RMetS

Royal Meteorological Society

Evaluation of the statistical cloud scheme in the ECHAM5 model using satellite data

T. Weber,[a][*][†] J. Quaas[a][‡] and P. Räisänen[b]

[a]*Max Planck Institute for Meteorology, Hamburg, Germany*
[b]*Finnish Meteorological Institute, Helsinki, Finland*

[*]Correspondence to: T. Weber, Helmholtz-Zentrum Geesthacht, Climate Service Center (CSC), Fischertwiete 1, 20095 Hamburg, Germany. E-mail: torsten.weber@hzg.de
[†]Current address: Helmholtz-Zentrum Geesthacht, Climate Service Center (CSC), Hamburg, Germany.
[‡]Current address: Institute for Meteorology, Universität Leipzig, Germany.

An evaluation of a statistical cloud scheme taking into account subgrid-scale variability for water vapour and cloud condensate in the ECHAM5 general circulation model of the atmosphere is presented. Three-dimensional modelled water vapour, cloud liquid water and cloud ice were distributed stochastically into subcolumns of each grid box and vertically integrated to total water path (TWP). Thus the lower atmosphere is emphasized in the evaluation of TWP due to its exponential profile. The edited model dataset was compared with the globally analyzed distribution of TWP measured by the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite instrument. The results show that the mean TWP and mean cloud cover are on average relatively well simulated. However, large deficiencies are revealed by the evaluation of both variance and skewness of the probability density function (PDF). Systematically negative deviations of variance are found for almost all regions of the globe. Skewness of the TWP is overestimated in the Tropics and underestimated at high latitudes. Moreover, sensitivity experiments were performed to reveal the deficiencies in the parametrization leading to the observed deviations of variance and skewness of TWP. It was found that the positive bias in skewness in the Tropics can be reduced by modifying the influence of convection on the PDF. Copyright © 2011 Royal Meteorological Society

1. Introduction

The modelling of clouds in general circulation models (GCMs) needs a proper theoretical framework for the representation of the subgrid-scale humidity and temperature variations and a balance between the complexity of the parametrization and the computational efficiency. One approach to handle these demands is to apply statistical cloud schemes. They are able to take into account the spatial variability of a modelled quantity in the grid box in terms of a probability density function (PDF). Pincus and Klein (2000) found that by accounting for the subgrid-scale variability in microphysical processes in cloud schemes, the bias in the nonlinear process rates caused by using averaged quantities can be reduced and arbitrary tuning of parameters can also be avoided. Early statistical schemes were developed by Sommeria and Deardorff (1977) and Mellor (1977). They used a joint PDF for the subgrid-scale liquid temperature and total water content to estimate the cloud fraction by integrating over the saturated part of the PDF. This concept implies that

the saturated part of the PDF condenses immediately and that the chosen PDF shape is able to reproduce the spatial distribution of the total water content.

The simplest form of a PDF is based on a symmetric uniform distribution and is used by LeTreut and Li (1991) to describe the horizontal variability of the sum of water vapour and cloud condensate in a grid box. A similar approach was made by Smith (1990), who uses a relative humidity threshold function derived from a symmetric triangular PDF diagnosing the variance of cloud water content. This simple condensation scheme was adopted by Rotstayn (1997) and Nishizawa (2000). Likewise a symmetric, but Gaussian, form of PDF was applied by Bechtold *et al.* (1995) to model the cloud water content, partial cloudiness and liquid-water flux in stratocumulus cases. It is evident that these simple PDFs are not able to reproduce the spatial distribution of the respective quantity in all cases, but they are easy to handle and have a low computational cost. Using a PDF that is also skewed, a quantity can be described by the same PDF in different cases. A beta function having this capability is employed by Tompkins (2002) to represent the spatial distribution of total water mixing ratio, and Bony and Emanuel (2001) used a log-normal distribution to model the total water content. Bougeault (1981) showed that a skewed exponential distribution can be applied to parametrize clouds in a one-dimensional model.

More comprehensive concepts were also developed by Watanabe *et al.* (2009), who use double-uniform and skewed-triangular distributions in different conditions for temperature and total water content fluctuations. Golaz *et al.* (2002) apply a joint PDF of a double-Gaussian function of vertical velocity, liquid water potential temperature and total specific water content to characterize the unresolved subgrid variations in a grid box. A better representation of low-level boundary-layer clouds and an improved atmospheric boundary-layer structure were achieved by Kuwano-Yoshida *et al.* (2010) using joint-Gaussian PDFs of the liquid water potential temperature and total water content in the atmospheric general circulation model for the Earth Simulator (AFES).

This short overview of various statistical PDF schemes shows that many different forms of PDFs are used to describe a certain quantity in cloud microphysical processes. The reasons are on the one hand the complexity of the scheme, e.g. how many parameters are needed to determine the shape of the PDF, and on the other hand the particular humidity-related quantity to be simulated in the model. Although the choice of assumed shape of the humidity distribution is one crucial criterion for successful performance of a statistical PDF scheme, there is an absence of information about the respective PDF from measurements. Several observational studies were carried out to analyze the shape of the distributions for humidity-related quantities in different conditions. Wood and Field (2000) found complex and often bimodal PDFs of total water content with large values of skewness in data from flights through stratocumulus clouds, which capped a well-mixed planetary boundary layer (PBL) or were decoupled from it. Similar results for PDFs of liquid water content that were bimodal as well as positively and negatively skewed were reported by Davis *et al.* (1996), measured during flights through marine stratocumulus clouds. Moreover, in the PBL various shapes of humidity distributions such as Gaussian, skewed,

platykurtic and multimodal were classified by Price (2001), examining data from tethered-balloon measurements.

Another approach to obtain information regarding the distribution of humidity-related quantities is to utilize cloud-resolving models (CRMs) simulating meteorological cases in a determined region. Such an approach was taken, additionally to the use of observational data, by Bony and Emanuel (2001), who found close to Gaussian forms of PDFs of total water at low levels and skewed ones at high levels in CRM simulations of the GATE Phase III experiment. Different results were achieved by Tompkins (2002), who found more beta distribution forms of PDF of total water mixing ratio in data produced from the large-eddy simulation model of the Met Office. The studies show that the shape of the measured PDF varies a lot depending on the analyzed data (flight measurements or CRMs). Flight datasets have the advantage of being results from measurements of the atmosphere, but they depend on the environmental conditions and are only available on a small path along the flight route, i.e. it is not possible to measure the quantity over a large area at the same time. Simulated data from CRMs possess this feature, however the results depend on the chosen spatial and temporal resolution as well as on the initial conditions needed to run the model. Satellite data combine the advantages of covering large regions and being real measurements of the atmosphere.

New, high spatial resolution (of order of 5×5 km^2) retrievals of column water vapour and column cloud condensate are available from satellites globally. These observational data allow construction of horizontal PDFs at the much coarser resolution of typical GCMs. Unfortunately, data with high spatial resolution only allow for the retrieval of vertically integrated total water (total water path, hereafter TWP), while instruments providing vertically resolved water vapour and cloud-condensate mixing ratio retrievals (such as infrared sounders) still have resolutions that are spatially too coarse (of the order of 20×20 km^2 and greater) to allow the construction of PDFs.

This article presents an evaluation of the subgrid-scale variability scheme of total water mixing ratio developed by Tompkins (2002) in the ECHAM5 general circulation model of the atmosphere. The defining parameters of the PDF, the modelled mean of TWP, its variance and skewness as well as the total cloud cover are compared with high-resolution satellite data. To make the three-dimensional modelled total water mixing ratio comparable to the two-dimensional satellite column total water retrieval, the modelled water vapour, cloud liquid water and cloud ice are vertically integrated using maximum overlap by the stochastic subcolumn generator developed by Räisänen *et al.* (2004) and added to TWP. Moreover, in sensitivity experiments processes are identified explaining part of the discrepancy between the parametrized variance and skewness and the observational data. In section 2, the model experiments and the methods of analyzing the model and satellite data are explained. Afterwards, the data derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite instrument are discussed in section 3. In section 4, the results of the comparison between the modelled parameters and the ones derived from MODIS as well as the results of the sensitivity experiments are presented. The evaluation closes with a summary of the results and conclusions in section 5.

2. Data and methodology

2.1. Model simulations

The parametrization for prognostically computing the subgrid-scale PDF of total water mixing ratio, as developed by Tompkins (2002) and implemented in the ECHAM5 GCM (Roeckner *et al.*, 2003) is briefly summarized in Appendix 5. For the evaluation of this cloud-cover scheme, the year 2004 (as characterized by the prescribed sea ice and sea-surface temperatures) was simulated with the ECHAM5 model starting three months earlier to reach an equilibrium for the cloud microphysical processes. The model experiments were carried out using a horizontal resolution of $2.8° \times 2.8°$ (approx. 310 km at the Equator) and a vertical resolution of 19 vertical levels with the uppermost pressure level at 10 hPa ($T42L19$). As boundary conditions for the model experiments, the observed monthly mean sea-surface temperature and sea ice data of 2004 were used as in the Atmospheric Model Intercomparison Project (AMIP: Gates *et al.*, 1999). The considered variables are the grid-box mean of total cloud cover, TWP and its variance and skewness calculated every hour. While Tompkins (2002) uses the distribution width in his scheme, here we employ the variance because the calculation of variance is less sensitive to outliers in the reference observational data and is therefore more robust. For comparison, one model simulation was done using the large-scale cloud-cover scheme by Sundqvist *et al.* (1989), based on a relative humidity threshold.

The model mean (vertically resolved) total water mixing ratio, its variance and skewness could not be compared with satellite data directly since these data are available only as a vertical integral on a two-dimensional grid. To solve the problem, the stochastic subcolumn-generator developed by Räisänen *et al.* (2004) was applied. It creates a number of subcolumns within each model grid box and distributes water vapour, cloud liquid water and cloud ice over the subcolumns using the total water mixing ratio PDF and a prescribed vertical overlap assumption to describe horizontal and vertical variations of each variable, respectively. In the experiments, the model grid boxes were divided into 1000 subcolumns and the beta function PDF from the statistical cloud scheme of the model was applied. Since the TWP is dominated by water vapour, the maximum overlap was chosen assuming that the distribution of water vapour overlaps in adjacent layers. In order to obtain a two-dimensional field for each grid box, the subcolumns were vertically integrated and subsequently the fields of water vapour, cloud liquid water and cloud ice were added to TWP. From that two-dimensional field, the mean of TWP and its variance and skewness were calculated. The total cloud cover determined with the maximum-random-overlap assumption in the model was taken directly without any modifications.

The application of the subcolumn-generator produces artificial negative values of skewness for the TWP, which were not simulated by the statistical cloud scheme in single model layers, since the scheme allows only positively skewed or symmetric distributions of total water mixing ratio. Three months experiments show that model grid boxes with negative skewness increase when random overlap or a vertical decorrelation length of 1000 m is applied. A similar result was observed when the number of subcolumns is decreased. The different overlap assumptions affect the variance of TWP as well. It was found that the mean variance is 83% lower when random overlap is used and 64% lower when a decorrelation length of 1000 m is applied, compared with the variance using maximum overlap. For the TWP skewness, the effect of the overlap assumption was less strong. Importantly for our main conclusions, the geographical distributions of variance and skewness are similar for all overlap assumptions tested. The main conclusions are not affected by the choice of the overlap assumption.

Because of the vertical integration of TWP and the approximately exponential profile of water vapour, which exhibits the largest amounts in the lower atmosphere, the TWP is mainly dominated by the PBL humidity. Consequently, the variance and skewness derived from the TWP are also mainly affected by the lower atmosphere and it can be assumed that both variables are governed more by the water vapour than by the cloud condensate, due to the higher contribution of the water vapour to the TWP. It is noteworthy that the PDF of TWP (the vertical integral) may not follow a beta function even when the distribution of total water mixing ratio has this form at most levels.

2.2. Satellite data

In order to evaluate the scheme, the observational dataset L2 Joint Product of MODIS on the Terra satellite (King *et al.*, 2003; Platnick *et al.*, 2003) was employed for the year 2004. Its high spatial resolution of $5 \times 5 \text{ km}^2$ permits us to capture the horizontal spatial subgrid-scale variability of column water vapour and cloud condensate very well at the much coarser model resolution of $T42$. As MODIS does not provide the TWP in one variable, the instantaneous retrievals of water vapour (Precipitable Water Infrared Clear) and cloud condensate (Cloud Water Path) including cloud ice had to be used. The reason for this is that MODIS can either retrieve the water vapour in clear sky or the cloud condensate in overcast or cloudy sky. The total column precipitable water vapour is determined in the retrieval algorithm by using the infrared wavelengths at 6.535–6.895, 7.175–7.475 and 8.400–8.700 µm. The instrument error for the water-vapour retrievals is estimated at about 10% (Seemann *et al.*, 2006).

To obtain reliable retrievals of water vapour that are not disturbed by clouds, only data pixels where the cloud mask of MODIS indicates 'confident clear' or 'probably clear' sky were selected. These conditions were also applied to the cloud mask in order to determine the total cloud cover, which was necessary to be consistent with water-vapour retrievals. The gaps occurring in the images due to clouds or missing retrievals were filled by a linear interpolation of the 8 nearest data pixels when at least 25% of the entire scene was cloud-free. If less than 25% of a scene was cloud-free, then the MODIS image was discarded from the analysis. Linear interpolation was applied under the assumption that the water vapour in columns containing clouds is similar to the concentration in surrounding clear skies. While this assumption is probably not strictly correct, we consider it good enough for this statistical analysis. The chosen threshold level of 25% is the result of a compromise between having enough data pixels for the interpolation and obtaining a large number of images (filled grid boxes) for a reliable statistical analysis, especially in those regions where cloudy conditions prevail. For instance, using this threshold,

about 40% of the images were rejected due to this constraint, compared to a 50% threshold rejecting 79% of the images. We assume that the large number of grid boxes from which the statistical quantities were calculated decreases the impact of potential errors caused by the interpolation.

The Cloud Water Path from the MODIS dataset contains both liquid and ice phase and was determined from the optical thickness and droplet effective radius using the visible and near-infrared wavelengths 0.645 and 0.858 or 1.240 and 2.130 μm (King *et al.*, 1997). Compared with other measurements such as those from the Advanced Microwave Scanning Radiometer–EOS (AMSR-E), MODIS overestimates the liquid water path in marine stratocumulus regions and poleward of 40° at all longitudes, while it underestimates the liquid water path in regions with cumuliform clouds (Seethala and Horváth, 2010). To ensure a high level of confidence, only data pixels of the cloud water path image were selected which possess 'very good confidence' or 'good confidence' from the Water Path Confidence Quality Assurance of MODIS. Afterwards, the reduced datasets of water vapour (Precipitable Water Infrared Clear) and cloud condensate (Cloud Water Path) of each image were added together to form the TWP under the condition that data pixels for both variables exist. In contrast to the model data, which is available globally and at hourly intervals, the MODIS data are recorded along the satellite swath at solar local time. Thus, these data had to be assigned to the model grid boxes by means of the latitude/longitude information and time (+/- 30 min, i.e. one model time step) of the data pixels. Additionally, only model grid boxes containing at least 100 MODIS data pixels were selected in order to obtain reliable results of variance and skewness (Figure 1). The ratio of remaining data of TWP related to the possible MODIS retrievals in each grid box after applying the above-mentioned constraints is shown in Figure 2. As expected, the number of data is rather small in regions characterized by large cloud cover, in particular over the southern oceans around 60°S. The statistical parameters of TWP (mean, variance and skewness) were then estimated from each grid box of the reduced dataset as shown exemplarily in Figure 3. Although Figure 3(a) depicts a unimodal PDF of TWP, we also found PDFs the forms of which were bimodal (Figure 3(b)) or multimodal (Figure 3(c)). Finally, the mean deviations of total cloud cover, mean TWP and its variance and skewness (model minus MODIS) were calculated.

3. Discussion of satellite results

3.1. Variance and skewness

In this subsection, the derived variance and skewness of TWP from the MODIS retrievals are briefly discussed. The high variance we find at low latitudes (Figure 4(c)) is likely caused by humidity-related processes, e.g. convection or precipitation. In particular, the high values of variance over the Sahel in Africa are probably produced by the occurrence of very dry and humid areas within the grid boxes, indicating convective activity. The very high variance found in regions with strong gradients in orography, however, is introduced by considering the vertical integral of the total water. In mountain regions, this integral will be much smaller than in the valleys, so that variance is high without meteorological reason. In general, the spatial pattern of variance over the
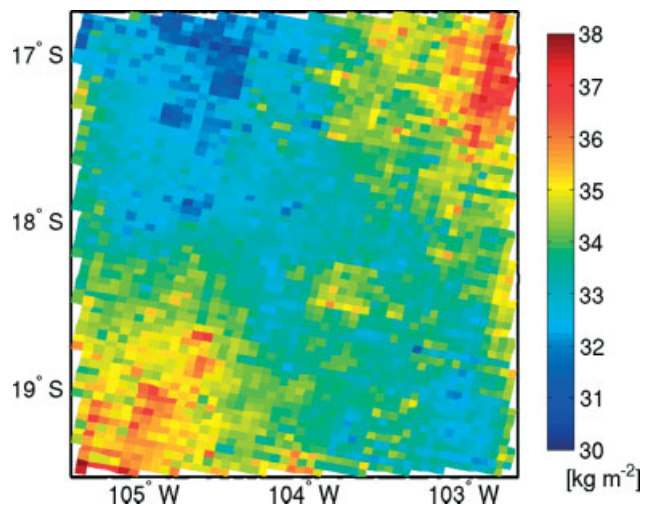

Figure 1. Spatial distribution of TWP in an example model grid box filled with MODIS data pixels.
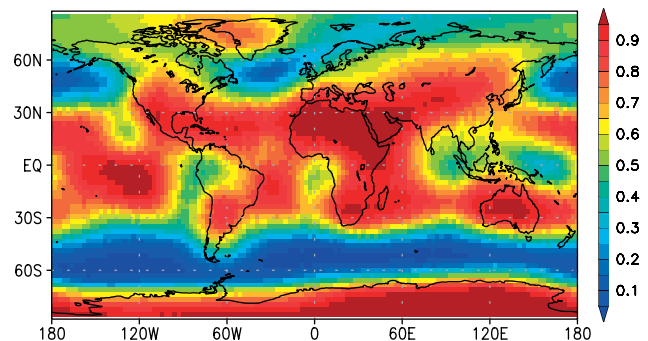

Figure 2. Ratio of reduced data of TWP related to the possible MODIS retrievals in each grid box.

globe indicates that the distribution of variance is a product of the natural variability of TWP rather than noise of the MODIS instruments.

The most interesting feature of the global mean distribution of skewness of TWP is the occurrence of negative values in the Tropics (Figure 4(d)). The distribution of TWP within a grid box is negatively skewed when the left tail of the distribution is longer than the right one (i.e. high values of TWP are more frequent than low ones). Low values of TWP can be attributed to convective downdraughts transporting dry air from the upper atmosphere down to lower levels and the PBL. The fact that negative skewness dominates over column mean total water can be explained by the method (vertical moisture integral) and the fact that water-vapour mixing ratio decreases approximately exponentially with height. Positive skewness is produced when low values of TWP are more frequent than higher ones in the domain. In this case, the high values that accompany clouds can be related to convective activity occurring more often over land than ocean.

3.2. Caveats

It is known that cloud detection from satellites over bright surfaces (sand, snow and ice) is flawed, since the retrieval algorithms use thresholds in reflectivity with which it is difficult to distinguish between clouds and these kinds of surfaces. In particular, MODIS often misses high thin cirrus
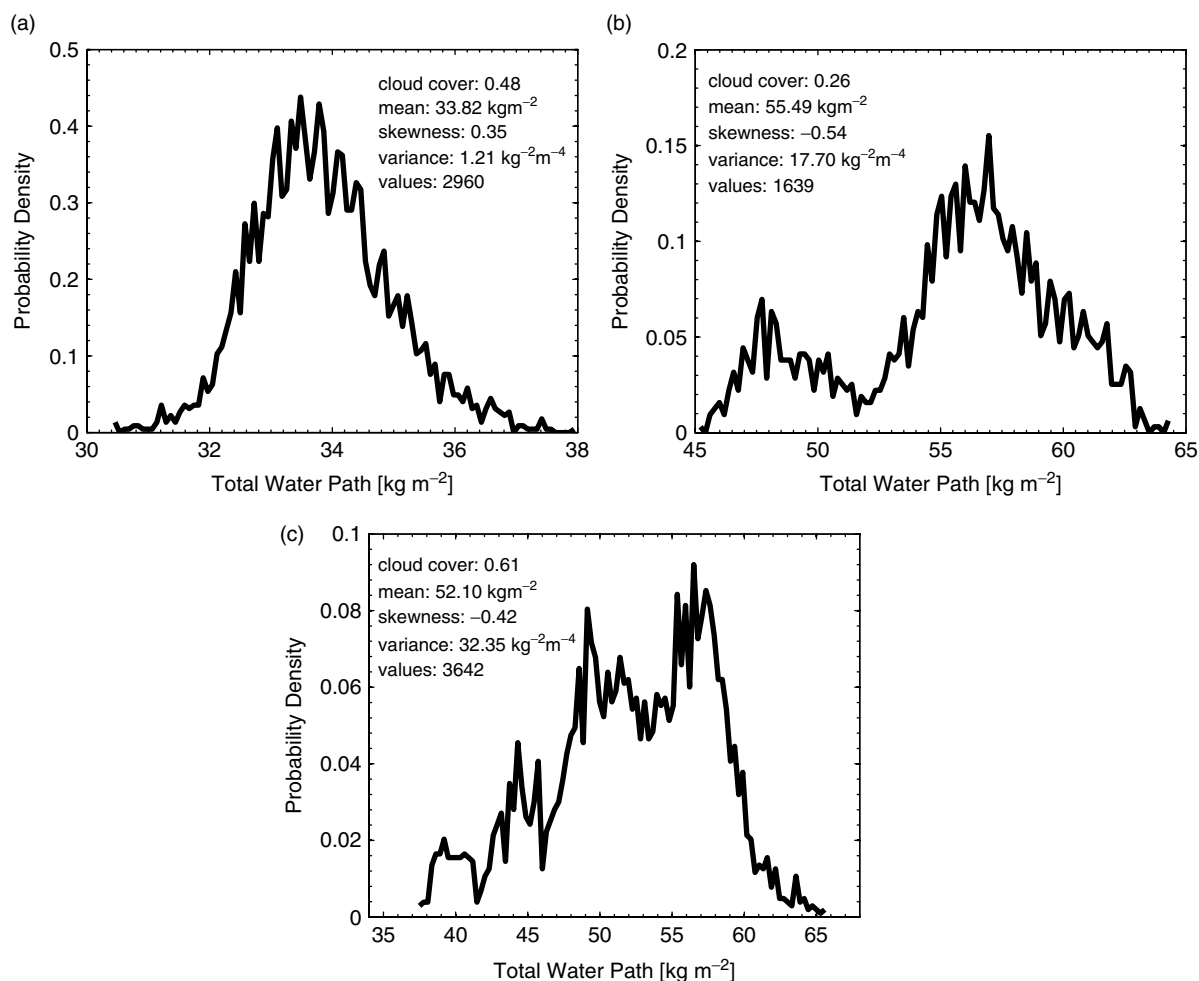
Figure 3. Observation-derived distributions of TWP from model grid boxes and corresponding statistics. (a) A unimodal distribution derived from the data shown in Figure 1, (b) a bimodal and (c) a multimodal distribution from other grid boxes as an example.

clouds over deserts (Holz *et al.*, 2008). Furthermore, the assignment of pixels of the cloud mask labelled 'probably clear' to clear sky and 'probably cloudy' to overcast sky leads to an overestimation of cloud cover over the ocean (Figure 4(a)) by 0.087 compared with that derived from the high-resolution Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER: Kotarba, 2010). The interpolation method we applied to fill the gaps in water vapour caused by overcast conditions could lead to an underestimation of water vapour below clouds (and consequently TWP; Figure 4(b)). This would reduce the variance of TWP in cloudy regions in our results and could be one reason for the lower variance over the subtropical oceans where low-level clouds are abundant, especially in stratocumulus regions on the western side of the continents. Therefore, it can be assumed that the variance over the oceans derived from the MODIS data as shown in Figure 4(c) may be slightly too low.

The high positive values of skewness over bright surfaces, e.g. Sahara and the polar regions, may not be trusted. In those regions, the water vapour is very low so that the cloud condensate being flawed there affects the TWP more. In general, the uncertainty in TWP is manly affected by the uncertainty in water vapour, since the TWP is dominated by water vapour due to its higher amount compared with the condensate. The intention in using MODIS measurements was to have a high spatial resolution on a global scale and a high number of retrievals, which should decrease the

uncertainty of the measurements. Furthermore, one has to keep in mind that because of the measurement method of MODIS, the TWP and its derived statistical parameters mainly represent the lower atmosphere, as explained in this subsection. The discussion shows that the variables derived from MODIS retrievals cannot be regarded as the absolute truth on an individual basis, but may be used for statistical comparisons. For the evaluation of the distributions of the statistical moments of TWP it seems sufficiently reliable.

4. Evaluation results

4.1. Model versus observation

To assess the agreement between model results and observations, the relative mean error and the root-mean-square error (RMS) of the deviations of total cloud cover, TWP and its variance and skewness were calculated. Furthermore, the spatial deviations of these parameters are displayed in the resolution $T42L19$ at a 0.95 significance level using a two-sided t-test (Figure 4(e)–(h)). For comparison the mean values of analyzed quantities derived by MODIS are shown in Figure 4(a)–(d).

In order to allow for a broad impression of the parametrization, we first compare the total cloud cover retrieved by MODIS with the simulated one. An in-depth
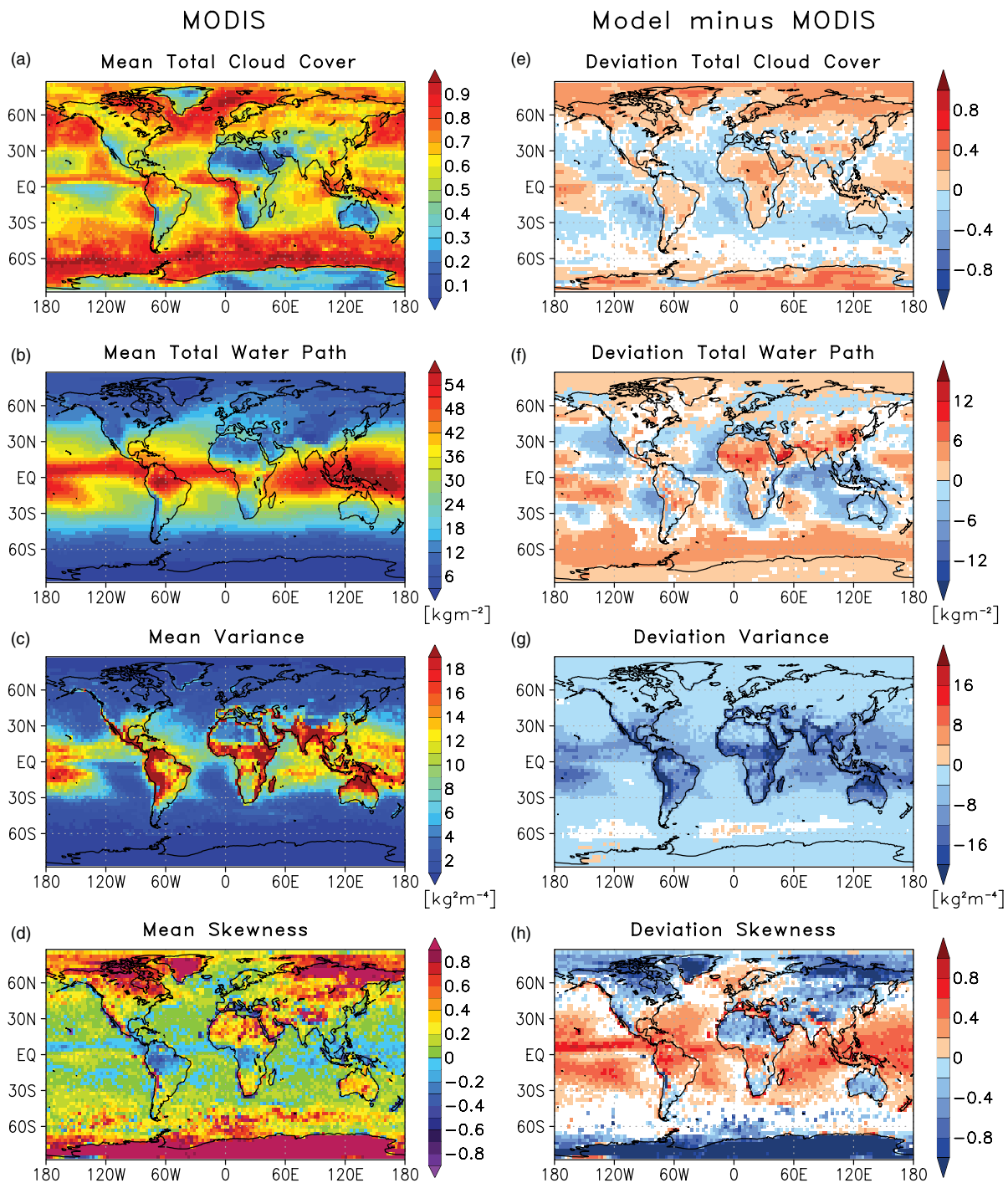
Figure 4. The left column depicts the mean values derived from MODIS in 2004 and the right column shows the deviation of modelled total cloud cover, TWP and its variance and skewness (*T42L19*) from MODIS data (white areas mask out differences not significant at a 0.95 statistical significance level).

evaluation of this quantity, though, is not the aim of this study. Therefore we did not apply a MODIS-simulator, which would be necessary for a full quantitative comparison. The modelled total cloud cover, with a relative mean error of 3.2% and an RMS of 0.14, differs from the MODIS data having a mean cloud cover of 0.63. Figure 4(e) shows an underestimation of total cloud cover by the model, mainly over the oceans, in particular off the western coasts of the continents, and an overestimation of clouds in high latitudes, over Africa and South America as well as over the central Pacific. Negative deviations on the western sides of the continents, where stratocumulus clouds usually prevail, indicate that the model has a deficiency in simulating

this type of shallow marine boundary-layer cloud. Positive deviations over North Africa, Antarctica and the North Pole region, on the other hand, might be caused by a systematic underestimation of clouds in the MODIS data rather than model errors, as explained in section 3.2.

A marginally better agreement with the observational data is found in the TWP, with a relative mean error of 1.2% referred to the MODIS mean of 26.55 kg m^{-2} and an RMS of 3.37 kg m^{-2}. In detail, the deviation pattern at low latitudes is similar to that of the total cloud cover. The modelled TWP is mostly too low over the oceans on the western sides of the continents as well as over Indonesia, Australia and the northern high latitudes, and too high over the central

Pacific, Africa, Asia and around 60°S over the southern oceans (Figure 4(f)). The results north of 60°N and south of 60°S as well as over North Africa and Arabia are not reliable, because of the aforementioned uncertainties in the MODIS data. However, evaluation of the mean values of TWP has been carried out earlier (e.g. Chen *et al.*, 1996) and is not of central interest to this study. It may be sufficient to note that despite several shortcomings, the model is broadly consistent with the data for these two quantities.

The third analyzed quantity, the variance, is almost globally underestimated by the model, which can be seen in the large relative mean error of -85.1% referred to the MODIS mean of $7.56\,\mathrm{kg}^2\,\mathrm{m}^{-4}$ and the RMS of $10.27\,\mathrm{kg}^2\,\mathrm{m}^{-4}$. Particularly, large differences from the observations are found at low latitudes and over the continents, potentially indicating deficiencies in the turbulence parametrization (Figure 4(g)). However, higher deviations towards the Equator suggest that, additionally, a humidity-related process is not well described or does not work properly in the cloud-cover scheme. One possible explanation could be that the precipitation process removes too much liquid water out of the clouds, resulting in too low a variance. However, it has to be mentioned that because of the coupling of variance (distribution width) and skewness (shape-parameter q) through Eq. (A9), the flawed skewness also contributes to the errors of the variance (for more details see section 4.2). Furthermore, an analysis of the frequency of cloud cover indicates that the prognostic equation for variance (distribution width) (Eq. (A9)) is applied in 90.1% of cases (clear sky or overcast) and diagnosed in only 9.9% of cases (partly covered). This suggests that the prognostic equation for variance (distribution width) contributes more to the bias than diagnostic closure of the equation system.

Deviations of the skewness of TWP are differently distributed geographically from the quantities analyzed before. The model overestimates the skewness in the Tropics and at low latitudes over the oceans. It underestimates the skewness over North America, North and South Africa, continental Asia, Australia and the polar regions. Despite the positive deviations of skewness in the Tropics shown in Figure 4(h), the mean deviation is dominated by negative values, which can be seen in the relative mean error of -40% compared with the MODIS mean of 0.20. Because of the fact that the positive deviations are overcompensated by negative ones, it is more reasonable to rely on the RMS of 0.46 or try to separate the regions with different sign of deviation. However, if the analysis is confined to the region between 30°S and 30°N, the RMS decreases to 0.31 and the relative mean error becomes positive, with 200% referred to the MODIS mean of 0.05. One reason for these positive errors is certainly that only positive skewness is allowed in the model, whereas the data derived from MODIS exhibits negative skewness, especially in the Tropics (Figure 4(d)). Negative skewness can be caused by strong convective downdraughts, which transport dry air from upper layers to lower ones or into the PBL. Neglecting these downdraughts, as is done in the prognostic equation for skewness in the model, may also contribute to the above-mentioned errors. In general, the discrepancy between skewness and observational data is likely to be caused by shortcomings in the cloud-cover scheme's link to the convection parametrization of the cloud-cover scheme. Moreover, the large-scale cloud-cover scheme by Sundqvist *et al.* (1989), based on a relative humidity threshold without a prognostic PDF, was also

Table I. Description of the sensitivity experiments.

Experiment	Description
control	standard model setup
vdiff	disabled vertical diffusion of distribution width of TWP
hedd	reduced dissipation of distribution width caused by horizontal eddies by the factor of 10
vedd	reduced dissipation of distribution width caused by vertical eddies by the factor of 10
rsk	calculating skewness of TWP only for deep convection
pvar	shape-parameter p is set to vary as $p = (q + 1)/(q - 1)$

evaluated and compared with the PDF scheme in terms of the total cloud cover and the TWP. It was found that the relative mean errors of cloud cover are the same and the RMSs only slightly differ by 0.01. The disagreement with observations of TWP is slightly higher in both the relative mean errors (2.6%) and the RMS ($3.46\,\mathrm{kg\,m}^{-2}$) compared with the PDF scheme.

4.2. Sensitivity experiments

Sensitivity experiments were carried out to identify specific parts of the parametrization that may be responsible for some of the deviations of modelled variance and skewness of TWP from the observations and potentially to allow for suggestions for improvements by modifying some choices within the parametrization. In these experiments, diabatic and dynamical processes increasing or decreasing variance and skewness, as parametrized by Tompkins (2002), were switched off or modified (an overview of the experiments can be seen in Table I). One of the sensitivity experiments also changed the parametrization at a slightly more fundamental level to allow for negative skewness in the model. Each sensitivity run encompasses a simulation of the first three months of 2004 (starting three months earlier to reach an equilibrium for cloud microphysics) and is compared with MODIS measurements for the same time period. For these sensitivity experiments, an analysis of just three months may be considered sufficient to identify impacts of the modifications on simulated variance and skewness of TWP.

In the first experiment (labelled *vdiff*) the vertical diffusion of the distribution width was switched off to analyze the impact of this mode of transport on the simulated TWP variance. This process, which transports the distribution width vertically, is described in the third term of Eq. (A9). The result shows that its impact on the variance is small, which can be seen in the RMS of $10.83\,\mathrm{kg}^2\,\mathrm{m}^{-4}$ compared with the RMS of $10.77\,\mathrm{kg}^2\,\mathrm{m}^{-4}$ for the control run (Table II). Consequently, the spatial pattern of this parameter shows only slight changes in this run (not shown). No changes are observed in the RMS of skewness (Table II) and its spatial pattern is almost identical to the control run, except for the northern hemisphere where the artificially negative skewness increases slightly (not shown).

Table II. High-order moments of the statistical PDF scheme in different sensitivity experiments (model minus observation).

Experiment	Variance (kg^2 m^{-4})			Skewness			Skewness (30°S–30°N)		
	Bias	RMS	Rel. Err.	Bias	RMS	Rel. Err.	Bias	RMS	Rel. Err.
control	−6.11	10.77	−84.0%	−0.18	0.71	−62.1%	0.10	0.37	166.7%
vdiff	−6.19	10.83	−85.1%	−0.18	0.71	−62.1%	0.10	0.37	166.7%
hedd	−6.11	10.75	−84.0%	−0.18	0.71	−62.1%	0.11	0.38	183.3%
vedd	−4.79	9.75	−65.9%	−0.15	0.70	−51.7%	0.12	0.37	200.0%
rsk	−5.44	10.15	−74.8%	−0.29	0.69	−100.0%	−0.07	0.31	−116.7%
pvar	−5.70	10.45	−78.4%	−0.39	0.80	−134.5%	−0.11	0.39	−183.3%

The next two experiments analyze the effect of dissipation caused by horizontal and vertical eddies on the distribution width (last term of Eq. (A9)), reducing the influence of these processes by a factor of 10. This high factor was chosen to amplify the effect of the modification on these processes. In the first run (*hedd*), where the dissipation caused by horizontal eddies is mitigated, the deviation of the modelled variance from the observations is somewhat lower, with an RMS of 10.75 kg^2 m^{-4}, caused by an increase of variance in the Tropics. The spatial pattern of skewness is quite similar to the one in the aforementioned experiment (*vdiff*), with a somewhat higher skewness in the Tropics. Nevertheless, its RMS of 0.71 is almost equal to that of the control run. In the second run (*vedd*), in which the dissipation caused by vertical eddies is reduced, the RMS of the variance decreases by 9.5% to 9.75 kg^2 m^{-4}, induced by an increase of variance mainly over the oceans and in the Tropics (Figure 5(b)). The modification of the dissipation also leads to an enhancement of skewness over the extratropical oceans (Figure 5(f)), resulting in a marginally lower RMS of 0.70. This can be explained by the coupling of variance and skewness in Eq. (A6), in which the minimum a and the maximum b (distribution width, $b − a$) are used to estimate the impact of the change of cloud condensate on skewness (shape-parameter q). Altogether, as a result of the analysis of the processes affecting the variance directly, the dissipation of distribution width caused by vertical eddies has the largest, albeit still small, impact on variance.

In the next two experiments the scheme was modified to analyze the causes for the too-large positive skewness, especially in the Tropics. To achieve this, in the first run (*rsk*) only the detrainment by high-reaching convection (determined as convective clouds consisting of ice rather than liquid water) was allowed to increase the skewness of total water mixing ratio (first term of Eq. (A5)). Convection is the dominant source for positive skewness in the scheme. In the lower troposphere, convective regions may exhibit negative skewness, though, where dry downdraughts are effective. This sensitivity study crudely assumes that this generation of negative skewness by downdraughts and the generation of positive skewness by detrainment from updraughts roughly cancel for liquid-water clouds. This result shows a strong reduction of skewness at low latitudes, where low negative skewness now prevails (Figure 5(g)), highlighting that indeed convection is the main source of positive skewness in this parametrization. In the sensitivity study, simulated skewness compares better with the observations in the Tropics by exhibiting now only small values of skewness for the vertically integrated total water, instead of the strongly positive skewness in the control simulation.

Such significant changes in the spatial pattern of skewness, however, have only consequences for the RMS, which decreases only slightly by 2.8% from 0.71 to 0.69, whereas the relative mean error increases from −62.1% to −100.0%. This is explained by disappearance of high skewness values at low latitudes and thus less cancellation of errors. It becomes even more obvious if the area for calculation of deviation errors is confined to 30°S to 30°N. Then the relative mean error decreases from 166.7% of the control run to −116.7% in the modified run and the RMS from 0.37 to 0.31, which is a decline of 16.2%. The variance in this experiment shows an RMS of 10.15 kg^2 m^{-4}, somewhat lower than in the control experiment. From this sensitivity experiment we conclude that it is recommendable to allow for a source of negative skewness by convective downdraughts in order to match better the observed pattern of TWP skewness.

Although the current version of the cloud-cover scheme is not able to produce negative skewness of the total water mixing ratio, some slightly negative values for the skewness of total water path are derived when integrating vertically using the overlap assumptions within the subcolumn-generator. However, to model a distribution of negative skewness comparable to the MODIS measurements (as shown in Figure 4(d)), it is necessary to allow negative skewness in the scheme. In the general case, this would require an additional prognostic equation for the shape-parameter p. In order to avoid this, Tompkins (2008) suggests letting the shape-parameter p be tied to q and vary as $p = (q + 1)/(q − 1)$, which allows for relations of p and q leading to negative skewness without the need for an additional prognostic equation. The result of this modified parametrization (*pvar*) causes a decline of skewness to negative values in regions where low positive values already exist, but the spatial pattern of skewness is maintained (Figure 5(h)), since no source term for negative skewness exists. Accordingly, the RMS of skewness increases to 0.80. A better result is obtained for the RMS of the variance of 10.45 kg^2 m^{-4} induced by an increase of variance at low latitudes (Figure 5(d)). As shown in this experiment, negative skewness can in principle be modelled by a simple modification of the scheme. However, the skewness is only reduced and the high values are still simulated in the Tropics, which implies additionally that a modification of the influence of convection on skewness also has to be implemented. The RMS of the total cloud cover is almost equal to that of the control run in the afore-analyzed experiments (Table III), except for experiment *pvar* with a slightly higher RMS of 0.18. The RMS of TWP ranges between 4.41 kg m^{-2} in experiment *vdiff* and 5.07 kg m^{-2} in experiment *pvar*.
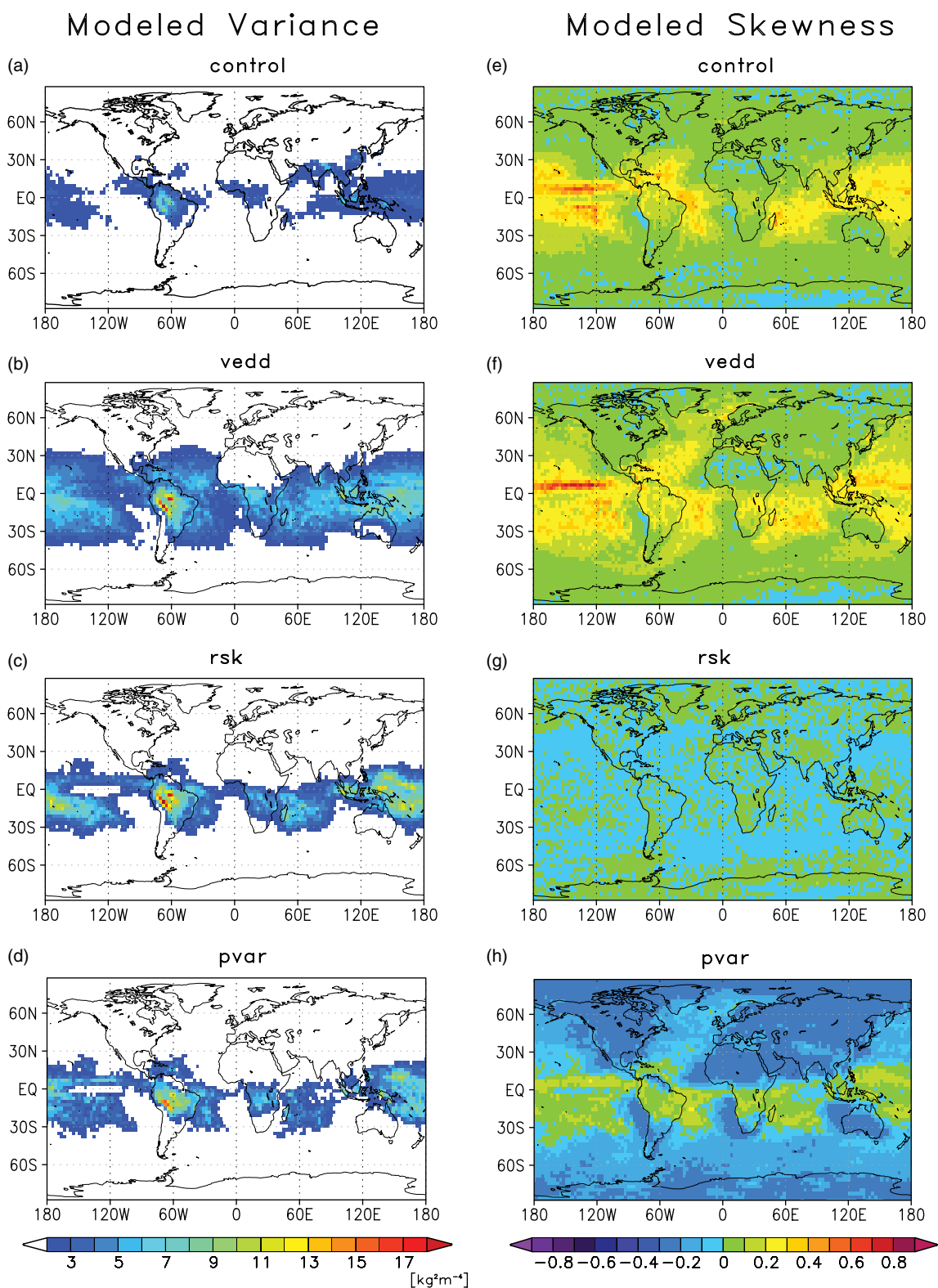
Figure 5. The left column shows the modelled variance and the right column the modelled skewness of simulations with different parametrizations for resolution *T*42*L*19 (mean of three months during 2004). The abbreviations of the experiments are explained in Table I.

5.   Summary and conclusions

The subgrid-scale variability scheme for water vapour and cloud condensate applied in the ECHAM5 climate model was evaluated globally with high-resolution satellite data. For this purpose, a model experiment with a horizontal resolution of $2.8° \times 2.8°$ was carried out. To make the three-dimensional model data comparable to the two-dimensional MODIS data, the subcolumn-generator developed by Räisänen *et al.* (2004) was used to integrate vertically the horizontal subgrid-scale distributions of water vapour, cloud liquid water and cloud ice mixing ratios. Model-simulated total cloud cover, TWP, its variance and skewness were compared with parameters derived from MODIS and deviations were calculated, where statistically significant at a 95% significance level.

Table III. First-order moments of the statistical PDF scheme in different sensitivity experiments (model minus observation).

Experiment	Total Cloud Cover			TWP ($kg\,m^{-2}$)		
	Bias	RMS	Rel. Err.	Bias	RMS	Rel. Err.
control	0.01	0.17	1.6%	0.07	4.64	0.3%
vdiff	0.02	0.17	3.1%	−0.02	4.41	−0.1%
hedd	0.01	0.17	1.6%	−0.01	4.48	−0.04%
vedd	0.00	0.17	0.0%	−0.44	4.56	−1.7%
rsk	0.02	0.17	3.1%	−0.01	4.73	−0.04%
pvar	−0.02	0.18	−3.1%	−0.32	5.07	−1.3%

The results show that the mean total cloud cover and mean TWP are – despite several noticeable deficiencies, which are not the focus of the present study – on average relatively well simulated.

Large deficiencies in both variance and skewness of the PDF were revealed through the evaluation. Systematically negative deviations of variance are found for almost all regions of the globe. Skewness of the TWP is overestimated in the Tropics and at lower latitudes over the oceans, and underestimated at high latitudes. The former is caused mainly by the model condition that allows for only positive skewness compared with the satellite data, which shows negative skewness in the Tropics in particular.

Sensitivity experiments were made to identify the processes that cause the respective deviations of modelled variance and skewness of TWP from the observations. These experiments show that the variance (distribution width) is increased by reducing the dissipation caused by vertical eddies, which has a stronger impact than the dissipation caused by horizontal eddies or the vertical diffusion of this quantity. This suggests that the scheme probably overestimates the vertical turbulence that dissipates the distribution width in the PBL and in the environment of deep convection cores. However, it should be noticed that the vertical gradient of total water mixing ratio being coupled with the turbulent moisture flux also affects the variance (distribution width, see Eq. (A9)). One process that possibly also contributes to the negative bias of variance is the precipitation process removing too much liquid water from clouds within one time step. Another point that should also be considered in future investigations is the underlying surface heterogeneity in the model, which may be too low. Additionally, the flawed skewness (shape-parameter q) coupled with the variance (distribution width) through its prognostic equation also contributes to the bias of modelled variance. However, the flawed variance also affects the skewness, as demonstrated in the experiment in which the reduced dissipation of distribution width caused by vertical eddies increases the skewness.

The positive bias in skewness in the Tropics is reduced considerably when detrainment only for deep (ice-forming) convection is used as a source of positive skewness. This sensitivity study suggests that in the boundary layer, convective downdraughts should also be considered as a source for negative skewness. The test confirmed that the contribution of skewness caused by cloud liquid water is one main source for the positive bias in skewness in the Tropics. In order to allow for negative skewness in the model, a relation between the two shape parameters of the PDF suggested by Tompkins (2008) was implemented. With this modified parametrization, the model was better able to produce mean negative skewness, mainly in high and mid-latitudes, but again positive skewness in the Tropics.

The reference satellite data contain some uncertainties. Most importantly, the interpolation of missing water-vapour data below clouds, a result of the measurement method of MODIS, might bias the evaluation in regions with large average cloud cover. In some regions, such as high latitudes, the satellite data are not trustworthy. Also, the vertical integration of modelled TWP making the data comparable to the satellite data introduces biases into evaluation. This integration strongly depends on the overlap assumption. Assuming maximum overlap results in larger variance and skewness for the vertical integral of total water mixing ratio compared with random or maximum-random overlap assumptions – a potential overestimation of the distribution moments. However we find that, even for this maximum overlap assumption, variance is strongly underestimated by the model. For skewness, the absolute value might be overestimated due to the overlap assumption but the qualitative results are not affected. Overall, given the strong conclusions regarding the deficiencies of the current parametrization, issues of data quality and methodological assumptions are of minor importance.

An important caveat on this evaluation is the vertical integration of TWP, which emphasizes the lower troposphere, in particular the PBL, but does not allow us to investigate the free and upper troposphere specifically. However, the vertical integral comprises the only data currently available for an evaluation of the PDF scheme at a global scale, so this study should be seen as a first step in the evaluation. It should be noted that simulation results and observational data are treated consistently. This evaluation shows that there is still an urgent need for further improvements of the subgrid-scale variability scheme for water vapour and cloud condensate. Once the overall performance of the parametrization is improved, it is important to dedicate a study to the variability of just the condensed water, which is of particular interest for the radiation. Several interesting features of the spatial distribution of TWP variance and skewness can be seen in the satellite data, which merit specific investigation.

In the current scheme, vertical downdraughts, which transport dry air from the upper atmosphere in the PBL and produce negative skewness, are not taken into account in the prognostic equation of skewness. However, negative skewness in the lower atmosphere can be seen in the satellite data and negative skewed distributions of total water mixing ratio in stratocumulus clouds have also been found by Zhu and Zuidema (2009), who analyzed various cloud

cases with large eddy simulations (LESs). Moreover, they revealed that the variance of moisture and temperature are equally important in controlling subgrid-scale clouds. Tompkins (2003) analyzed extensive aircraft data and found that the effect of temperature fluctuations for cloud cover is roughly half that of the humidity fluctuations. Thus, the inclusion of subgrid-scale variability of temperature, which is currently neglected, could also improve the scheme. The next logical step would be to allow for negative skewness in the parametrization of the scheme and to exploit the information regarding subgrid-scale variability of cloud condensate provided by the scheme in precipitation formation, such as autoconversion or the accretion process.

Acknowledgements

Appendix. Subgrid-scale variability scheme in ECHAM5

The ECHAM5 general circulation model of the atmosphere (Roeckner *et al.*, 2003) employs a statistical-dynamical approach to account for the subgrid-scale variability of water vapour and cloud condensate, developed by Tompkins (2002). It uses the prognostic equations of the water phases (vapour, liquid and solid) and the bulk cloud microphysics described by Lohmann and Roeckner (1996). The subgrid-scale variability scheme neglects the temperature fluctuation within a model grid box, assuming the subgrid-scale formation of clouds is caused only by fluctuations in the total water mixing ratio. Then, the horizontal cloud cover can be calculated by

$$C = \int_{r_s}^{\infty} G(r_t) \, dr_t, \qquad (A1)$$

where saturation vapor mixing ratio r_s is assumed to be constant within the considered grid box and $G(r_t)$ represents a PDF of the total water mixing ratio r_t, the sum of water vapour, cloud liquid water and cloud ice. In this scheme the PDF of the beta function is used to describe the distribution of the total water mixing ratio, which gives for the horizontal cloud cover

$$C = 1 - I_x(p, q). \qquad (A2)$$

The right-hand term of (A2) is the incomplete beta function ratio I_x with $x = (r_s - a)/(b - a)$:

$$I_x(p, q) = \frac{1}{B(p, q)} \int_0^x r_t^{p-1} (1 - r_t)^{q-1} \, dr_t, \qquad (A3)$$

where a and b are the minimum and maximum of total water mixing ratio, $p > 0$ and $q > 0$ the shape parameter of the beta function and B is the beta function:

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}, \qquad (A4)$$

with Γ the gamma function. Atmospheric processes such as turbulence and microphysics are taken into account through the parameters p and q, which determine the shape of the distribution of total water mixing ratio. Due to simplifications in the current version of the subgrid-scale variability scheme, p is set constant equal to 2 and q is allowed to assume values between 2 and 50. As a result of this relation of p and q, the distribution of total water mixing ratio can only be bell-shaped ($p = q$) or positively skewed ($p < q$). In the model, the temporal evolution of the shape parameter q is calculated by the following prognostic equation:

$$\frac{Dq}{Dt} = \frac{K}{\overline{\rho} r_s} \frac{\partial}{\partial z} (M^{cu} r_c^{cu}) + \frac{\Delta q^{micro}}{\Delta t}$$
$$+ (q_0 - q) \left(\frac{1}{\tau_v} + \frac{1}{\tau_h} \right). \qquad (A5)$$

The first term of Eq. (A5) is related to convection and describes the increase in skewness through detrainment of cloud condensate, where K is a dimensionless constant and specifies how quickly detraining water from convection increases the skewness, $\overline{\rho}$ is the air density, M^{cu} is the updraft mass flux and r_c^{cu} represents the mean cloud water in the convective updrafts. The present scheme neglects the impact of convective downdrafts and vertical transport of skewness by vertical turbulence. Microphysical processes in clouds are considered in the second term. A reduction of the cloud condensate $\Delta \overline{r}_c^{micro}$ caused by precipitation or evaporation, for instance, over one time step reduces the maximum of total water mixing ratio b and accordingly the shape parameter Δq^{micro} and shifts the PDF to a symmetric shape. Both variables are coupled through the equation of the mean of total water mixing ratio, which is expressed as

$$\Delta q^{micro} = \frac{\left[b + \frac{\Delta \overline{r}_c^{micro}}{\overline{r}_c}(b - r_s) - a \right] p}{\overline{r}_t - a} - (p + q). \qquad (A6)$$

Turbulence is taken into account in the third term using Newtonian relaxation, which drives the distribution in the presence of mixing toward a symmetric form. The dissipation time-scale divided into a horizontal component τ_h and vertical component τ_v is written as

$$\tau_v = \frac{\kappa l}{\sqrt{e}} \qquad \text{with } l = \frac{kz}{1 + kz/\lambda}, \qquad (A7)$$

$$\tau_h^{-1} = C_s^2 \sqrt{\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2}. \qquad (A8)$$

Eq. (A7) describes the dissipation caused by 3D turbulence in the PBL and in the vicinity of deep convection cores, where κ is a constant, e the turbulent kinetic energy, k von Kármán's constant, z the height and λ the asymptotic mixing length. Eq. (A8) represents the dissipation through large-scale 2D horizontal eddies due to horizontal wind-shear instability with the horizontal wind components u and v, and C_s^2 is a constant. In clear sky or in overcast conditions, the system is not closed and a quasi-prognostic equation is applied that estimates the distribution width $(b-a)$ of total water mixing ratio as proportional to its variance:

$$\frac{D(b-a)}{Dt} = \frac{(\overline{r}_t - a)(p+q+\Delta q_{conv})p^{-1} - (b-a)}{\Delta t}$$
$$- \frac{\eta}{(b-a)}\overline{w'r_t'}\frac{\partial \overline{r}_t}{\partial z} - \Lambda\sqrt{e}\frac{\partial(b-a)}{\partial z}$$
$$- (b-a)\left(\frac{1}{\tau_v} + \frac{1}{\tau_h}\right). \qquad (A9)$$

The first term of Eq. (A9) takes into account the effect of deep convection detrainment on the distribution width, where Δq_{conv} is equal to the first term of Eq. (A5). The second term of Eq. (A9) estimates the production of distribution width in the presence of a vertical moisture gradient with $\eta = (p+q)^2(p+q+1)(pq)^{-1}$ and $\overline{w'r_t'}$ as the turbulent moisture flux including the vertical wind w. Vertical transport of distribution width through subgrid eddies is described in the next term, where $\Lambda = lS_r$ and S_r is a stability function, and the last term describes the impact of dissipation.

References

Bechtold P, Cuijpers JWM, Mascart P, Trouilhet P. 1995. Modelling of trade wind cumuli with a low-order turbulence model: Toward a unified description of Cu and Sc clouds in meteorological models. *J. Atmos. Sci.* **52**: 455–463.

Bony S, Emanuel KA. 2001. A parametrization of the cloudiness associated with cumulus convection; Evaluation using TOGA COARE data. *J. Atmos. Sci.* **58**: 3158–3183.

Bougeault P. 1981. Modeling the trade-wind cumulus boundary layer. Part II: A high-order one-dimensional model. *J. Atmos. Sci.* **38**: 2429–2439.

Chen CT, Roeckner E, Soden BJ. 1996. A comparison of satellite observations and model simulations of column-integrated moisture and upper-tropospheric humidity. *J. Climate* **9**: 1561–1585.

Davis A, Marshak A, Wiscombe W, Cahalan R. 1996. Scale invariance of liquid water distributions in marine stratocumulus. Part I: Spectral properties and stationary issues. *J. Atmos. Sci.* **53**: 1538–1558.

Gates WL, Boyle JS, Covey C, Dease CG, Doutriaux CM, Drach RS, Fiorino M, Gleckler PJ, Hnilo JJ, Marlais SM, Phillips TJ, Potter GL, Santer BD, Sperber KR, Taylor KE, Williams DN. 1999. An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Am. Meteorol. Soc.* **80**: 29–55.

Golaz JC, Larson VE, Cotton WR. 2002. A PDF-based model for boundary-layer clouds. Part I: Method and model description. *J. Atmos. Sci.* **59**: 3540–3551.

Holz RE, Ackermann SA, Nagle FW, Frey R, Dutcher S, Kuehn RE, Vaughan M, Baum B. 2008. Global moderate resolution imaging spectroradiometer (MODIS) cloud detection and height evaluation using (CALIOP). *J. Geophys. Res.* **113**: 1–17. DOI:10.1029/2008JD009837.

King MD, Tsay SC, Platnick SE, Wang M, Liou KN. 1997. 'Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius and thermodynamic phase', MODIS Algorithm Theoretical Basis Document No. ATBD-MOD-05: 1–79. http://modis-atmos.gsfc.nasa.gov

King MD, Menzel WP, Kaufman YJ, Tanré D, Gao BC, Platnick S, Ackerman SA, Remer LA, Pincus R, Hubanks PA. 2003. Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapour from MODIS. *IEEE Transactions on Geoscience and Remote Sensing* **41**: 442–458.

Kotarba A. 2010. Estimation of fractional cloud cover for Moderate Resolution Imaging Spectroradiometer/Terra cloud mask classes with high-resolution over ocean ASTER observations. *J. Geophys. Res.* **115**: 1–11. DOI:10.1029/2009JD013520.

Kuwano-Yoshida A, Enomoto T, Ohfuchi W. 2010. An improved PDF cloud scheme for climate simulations. *Q. J. R. Meteorol. Soc.* **136**: 1583–1597.

LeTreut H, Li ZX. 1991. Sensitivity of an atmospheric general circulation model to prescribed SST changes: Feedback effects associated with the simulation of cloud optical properties. *Clim. Dyn.* **5**: 175–187.

Lohmann U, Roeckner E. 1996. Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model. *Clim. Dyn.* **12**: 557–572.

Mellor GL. 1977. The Gaussian cloud relations. *J. Atmos. Sci.* **34**: 356–358.

Nishizawa K. 2000. Parametrization of nonconvective condensation for low-resolution climate models: Comparison of diagnostic schemes for fractional cloud cover and cloud water content. *J. Meteorol. Soc. Jpn* **78**: 1–12.

Pincus R, Klein SA. 2000. Unresolved spatial variability and microphysical process rates in large-scale models. *J. Geophys. Res.* **105**: 27059–27065.

Pincus R, Hemler R, Klein SA. 2006. Using stochastically generated subcolumns to represent cloud structure in a large-scale model. *Mon. Weather Rev.* **134**: 3644–3656.

Platnick S, King MD, Ackerman SA, Menzel WP, Baum BA, Riédi JC, Frey RA. 2003. The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing* **41**: 459–473.

Price JD. 2001. A study of probability distributions of boundary-layer humidity and associated errors in parametrized cloud-fraction. *Q. J. R. Meteorol. Soc.* **127**: 739–758.

Räisänen P, Barker HW, Khairoutdinov MF, Li J, Randall DA. 2004. Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Q. J. R. Meteorol. Soc.* **130**: 2047–2067. DOI:10.1256/qj.03.99.

Roeckner E, Bäuml G, Bonaventura L, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblueh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A. 2003. '*The atmospheric general circulation model ECHAM5 part I: Model description*', Max Planck Institute for Meteorology, Report No. 349. Available from MPI for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany.

Rotstayn LD. 1997. A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes. *Q. J. R. Meteorol. Soc.* **123**: 1227–1282.

Seemann SW, Borbas EE, Li J, Menzel WP, Gumley LE. 2006. *MODIS atmospheric profile retrieval algorithm theoretical basis document*, Version 6: 1–37. Available at http://modis-atmos.gsfc.nasa.gov

Seethala C, Horváth A. 2010. Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds. *J. Geophys. Res.* **115**: 1–19. DOI:10.1029/2009JD012662.

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, Deardorff JW. 1977. Subgrid-scale condensation in models of nonprecipitating clouds. *J. Atmos. Sci.* **34**: 344–355.

Sundqvist H, Berge E, Kristjánsson JE. 1989. Condensation and cloud parametrization studies with a mesoscale numerical weather predicition model. *Mon. Weather Rev.* **117**: 1641–1657.

Tompkins AM. 2002. A prognostic parametrization for the subgrid-scale variability of water vapour and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.* **59**: 1917–1942.

Tompkins AM. 2003. Impact of temperature and humidity variability on cloud cover assessed using aircraft data. *Q. J. R. Meteorol. Soc.* **129**: 2151–2170. DOI:10.1256/qj.02.190.

Tompkins AM. 2008. 'Cloud Parametrization'. In *Proceedings of Workshop on Parametrization of Subgrid Physics Processes, Reading, UK, 1-4, September 2008*. ECMWF: Reading, UK. http://www.ecmwf.int/publications

Watanabe M, Emori S, Satoh M, Miura H. 2009. A PDF-based hybrid prognostic cloud scheme for general circulation models. *Clim. Dyn.* **33**: 795–816. DOI:10.1007/s00382-008-0489-0.

Wood R, Field PR. 2000. Relationships between total water, condensed water, and cloud fraction in stratiform clouds examined using aircraft data. *J. Atmos. Sci.* **57**: 1888–1905.

Zhu P, Zuidema P. 2009. On the use of PDF schemes to parametrize sub-grid clouds. *Geophys. Res. Lett.* **36**: 1–5. DOI:10.1029/2008GL036817.