radiative forcing, arising from discrepancies in the predicted cloud fields themselves as well as in the radiative effects of the clouds.

Table 8 describes statistics of the short-wave flux at all the stations. Systematic errors range in absolute values from less than 5 W/m^2 for REMO-GKSS to about 25 W/m^2 for RACMO. For a given

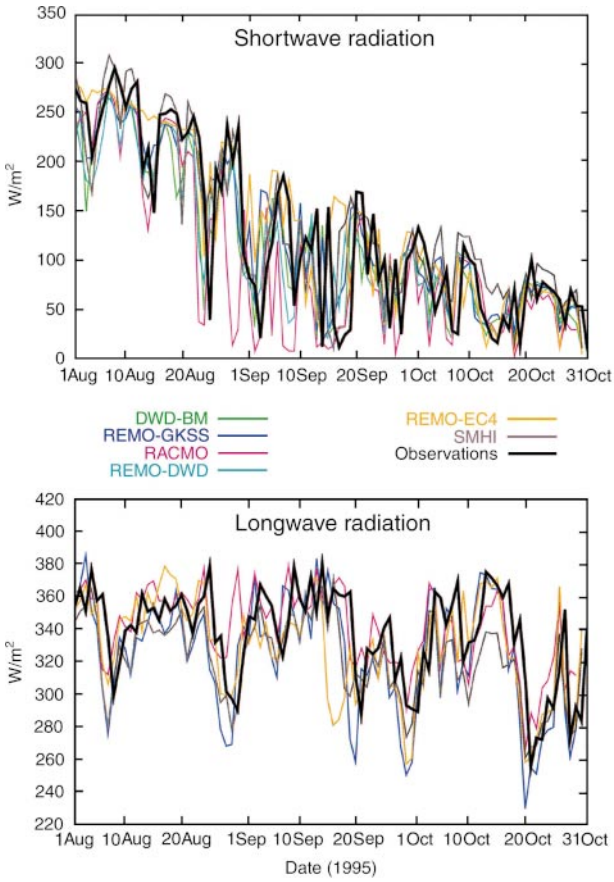

Fig. 6. Downwelling short-wave (upper panel) and long-wave (lower panel) surface radiation fluxes (W/m) at Lund in southern Sweden

model the bias tends to retain its sign from one station to the next. The variability is generally well described by the models. Correlation coefficients in the range of 0.82–0.97 indicate that the models explain roughly 70–95% of the observed variance. However, the models DWD-BM, REMO-GKSS, and REMO-DWD underestimate slightly the amplitude of the variability. The random errors must be dominated by the predicted cloud systems. Consistently with this statement, the models running shorter forecasts (DWD-BM and REMO-GKSS) have the smallest random errors as well as the highest correlation with observations. Overall, RACMO seems to provide the least reliable simulation of the short wave flux.

Figure 7 shows histograms of observed and modelled daily mean short-wave fluxes for the ensemble of all stations. The measurements display an almost linear decrease of frequency with increasing intensity. The negatively biased models, DWD-BM, REMO-DWD and RACMO, all capture the monotony of the distribution, but overestimate the slope in a pattern that could result from overestimated cloud forcing. It is worth noting that the pronounced pattern of RACMO arises in spite of an under-predicted daily mean cloud cover. Van Meijgaard et al. (2001) examined the shortwave cloud radiative properties of RACMO using observations collected in The Netherlands, and conclude that the ECHAM4 physical package as implemented in RACMO is too opaque for high cloud cover. Interestingly, the same physical package in REMO-EC4 shows a better correspondence to observed surface radiation, but a lower cloud cover (Fig. 4) compensates the excessive absorption and/or reflection of short-wave radiation in this model.

The positively biased models, REMO-EC4 and SMHI, both under-predict the frequency of low-flux cases as might be expected from the under-predicted mean cloudiness in these models. However, REMO-EC4 captures the linear shape of the observed distribution, while SMHI displays a maximum frequency between 60 and 120 W/m^2.

Long-wave fluxes were available only for the models REMO-EC4, REMO-DWD, SMHI and RACMO. These models show a fair agreement with the measurements, although episodic discrepancies exceeding 20 W/m^2 are frequently seen (Fig. 6, lower panel).

Results for the long-wave flux are summarised in Table 9. The three models REMO-GKSS, REMO-EC4 and SMHI are all negatively biased. By contrast, RACMO has a small positive bias everywhere except at Luleå where the bias is large. The ratio of the random error to the variability is larger than for the short-wave flux, and the correlation coefficients are lower, corresponding to about 60–85% explained variance. However, this may be not so much the effect of a worse description of the day to day variability as of the lack of a seasonal trend in the long-wave flux.

All of the models for which the downwelling long-wave radiation is available reproduce the shape of the observed frequency distribution (Fig. 8). For the three models that under-predict the flux, REMO-EC4, REMO-GKSS and SMHI, the whole distribution is shifted towards the lower fluxes in a similar manner, hinting at an under-predicted long-wave cloud forcing in these models. For RACMO, the distribution is shifted towards

Table 8. Mean, standard deviation (sd), and systematic error, error standard deviation (esd), and correlation coefficient (r) relative to the observed value for the daily mean downwelling short-wave flux (W/m^2) for the three-monthly period at the four stations. Bold face numbers indicate a bias significant at the 99.5% confidence level

Station	source	mean	sd	bias	esd	r
Schleswig	obs	132	80			
	DWD-BM	122	71	**−10.3**	21	0.97
	REMO-GKSS	136	71	3.9	25	0.95
	REMO-EC4	142	82	**10.1**	35	0.90
	REMO-DWD	124	71	−7.7	33	0.91
	RACMO	114	88	**−19.6**	45	0.86
	SMHI	149	79	**17.1**	34	0.91
Lund	obs	127	85			
	DWD-BM	116	74	**−10.9**	23	0.97
	REMO-GKSS	130	76	2.8	26	0.95
	REMO-EC4	139	83	11.6	46	0.85
	REMO-DWD	116	74	**−11.0**	34	0.92
	RACMO	104	86	**−25.7**	45	0.86
	SMHI	137	79	10.0	39	0.89
Borlänge	obs	108	81			
	DWD-BM	95	66	**−13.1**	32	0.93
	REMO-GKSS	113	72	4.6	33	0.91
	REMO-EC4	126	82	**17.6**	42	0.87
	REMO-DWD	102	69	−5.9	30	0.93
	RACMO	86	71	**−23.3**	46	0.82
	SMHI	130	79	**21.7**	31	0.92
Luleå	obs	108	71			
	DWD-BM	101	67	**−6.5**	19	0.96
	REMO-GKSS	104	68	−3.2	27	0.93
	REMO-EC4	114	77	6.6	31	0.91
	REMO-DWD	89	66	**−18.4**	38	0.85
	RACMO	78	67	**−31.2**	37	0.85
	SMHI	113	73	4.9	27	0.93

higher fluxes despite the under-predicted mean cloud amount noted earlier.

3.5 Surface sensible heat flux

Reliable estimates of the surface sensible heat flux over land can only be collected for small areas using tower measurements. Sensible heat flux over the Baltic Sea was measured at a small number of ships, but these were also not considered to be representative. Therefore, estimates on the spatial scale of the model domain were not available. Consequently, the results shown here reflect the variability between the participating models rather than a deviation from observations.

Figure 9 shows a comparison of daily averaged sensible heat flux over all land grid points in the Baltic sea catchment area. Large differences can be seen, in particular during August 1995. For land areas two major causes may be responsible for this variability: differences in modelled net radiation, or different partitioning of net radiation over sensible and latent heat flux at the land surface. From the time series of the Bowen ratio (defined as the ratio between daily averaged sensible and latent heat flux) shown in Fig. 10, a similar variability between models is displayed. This is clearly related to the land surface treatment by the models. A considerable contribution to the wide range of results shortly after initialisation is caused by differences in soil water initialisation. DMI-CSE soil moisture fields had to be converted into model specific soil properties, and this conversion is not always straightforward (Shao et al., 1994). Small differences in absolute soil moisture amount may lead to significant differences in transpiration stress in the evaporation submodel of the land surface parameterisation.

UKMO appears to simulate fairly high Bowen ratio values, whereas RACMO and REMO-DWD

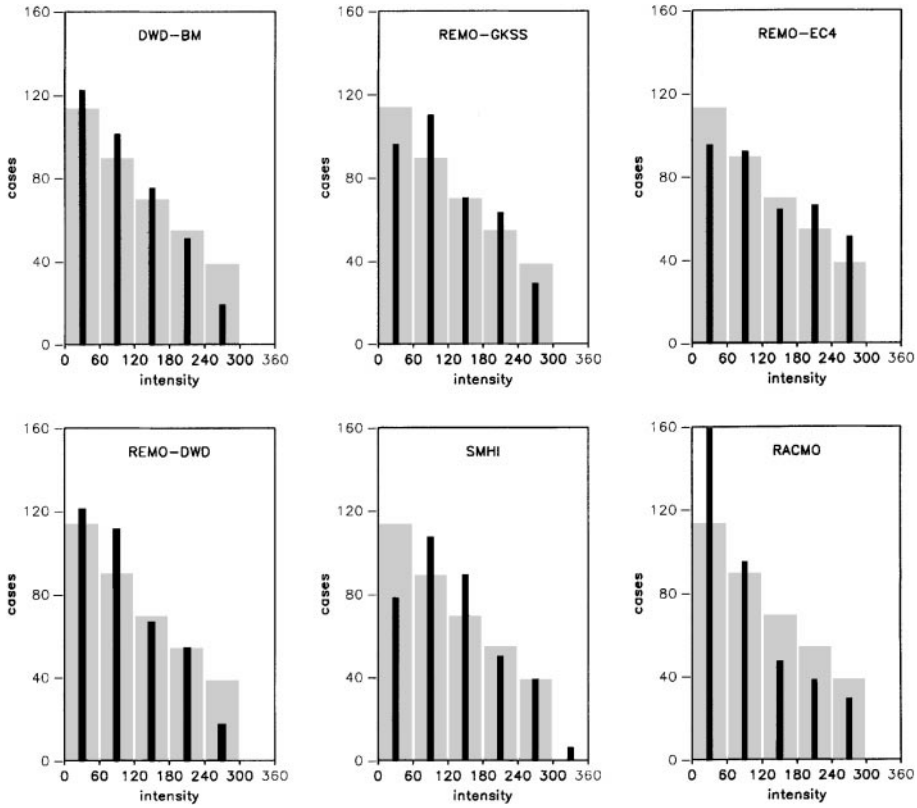

Fig. 7. Histograms of observed and modelled daily mean downwelling short-wave fluxes (W/m^2) at the surface during August, September and October 1995, for ensembles of the stations Schleswig, Lund, Borlänge and Luleå. The observed distribution is shown by shaded rectangles and the model results by black columns

are apparent outliers with low values. The remaining models simulate Bowen ratio's relatively close to each other for the first two months of simulation. RACMO is the only model in which downward sensible heat fluxes during stable stratification are enhanced by adopting the formulation of Mahrt (1987). This enhancement was introduced in RACMO during sensitivity experiments aiming at reducing the strong cold bias in the ECMWF land surface model (Viterbo et al.,

Table 9. As for Table 8, but for the downwelling longwave radiation flux

Station	source	mean	sd	bias	esd	r
Schleswig	obs	338	28			
	REMO-GKSS	327	32	**−10.7**	13	0.91
	REMO-EC4	334	26	−3.2	17	0.80
	RACMO	342	23	4.2	18	0.77
	SMHI	323	24	**−13.7**	15	0.83
Lund	obs	337	29			
	REMO-GKSS	324	36	**−13.0**	14	0.93
	REMO-EC4	330	31	−7.7	19	0.79
	RACMO	341	24	3.4	16	0.83
	SMHI	322	25	**−14.9**	14	0.88
Borlänge	obs	319	37			
	REMO-GKSS	307	41	**−11.8**	16	0.92
	REMO-EC4	305	35	**−14.0**	20	0.84
	RACMO	322	30	2.5	21	0.81
	SMHI	301	30	**−17.4**	15	0.91
Luleå	obs	293	34			
	REMO-GKSS	289	38	**−4.9**	14	0.93
	REMO-EC4	291	36	−2.0	16	0.90
	RACMO	313	30	**18.1**	17	0.85
	SMHI	290	34	−3.9	15	0.91

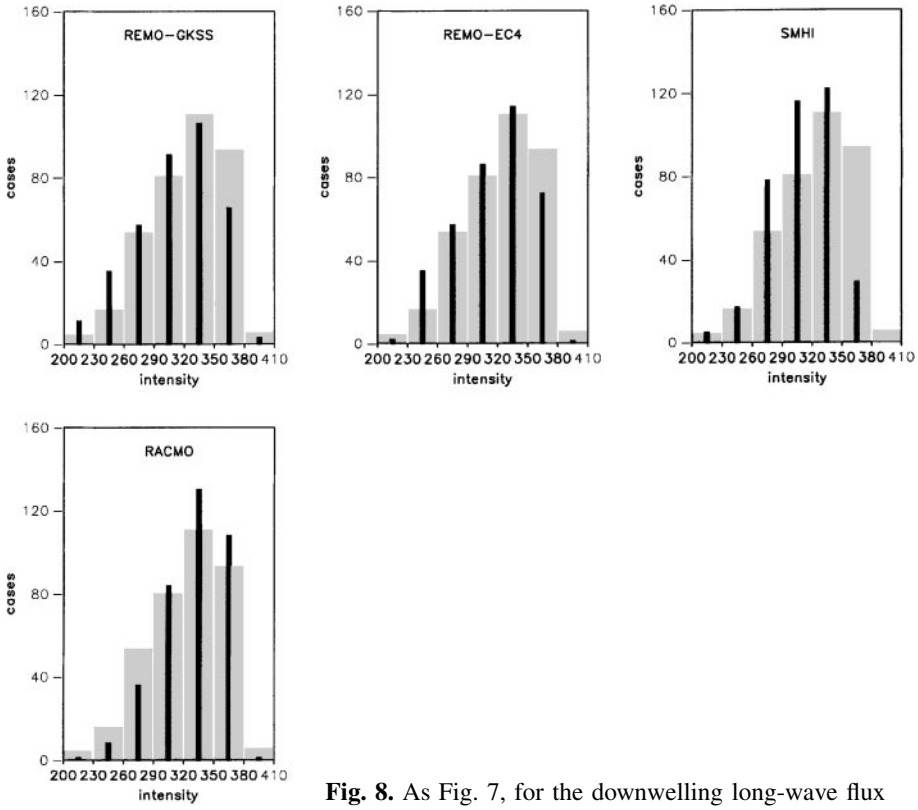

Fig. 8. As Fig. 7, for the downwelling long-wave flux

1999). The effect of this modification is clearly present in the comparison of RACMO to the other models, in particular later in the simulation period where stable conditions increasingly occur.

Differences over sea are much less pronounced (Fig. 9, lower panel), albeit that again RACMO seems an outlier giving lower fluxes on average. Apart from DWD-BM, all models use a common lower temperature boundary condition, and differences may result from either difference in the surface roughness parameterisation or in the simulation of the inner domain flow. Time-averaged surface heat fluxes remain directed upward, indicating that the sea is still a net source of sensible heat in this part of the season.

3.6 Evaporation

The inter-comparison of evaporation is performed separately for land grid boxes and for water grid boxes. As for the sensible heat fluxes, observations were not available as validation material. Because of the sparse data evaporation is one of the most uncertain parameters in the water cycle. For parts of the Baltic Sea surface evaporation data have been derived from COADS data set (Isemer, 1998, private communications).

Time series of daily mean surface evaporation rates over land are shown in Fig. 11. All models show a decrease of evapotranspiration from summer to autumn of roughly 0.5 mm/month. Most model results are quite close to each other, being less than 10 mm/month for August and September, and less than 5 mm/month in October (Table 10). Exceptions are the DMI-CSE assimilation runs at the lower end of the range, and REMO-DWD being particularly wet. As concluded earlier, this was related to the scaling of initial soil water content to appropriate initial conditions for the REMO-DWD model.

The land surface parameterisation in DMI-CSE and SMHI make use of a monthly prescribed soil moisture content in a deep layer, acting as a relaxation of the surface evaporation to a seasonal climatology. This relaxation is shown clearly in the temporal structure of the deviation of the simulated evaporation by DMI-CSE and SMHI in Fig. 11, suggesting a step-wise evaporation re-

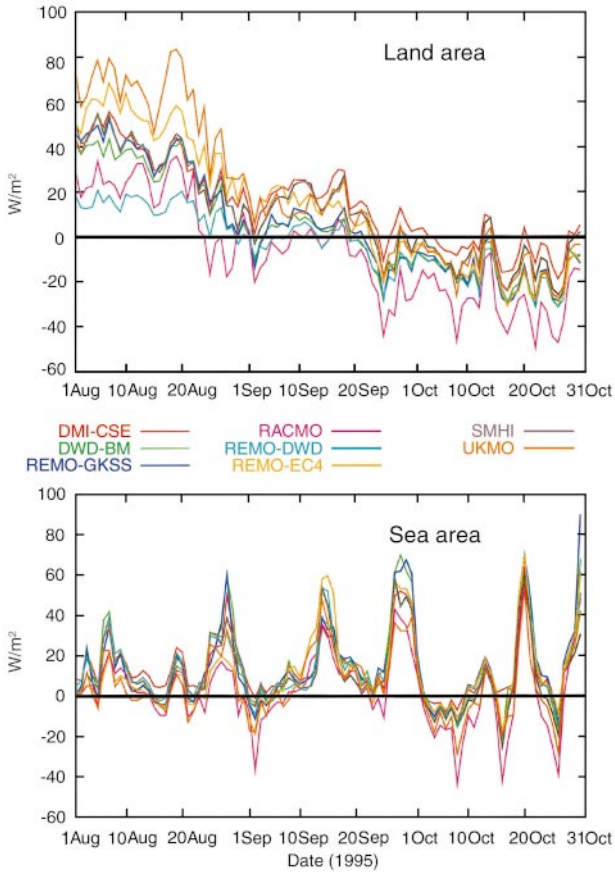

Fig. 9. Timeseries of simulated daily averaged sensible heat flux (W/m^2) over the Baltic land area (upper panel) and sea area (lower panel) during the PIDCAP period

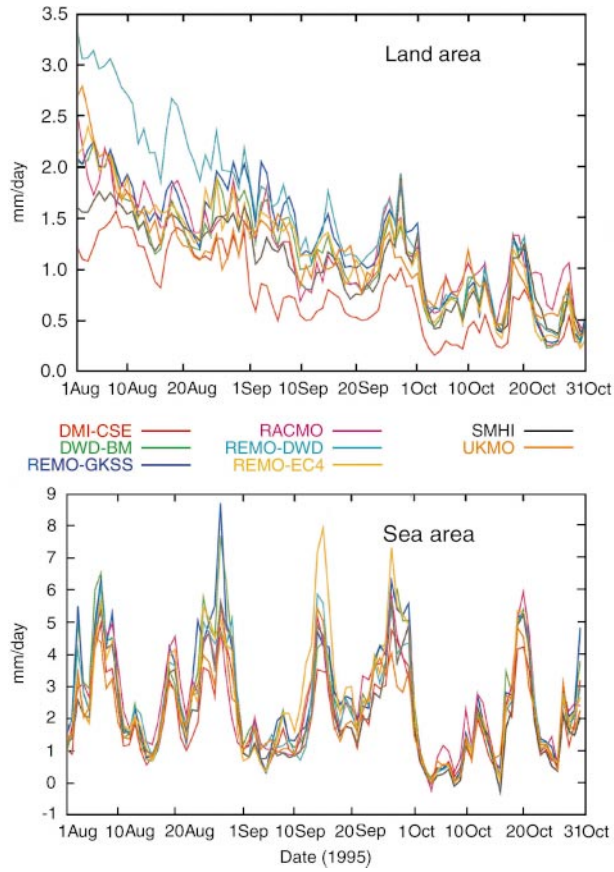

Fig. 11. Daily means of evaporation (mm/day) from the land surface (upper panel) and sea and lakes (lower panel) inside the Baltic Sea catchment area

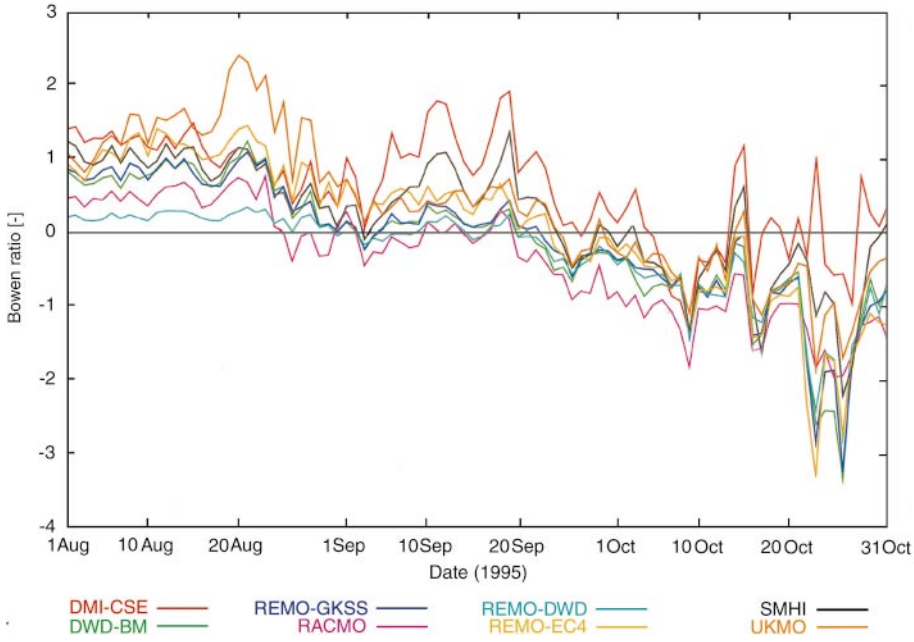

Fig. 10. Time series of Bowen ratio over the Baltic land area

Table 10. Monthly sums of evaporation from the land surface inside the Baltic Sea catchment area (mm/month)

	August 1995	September 1995	October 1995
DMI-CSE	38.5	20.6	12.1
DWD-BM	51.0	36.0	20.3
REMO-GKSS	55.6	41.7	20.8
REMO-DWD	75.9	43.0	21.7
REMO-EC4	53.1	37.0	18.8
SMHI	46.3	31.6	20.6
RACMO	60.0	34.4	25.9
UKMO	49.5	34.5	22.0

duction at monthly intervals. This behaviour was also demonstrated by Viterbo and Hollingsworth (1996).

All models show a modest delayed reaction to heavy precipitation events (e.g., precipitation on 15–16 August increasing evaporation on 17 August). For most days the spread of the evapotranspiration values between the models is less than 1 mm/day.

Over water there are no such systematic differences between the models as over land (see Fig. 11, lower panel). All models show almost the same temporal behaviour. Again the values are closest for October. The decrease in evaporation from August to October is not so obvious as for the land points, as to be expected. There is also a larger day by day variability in the model predictions. The evaporation over water changes sometimes more than 6 mm within a few days. This is twice as much as the overall maximum value in evapotranspiration over land. For two of these maxima some models produce significantly higher values than the rest of the models. At the end of August, the REMO-DWD and REMO-GKSS evaporate about 3 mm/day more than the other models. In the mid of September REMO-EC4 results are about 2.5 mm/day higher than those from the other models.

The spread in the results of the majority of the models ranges from approximately 1 mm/day for the periods with relatively low evaporation to about 2 mm/day for periods with higher evaporation. This spread is thus about twice as much as for the majority of models over the land area (Fig. 11). The total evaporation for the three months in the PIDCAP period are shown in Table 11. The monthly decrease in evaporation for the most models is about 20 mm. A stronger decrease is found in the evaporation derived from ship observations of meteorological standard parameters for the Baltic Proper and the western Baltic Sea (Isemer 1998, personal communication). The observations can only give a rough assessment because they are not valid for the whole Baltic Sea and there are no observations from lakes included. The maximum spread in monthly values among the models is about 30 mm in August and September, and 15 mm in October.

Table 11. Monthly sums of evaporation from the Baltic Sea and lakes inside the Baltic Sea catchment area (mm/month). Also shown are COADS observations. Data of Bothnian Bay, Bothnian Sea, and Gulf of Finland are not included in the observations

	August 1995	September 1995	October 1995
DMI-CSE	73.6	61.6	45.4
DWD-BM	99.6	78.4	55.2
REMO-GKSS	104.0	81.4	59.5
REMO-DWD	96.3	74.1	53.2
REMO-EC4	91.3	92.3	53.4
SMHI	87.5	63.6	46.3
RACMO	93.8	74.4	59.0
UKMO	77.5	69.3	49.3
COADS	92.9	51.3	38.3

Table 12. Precipitation minus evaporation (mm/day) averaged for the land and water area in the Baltic Sea catchment domain, separately for 2 periods

Model	1–23 August	24 August–31 October
DMI-CSE	−0.25	1.61
DWD-BM	−0.69	1.29
REMO-GKSS	−0.84	1.39
REMO-DWD	−1.62	0.83
REMO-EC4	−0.92	1.46
SMHI	−0.87	1.41
RACMO	−1.08	1.65
UKMO	−1.10	1.12

The differences in the surface evaporation have a clear impact on the hydrological cycle of the Baltic Sea catchment area, even in a model configuration where lateral boundaries are standardised (Table 12). The total amount of water advected to or from the entire Baltic Sea catchment (including the land, sea and lake area) was calculated as precipitation minus evaporation

($P - E$) for two separate episodes in the simulation period: the summer weather period preceding the first storm (1–23 August) and the remaining part of the simulation period. All models show a net atmospheric export of water from the catchment ($P - E < 0$) for the early summer period and a net import in the second part. However, the high land evaporation shown in REMO-DWD and the low evaporation rates in DMI-CSE are not totally compensated by modified precipitation but affect the net $P - E$, as shown in Table 12.

3.7 Runoff

Average total daily "runoff generation" to the Baltic Sea Drainage Basin from different models during the PIDCAP period is plotted in Fig. 12. Runoff generation is point runoff in each grid box and corresponds to what is commonly referred to as "effective precipitation" by hydrologists. This is the instantaneous excess water without any translation or transformation for either groundwater, lake and channel storage, or transport time. It should not be confused with the direct runoff concentrated in rivers. It cannot be directly compared to recorded river discharge observations without additional calculations to route the runoff to river mouths.

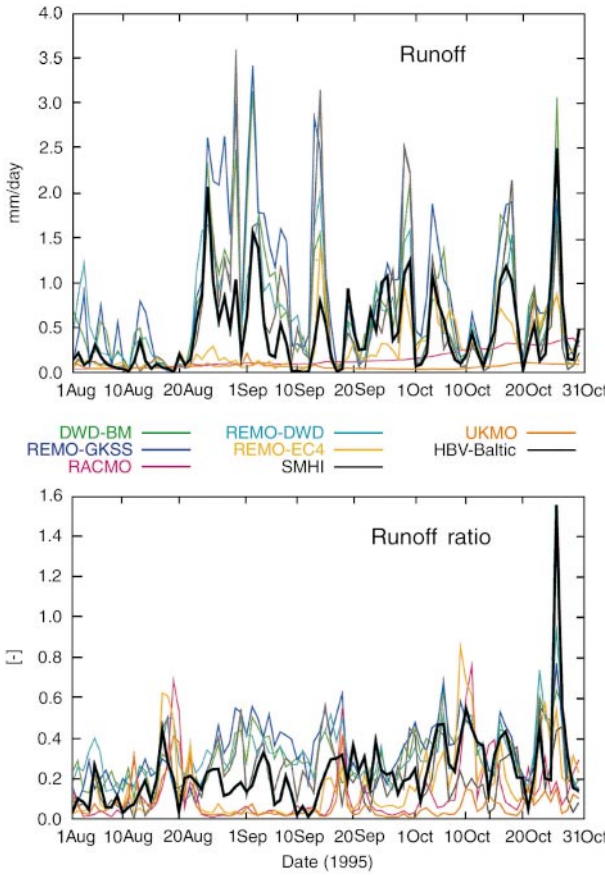

Fig. 12. Upper panel: simulations of total runoff (mm/day) to the Baltic Sea during the PIDCAP period; Lower panel: total runoff normalised by corresponding precipitation. For each curve, the precipitation generated by the corresponding atmospheric model is used. For HBV-Baltic, the precipitation forcing based on precipitation gauge analysis is used

Corresponding output from the HBV-Baltic model (Graham, 1999) is also plotted in Fig. 12. HBV-Baltic is a large-scale hydrological model of the water balance for the total Baltic Sea drainage basin. The precipitation forcing is derived from an analysis of synoptic station data, slightly different from the precipitation analysis of Rubel and Hantel (1999) shown in Sect. 3.1. The model is calibrated against river flow observations in the catchment area. Runoff generation from this model thus represents the signal of river flow in a variable that can be compared to atmospheric model runoff generation. HBV-Baltic has been used in other studies investigating the runoff generation processes in GCMs (Graham and Jacob, 2000; Van den Hurk et al., 2000).

Figure 12 shows a wide range of variation between the models. Of the seven atmospheric models compared, three indicate markedly little runoff response. These are the UKMO, REMO-EC4 and RACMO models. They all show low values of runoff for most of the period. The UKMO runoff is almost constant, whereas RACMO appears to linearly increase over the period. In both cases the surface runoff component, which should respond rapidly to rainfall events, is very small. This could result from specifying the soil texture to be too coarse or from inadequacies in the surface runoff formulation such as the lack of a representation of soil heterogeneity. REMO-EC4 runoff starts the period with low values and then begins to respond with more intensity toward the middle of the period. In spite of the parameterisation of a variable soil saturation in the REMO-EC4 model (Dümenil and Todini, 1992) which generally enhances the amount of quick runoff in response to precipitation, it takes nearly one month for the REMO-EC4 model to fill the soil water reservoir enough to produce runoff rates that are comparable to HBV-Baltic. The other four models show more

daily variation, indicating a response more in accordance with the runoff signal indicated by the HBV-Baltic results.

In order to evaluate the impact of the differences in modelled precipitation amounts, Fig. 12 also shows the runoff ratio, which is the daily runoff divided by the daily precipitation. Model generated precipitation is used for the normalisation, whereas for the HBV-Baltic series the analysed precipitation forcing is used. The results show that the mutual differences are maintained, and are thus inherent in the parameterisation or initialisation of the runoff in the participating models. Simulations by REMO-GKSS, DWD-BM, REMO-DWD and SMHI are fairly similar, but appear to generate more runoff during the late August and early September storm events than HBV-Baltic. RACMO, REMO-EC4 and UKMO show high runoff ratio's early in the simulation period only at occasions with little precipitation, but in general produce relatively little runoff.

3.8 Comparison with near surface observations

The models have been verified against synop observations from 00, 06, 12 and 18 UTC that are evenly distributed over the catchment area (see Fig. 1). Observations used are temperature and relative humidity at 2 m and wind speed at 10 m. The statistics for all models and stations are presented in Table 13.

The average diurnal cycle of temperature at 2 m is relatively well simulated by the models (Fig. 13). All models, but RACMO, have got a small mean negative bias, and a RMS error close to 2 °C. RACMO uses a more efficient vertical exchange under stable conditions, which was implemented in order to avoid cold wintertime bias as commonly observed in many NWP models (see Sect. 2). The models have the similar daily variation in bias with an underestimation of the diurnal cycle. The seasonal cycle of the temperature is well captured by all models, although REMO-EC4 has far too low temperature at the end of the period (not shown). Agreement with the relative humidity observations is also considered to be good. The RMS is in the range of 10–15% and the bias is less than 5%. The diurnal variation of the bias is opposite to the temperature bias.

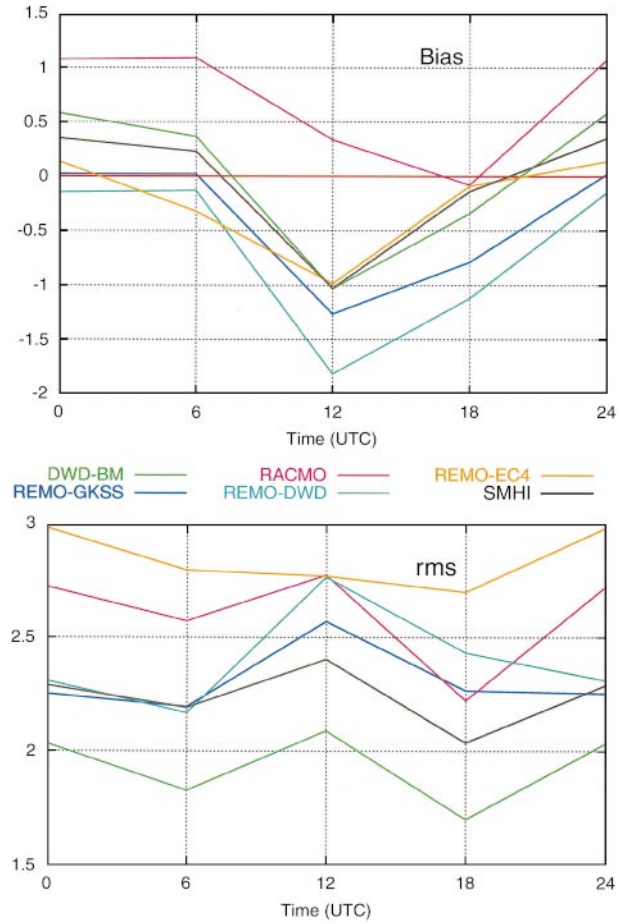

Fig. 13. Diurnal cycle of average bias (top) and rms (bottom) of 2 m temperature (°C) for the Synop locations shown in Fig. 1

Table 13. Statistics from model results compared with observations from 8 stations at 00, 06, 12 and 18 UTC

	Temperature 2 m (°C)		Relative humidity 2 m (%)		Wind speed 10 m (m/s)	
Model	Bias	RMS	Bias	RMS	Bias	RMS
DWD-BM	−0.14	1.9	−0.69	12.33	1.14	2.42
REMO-GKSS	−0.53	2.33	−0.93	13.16	1.11	2.41
SMHI	−0.17	2.23	3.66	11.74	0.26	2.19
RACMO	0.58	2.57	−3.41	13.66	0.74	2.52
REMO-DWD	−0.84	2.43	3.06	14.11	0.90	2.39
REMO-EC4	−0.34	2.80	−0.93	15.22	1.27	2.78

All models are overestimating the average wind speed, with a general diurnal cycle pattern consistent with the reduced amplitude in temperature (not shown).

As can be seen from Table 13 the difference between the models is very small. All models are capable of reproducing the surface variables well and the difference between the integration with and without data assimilation are surprisingly small. The near-surface variables considered are determined more by the physics in the models than by the initial state. The fact that the different physical parameterisations display the same diurnal error variations points at a common weakness in the representation of diurnal variations of the boundary-layer structure.

4. Discussion and conclusions

In a comparison of a group of 8 limited area atmospheric model integrations of a 3-month summer period in the Baltic Sea area, a fair degree of correspondence between the calculated characteristics of the energy and water cycles in the area was found. Five models were operated in climate mode, implying that analysed large scale atmospheric patterns entered the model domain only through the lateral boundaries. These models were generally able to reproduce the large scale weather phenomena rather well, providing a valuable environment to evaluate the model performance under constrained and close to observed weather conditions.

In qualitative sense, time series of *precipitation* produced by either climate mode model simulations and runs performed in assimilation or short-range forecast modes agreed relatively well. Correspondence was optimal in August, but significantly different precipitation time series with relatively low correlation with rain-gauge data were produced, in particular during September 1995. Convective precipitation has probably played a larger role during this month than in August and October, giving rise to larger RMS-errors due to the scattered and more localised nature of this form of precipitation. However, a systematic overestimation of precipitation produced by the ECHAM4 scheme present in REMO-EC4 and RACMO is apparent from this comparison.

In contrast to earlier findings (such as reported by Christensen et al., 1997) no systematic low summer precipitation bias was found in this study. The main reason is probably the fact that the most important mechanism leading to a low summer precipitation bias (positive feedback leading to a decrease in land surface evaporation, precipitation and soil moisture content) does not play a major role in the area under study here. Although the evaporation-precipitation feedback was not explicitly quantified, the main portion of the precipitation in the area probably originates from the Atlantic Ocean and Baltic Sea, thereby reducing the impact of the land surface feedback. This was also suggested by Koster et al. (2000).

In the data set used, the integrated specific humidity derived by GPS is typically underestimating the amount of atmospheric water vapour by approximately 1 kg/m^2 (up to 3 kg/m^2 in some cases), compared to radiosonde and microwave radiometer data. During the PIDCAP period the different GPS networks were being installed and changes in the electromagnetic environment at the GPS antennas, e.g., different types of radomes, influenced the estimates in a systematic way. Emardson et al. (2000) discuss the long term stability effects in detail. An evaluation of the GPS data quality especially for the PIDCAP period was carried out by Emardson et al. (1998). Being aware of the importance of a stable electromagnetic environment at the GPS antennas it shall in the future be possible to keep systematic effects below the 1 kg/m^2 level.

Generally the high variability of the observed IWV is in good agreement with the simulation results of all models. However, certain deviations mainly caused by specific weather situations like cyclones, simulated differently by climate mode and assimilation/forecast mode runs show a larger scatter. RACMO and all models which are based on DWD-physical parameterisation schemes (DWD-BM, REMO-DWD, REMO-GKSS in this order) overestimate the IWV. The reason for this is not clear until now. The positive bias in DWD-BM may also be related to the use of EM analyses rather than DMI-CSE. EM analyses have a positive bias probably caused by the use of surface cloud observations.

The investigation documented that the integrated specific humidity extracted from GPS-data is a valuable source of information which contributes to understanding and describing the large scale flow in atmospheric models for climate

and forecast issues. The usefulness for assimilation of GPS data into atmospheric models has recently been explored by several groups (Higgins, 2000). Ideally, accurate calibrated long term measurements of water vapour in the free troposphere are required.

The overall range of modelled average *cloud cover* is as large as 20–35% (depending on month and land or sea area). However, apart from the group of models with an apparent underestimation of cloudiness, a good agreement of within 10% from the cloud analysis is obtained. This range is well comparable to results from the Atmospheric Model Inter-comparison Experiment (AMIP; Gates et al., 1999), where cloud cover also shows up as a parameter for which models yield a wide range of results. In AMIP, differences of 20–30% or more in latitudinal averages were found in many models compared to ISCCP analysis data. The cloud cover in the physical parameterisations of ECHAM4 (REMO-EC4 and RACMO) and SMHI is significantly lower than observed.

Surprisingly, models using essentially identical physical parameterisation but with different dynamical formulations (REMO-EC4 and RACMO; REMO-DWD and DWD-BM) produced different average cloud amounts of about 10% in this study. Initialisation of soil moisture and small differences in scheme dependent coefficients are considered responsible for this feature. As a result of this model inter-comparison the cloud parameterisation in REMO-EC4 has been changed slightly and the new simulations of total cloud cover are in good agreement with observations.

The local day-to-day variability of both components of the *downwelling radiation* at the surface is reproduced fairly well. The models are able to explain 60–90% of the observed variance. However, systematic errors are present in all the models participating in the radiation comparison. The data suggest that these errors are related to deficiencies in modelling the clouds or their radiative effects. Short-wave radiation is systematically underestimated by the models carrying the DWD physical package pointing at too strong cloud forcing on the radiation. The reverse holds for REMO-EC4 and SMHI, who tend to overestimate surface short-wave radiation by a reduced cloud impact on the short-wave radiation. This is consistent with an underestimated cloud cover in these models, as noted earlier. This cloud cover underestimation was less pronounced in RACMO, which as a consequence under-predicts the downwelling short-wave flux. Whether this is a result of over-predicted liquid water content or too active radiative absorption by clouds is not clear.

The models REMO-EC4, SMHI and REMO-GKSS, compared to ground observations of longwave downwelling radiation, produce too low radiation amounts. REMO-GKSS and SMHI were also found to produce too low longwave radiation values in the study by van Meijgaard et al. (2001), who compared surface radiation data in The Netherlands to model output for a short period during PIDCAP. However, the flux in RACMO tends to be too large, but the bias varies from station to station.

Differences in simulated *sensible heat flux* over land shortly after the initialisation of the intercomparison experiment are considerable. This is likely to be caused by an incompatibility in the initial soil water content among the models. Changes in the heat exchange for stable stratification employed in RACMO have a clear effect on daily averaged sensible heat flux both over land and sea. However, as sensible heat flux data are not present it is difficult to decide whether this modification is to be considered beneficial or not. It helped switching the negative near surface temperature bias to a positive value, which suggests that the enhancement of downward sensible heat flux in RACMO is a bit too large.

Time series of simulated *evaporation* show similar patterns during the three months, both over land and over sea. Outliers are formed by REMO-DWD (high evaporation in August, too wet initial soil) and DMI-CSE (lowest evaporation in the whole period, with unknown cause). SMHI and DMI-CSE underestimate evaporation over land, and their simulations bear the signature of a climatological control. The results shown here are similar to those found by Källen (1999), who found an underestimation of the evaporation during in particular the summer season. Monthly evaporation over sea is generally about twice as high as over land, due to the lack of limited water supply. However, day-to-day variability of evaporation over sea is also stronger than over land. Differences in the surface evaporation rates lead to differences in the net atmospheric transport of

water to or from the catchment area, even for identical lateral boundary conditions.

Runoff generation from the UKMO, REMO-EC4 and RACMO models is consistently underestimated and relatively insensitive to changes in rainfall when compared to runoff generation from the calibrated water balance model HBV-Baltic. The REMO-GKSS, DWD-BM, REMO-DWD and SMHI models show a sharper temporal variation in runoff generation, but large differences in the magnitude of runoff are apparent between the models. Compared to HBV-Baltic too much precipitation is removed as runoff.

As far as comparisons to *near surface temperature, relative humidity and wind speed* are concerned, all models show a fair correspondence with the observations. The high horizontal resolution of the models appears adequate for describing local variations of these parameters. Near surface temperature remains systematically underestimated by all models except RACMO, and the amplitude of the diurnal cycle is underestimated. Average correspondence with observed relative humidity is within 5% and RMS values are not exceeding 15%, which is considered to be good. Wind speed is biased high in all models. A remarkably small difference is shown between models operated in forecast or assimilation mode and the models run freely in climate mode. The boundary conditions provided by the lateral boundaries and sea surface temperature apparently cause a firm constraint.

The models participating in this inter-comparison are under continuous development. The results presented here must therefore be regarded as a snap-shot taken from this suite of models. However, the model inter-comparison effort was successful in revealing differences between the models and observations, and it gives a first impression on the uncertainties involved in the modelling of the components of the water and energy cycles. Furthermore, the inter-comparison was an effective exercise for indicating areas of specific weaknesses of the different models, leading to further model developments. For example, a systematic underestimation of total cloud cover using REMO-EC4 has been detected. Changes in the treatment of clouds in REMO-EC4 has cured the problem for future applications. Problems with the runoff formulation in the ECMWF land surface scheme in RACMO, as revealed in this study, were subject to a further investigation (Van den Hurk et al., 2000). Also the poor simulation of runoff by the UKMO model indicates that its representation of surface runoff generation needs to be improved. Work has been initiated to incorporate and test surface hydrology schemes which represent soil heterogeneity such as those of Moore (1985), and Sivapalan et al. (1987).

Acknowledgements

Prof. Lennart Bengtsson (MPI) has successfully taken the lead in the NewBaltic I and II research projects. The cloud cover analysis has been kindly carried out by Dr. Rosenow, Meteorologisches Observatorium Potsdam (DWD). Erik van Meijgaard (KNMI) and Thomas Winter (DWD) have provided useful contributions to discussions and analyses. Two anonymous reviewers provided many useful and adequate comments, and this is appreciated. This project is partially supported by the European Committee under contract numbers ENV4-CT95-0072 and ENV4-CT97-0626.

References

Bengtsson L (1995) Baltic Sea Experiment BALTEX: Initial implementation plan. International Baltex Secretariat Publ No 2, 84 pp. (Available from Max-Planck-Institute for Meteorology, Hamburg)

Bengtsson L (ed) (1998) Numerical studies of the energy and water cycle of the Baltic region. Final report covering the period 1/1/1996–30/4/1998; EU-project ENV4-CT95-0072, 247 pp. (Available from Max-Planck-Institute for Meteorology, Hamburg)

Bengtsson L (ed) (2000) Numerical studies of the energy and water cycle of the Baltic region (NEWBALTICII). Final report covering the period 1/4/1998–31/3/2000; EU-project ENV4-CT97-0626, 204 pp. (Available from Max-Planck-Institute for Meteorology, Hamburg)

Brinkop S, Roeckner E (1996) Sensitivity of a general circulation model to parameterizations of cloud-turbulence interactions in the atmospheric boundary layer. Tellus 47A: 197–220

Christensen JH, Christensen OB, Lopez P, van Meijgaard E, Botzet M (1996) The HIRHAM4 Regional Atmospheric Climate Model. Danish Meteorol Institute, Copenhagen, Scientific Report 96-4, 51 pp

Christensen JH, Machenauer B, Jones RG, Schär C, Ruti PM, Castro M, Visconti G (1997) Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. Climate Dynamics 13: 489–506

Cox PM, Betts RA, Bunton C, Essery RLH, Rowntree PR, Smith J (1999) The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Climate Dynamics 15: 183–203

Cullen MJP (1993) The unified forecast/climate model. Meteorol Mag 122: 81–94

Dolman AJ, Gregory D (1992) The parametrization of rainfall interception in GCMs. Q J R Meteorol Soc 118: 455–467

Dümenil L, Todini E (1992) A rainfall-runoff scheme for use in the Hamburg climate model. In: O'Kane JP (ed) Adv Theor Hydrology. European Geoph Soc Series on Hydrological Sciences, Vol 1, Elsevier Science Publishers, Amsterdam, pp 129–157

Emardson TR, Elgered G, Johansson JM (1998) Three months of continuous monitoring of atmospheric water vapor with a network of Global Positioning System receivers. J Geophys Res 103: 1807–1820

Emardson TR, Johansson JM, Elgered G (2000) The systematic behavior of water vapor estimates using four years of GPS observations. Trans IEEE Geoscience and Remote Sensing GE-38: 324–329

Gates W et al. (1999) An overview of the results of the Atmospheric Model Intercomparison Project (AMIP). Bull Am Met Soc 80: 29–55

Graham LP (1999) Modeling runoff to the Baltic Sea. Ambio 28: 328–334

Graham LP, Jacob D (2000) Using large-scale hydrologic modeling to review runoff generation processes in GCM climate models. Meteorol Z 9: 49–57

Gregory D, Rowntree PR (1991) A mass flux convection scheme withrepresentation of ensemble characteristics and stability dependent closure. Mon Wea Rev 118: 1483–1506

GTOPO30: Global 30 sec topographic data set. By Internet from: Earth Resources Observation Systems (EROS) Distributed Active Archive Data Center (EDC DAAC) at US Geological Survey's EROS Data Center, Sioux Falls, South Dakota, USA

Higgins M (2000) Progress in 3D-variational assimilation of total zenith delay at The Met Office. To appear in the COST-716 Workshop edition of Physics and Chemistry of the Earth

Holtslag AAM, de Bruin HAR (1988) Applied modeling of the nighttime surface energy balance over land. J Applied Meteorol 27: 689–704

Holtslag AAM, Boville BA (1993) Local versus non-local boundary layer diffusion in a global climate model. J Climate 6: 1825–1842

Jacob D (2001) Investigation of the annual and interannual variability of the water budget over the Baltic Sea drainage basin using the regional climate model REMO. Meteorol Atmos Phys (this issue)

Jacob D, Podzun R (1997) Sensitivity studies with the regional climate model REMO. Meteorol Atmos Phys 63: 119–129

Jacobsen I, Heise E (1982) A new economic method for the computation of the surface temperature in numerical models. Beitr Phys Atm 55: 128–141

Källen E (ed) (1996) HIRLAM documentation manual. System 2.5. Available from Swedish Meteorological and Hydrological Institute, SE-60176 Norrköping, Sweden

Källen E (ed) (1999) SweClim Newsletter 6. Available from Swedish Meteorological and Hydrological Institute, SE-60176 Norrköping, Sweden

Koster RD, Suarez MJ, Heiser M (2000) Variance and predictability of precipitation at seaonal-to-interannual timescales. J Hydrometeorol 1: 26–46

Lehmann A, Zimmermann K, Graute C, Stein D (1999) Meteorological Data Centre for BALTEX – BALTEX MDC – Status Report No 3, Selbstverlag des DWD, Offenbach/M, 106 pp

Lenderink G, Van Meijgaard E (2001) Impacts of ccloud and turbulence schemes on integrated water vapor: comparison between model predictions and GPS-measurements. Meteorol Atmos Phys (this issue)

Loennberg P, Shaw D (1987) ECMWF Data Assimilation – Scientific Documentation. ECMWF Research Manual 1, 10/87, 2nd revised ed. Available from: ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, England

Louis J-F (1979) A parametric model of vertical eddy fluxes in the atmosphere. Boundary-Layer Meteor 17: 187–202

Mahrt L (1987) Grid averaged surface fluxes. Monthly Weather Rev 115: 1550–1560

Majewski D (1991) The Europa-Modell of the Deutscher Wetterdienst. ECMWF Seminar on Numerical Methods in Atmospheric Models, Vol 2, 147–191

Meijgaard E van, Andræ U, Rockel B (2001) Comparison of model predicted cloud parameters and surface radiative fluxes with observations on the 100 km scale. Meteorol Atm Phys (this issue)

Mellor GL, Yamada T (1974) A hierarchy of turbulence closure models for planetary boundary layers. J Atmos Sci 31: 1791–1806

Milton SF, Wilson CA (1996) The impact of subgrid-scale orographic forcing on systematic errors in a global NWP model. Mon Weather Rev 124: 2023–2045

Moore RJ (1985) The probability-distributed principle and runoff production at point and basin scales. Hydrol Sci J 30, 2: 273–279

Morcrette J-J (1989) Description of the radiation scheme in the ECMWF model; ECMWF TechMemo 165. Available from ECMWF, Reading, UK

Nordeng TE (1994) Extended versions of the convective parameterization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics; ECMWF TechMemo 206. Available from ECMWF, Reading, UK

Palmer TN, Shutts GJ, Swinbank R (1986) Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. Quart J Roy Meteorol Soc 112: 1001–1039

Ritter B, Geleyn J-F (1992) A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Mon Wea Rev 120: 303–325

Robock A, Schlosser CA, Vinnikov KY, Speranskaya NA, Entin JK, Qiu S (1998) Evaluation of the AMIP soil moisture simulations. Glob Plan Sci 19: 181–208

Roeckner E, Arpe K, Bengtsson L, Christoph M, Claussen M, Dümenil L, Esch M, Giorgetta M, Schlese U, Schul-

weida U (1996) The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate. Max-Planck Institute für Meteorologie, Report no 218

Rosenow et al. (1992) Proceedings of the 9th Meteosat Scientific Users' Meeting, Locarno, Switzerland

Rubel F (1998) PIDCAP – Ground truth precipitation atlas. Österreichische Beiträge zu Meteorologie und Geophysik 18: 76 pp

Rubel F, Hantel M (1999) Correction of daily rain gauge measurements in the Baltic Sea drainage basin. Nord Hydrol 30: 191–208

Sass BH, Yang X (1997) PIDCAP Re-analysis and Atmospheric Budget Diagnosis. Hydrological, Oceanic and Atmospheric Experience from BALTEX. Alesto M, Isemer HJ (eds) BALTEX, 135–142

Sass BH, Rantu L, Raisanen P (1994) Hirlam-2 radiation scheme: documentation and tests, HIRLAM Technical Report 16, 43 pp

Savijarvi H (1990) Fast radiations parameterization schemes for mesoscale and short range forecast models. J Appl Meteor 20: 437–447

Shao Y, Anne RD, Henderson-Sellers A, Irannejad P, Thornton P, Liang X, Chen TH, Ciret C, Desborough C, Balachova O, Haxeltine A, Ducharne A (1994) Soil moisture simulation – a report of the RICE and PILPS Workshop. IGPO Publ Series No 14, Clim Impact Centre, 179 pp

Sivapalan M, Beven K, Wood EF (1987) On hydrologic similarity. 2. A scaled model of storm runoff production. Water Resources Research 23: 2266–2278

Slingo A (1989) A GCM parametrization for the shortwave radiative properties of water clouds. J Atmos Sci 46: 1419–1427

Slingo A, Wilderspin R (1986) Development of a revised longwave radiation scheme for an atmospheric general circulation model. Q J R Meteorol Soc 112: 371–386

Smith RNB (1990) A scheme for predicting layer clouds and their water content in a general circulation model. Q J R Meteorol Soc 116: 435–460

Sommeria G (1985) Parametrization of land surface processes. Proc 1985 ECMWF seminars on physical parameterizations for numerical models of the atmosphere. Reading, UK, 233–264

Sundqvist H (1978) A parameterization scheme for non-convective condensation including prediction of cloud water content. Q J R Meteorol Soc 104: 677–690

Sundqvist H, Berge E, Kristjansson JE (1989) Condensation studies with a mesoscale numerical weather prediction model. Mon Wea Rev 117: 1641–1657

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon Wea Rev 117: 1779–1800

Van den Hurk BJJM, Graham LP, Viterbo P (2000) Atmospheric simulations of runoff to the Baltic Sea. KNMI Preprints 2000–04; submitted to J Hydrol

Viterbo P, Beljaars ACM (1995) An improved land surface parametrization scheme in the ECMWF model and its validation. J Climate 8: 2716–2748

Viterbo P, Hollingsworth A (1996) Comments on ''Observations of a monthly variation in global surface temperature data''. Geophys Res Lett 23: 693–694

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

Yang X, Sass BH, Elgered G, Johansson JM, Emardson TR (1999) A Comparison of Precipitable Water Vapor Estimates by an NWP Simulation and GPS Observations. J Appl Meteor 38: 941–956

Zhang Y, Rockel B, Stuhlmann R, Hollmann R, Karstens U (2000) REMO Cloud Modeling: Improvements and Validation with ISCCP DX Data (forthcoming in J Appl Meteor)

Corresponding author's address: Dr. Daniela Jacob, Max-Planck-Institute for Meteorology, Bundesstr. 55, D-20146 Hamburg, Germany (E-mail: jacob@dkrz.de)