6A.3

EVALUATION OF OPERATIONAL MODEL CLOUD REPRESENTATION USING ROUTINE RADAR/LIDAR MEASUREMENTS

Dominique Bouniol
GAME/CNRM, CNRS/Météo-France, Toulouse, France
Alain Protat
CETP, CNRS, Vélizy, Paris
Julien Delanoë
Meteorology Department, University of Reading, Reading, United Kingdom

1. INTRODUCTION

The Cloudnet project (Illingworth *et al* 2007) took place between 1st octobre 2002 and 30th September 2004. One aim of this project was to collect continuous observations of cloud profiles (using cloud radar and lidar) during the two years from three observational sites in Northern Europe (Cabauw, NL, Chilbolton, UK, Palaiseau, F) in order to evaluate the cloud representation of four operational weather forecast models (ECMWF, ARPEGE, RACMO, Met Office) which profiles were stored at the measurements points every hours.

In this paper the main results of the evaluation are presented but the extensive details may be found in Bouniol *et al.* (2007a and b). At a first step the ability of a given model to produce a cloud at the right place and at the right time (*i.e.* cloud occurrence) is evaluated. The variables of cloud parametrisations (diagnostic or prognostic depending of the models): cloud fraction and ice water content are then compared between model and observations when model and observations agree on the cloud occurrence. In Bouniol *et al.* (2007 a and b) the comparisons are analysed at the scale of the whole data set as well at the seasonal scale but this paper will only summarised the results for the whole date set as interested readers may go in full details in Bouniol *et al.* (2007 a and b). This statistical comparison may highlight potential bias in the model. As an extra point, even if this evaluation is related to four models versions, this paper aim is to propose a methodology that allows to evaluate the model parametrisation, tuning and evolution.

This paper in then organised in three main parts: dealing with the cloud occurrence evaluation, the cloud fraction and the IWC.

2. CLOUD OCCURRENCE

As mentioned in the introduction instruments are recording instantaneous profiles. It is then necessary to average the cloud characteristics from the instrumental resolution to the model resolution. To do so, the same strategy as previous authors (Mace *et al.* 1998 or Hogan *et al.* 2001) has been used: a temporal average to yield the equivalent of a two-dimensional slice through the three dimensional grid-box. Using the model wind speed as a function of height and the horizontal model grid-box size (39 km for ECMWF, 24 km for ARPEGE, 18 km for RACMO and 12 km for Met Office) the appropriate averaging time is calculated. The underlying hypothesis is that in this time the cloud structure observed is predominantly due to the advection of structure within the grid-box across the site, rather than evolution of the cloud during the period as well as this two dimensional slice is representative of what happens in a three dimensional grid-box.

The fraction of the box that is filled by cloud is then the so called cloud fraction and the amount of ice corresponds to the IWC at the scale of the model. When a cloud fraction is larger than 3% the grid box is then considered as cloudy. By the end we compute the frequency of cloud occurrence from model and observation as the ratio of cloudy hours to the total number of observational hours for a given level.

Figure 1 shows a comparison of the frequency of cloud occurrence derived from the observations in blue (taking into account the model resolution) and the model in green at the three site (Cabauw, Chilbolton and Palaiseau, from top to bottom) and for the four models (ECMWF, ARPEGE, RACMO and Met Office from left to right). The statistics for the ARPEGE model (and the corresponding statistics from observations) has been split in two separate time period because this model encountered major changes in its cloud scheme on the 14th April of 2003. Therefore the arpege1 period (light green for the model and light blue for the observations) corresponds to cloud profiles before this date, the arpege2 period (dark green for the model, dark blue for the observations).

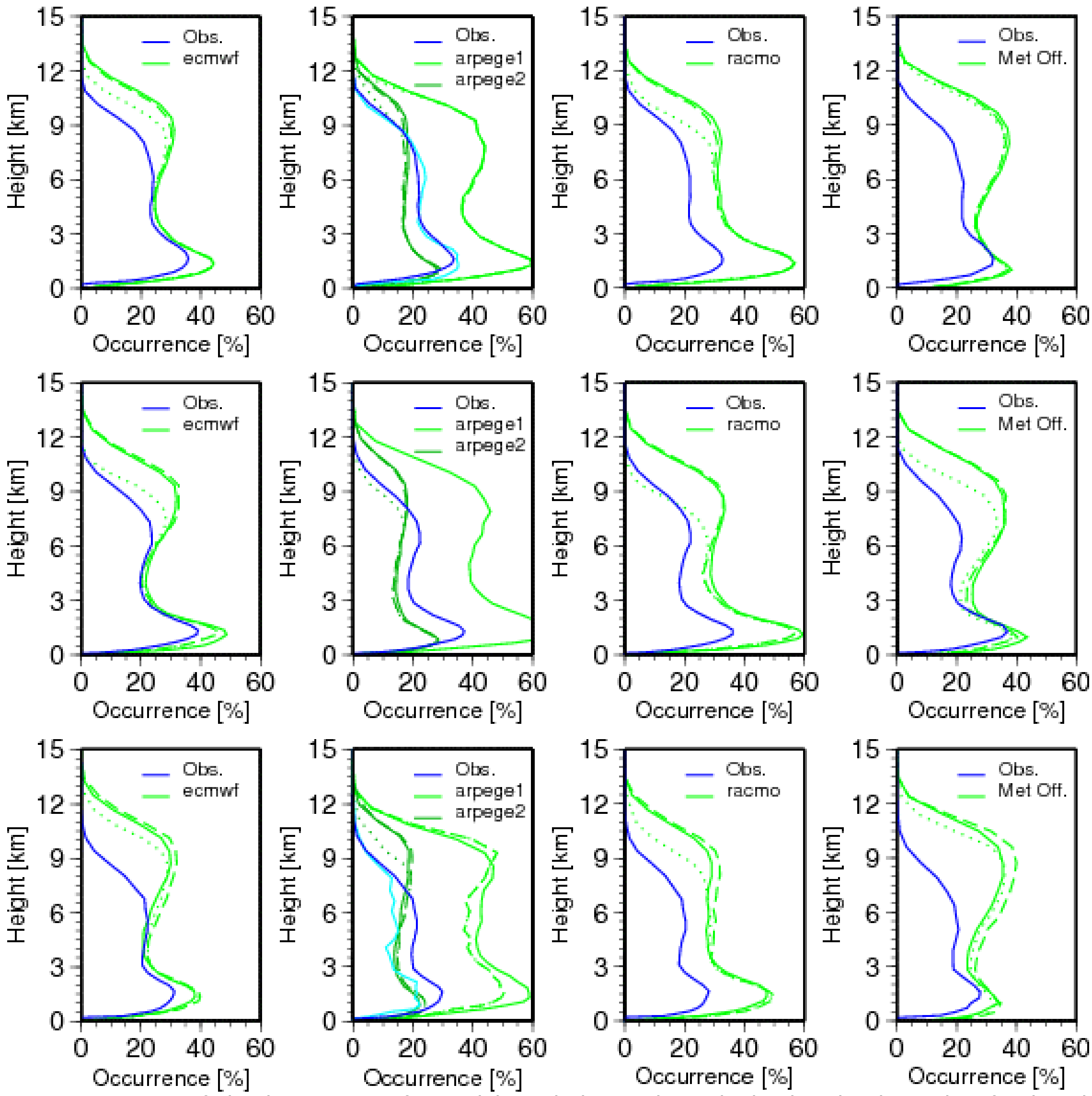

Figure1: Frequency of cloud occurrence for models and observations obtained at the three sites for the whole Cloudnet period (01~october~2002 to 30~september~2004). Blue line shows the frequency of cloud occurrence obtained from the observations. The green lines with different styles correspond to the model samples : solid line for whole model sample, dashed line for model sub-sample corresponding to instrument hours of operations and dotted line for model sub-sample corresponding to instrument hours of operations and including instrumental effects. Each line corresponds to an observatory from top to bottom: Cabauw, Chilbolton and Palaiseau. Each panel is dedicated to a comparison with a particular model : from left to right ECMWF, ARPEGE (1 and 2), RACMO and Met Office. For ARPEGE, the darkest line corresponds to arpege1 period, the lightest to arpege2.

Several curves, corresponding to the model sample (in green), are also superimposed in this figure. The solid line is the frequency of cloud occurrence for the whole model sample, the dashed line is the model sub-sample corresponding to instrument hours of operation and the dotted line is the model sub-sample corresponding to instrument hours of operations and including instrumental effects. The reason for superimposing these curves is that a instrumental network cannot operate strictly continuously, indeed depending of the radar or lidar type one needs time for maintenance, failure can occur on an instrument or a given instrument cannot operate in particular weather conditions. If not taken into account these under-sampling (dashed green line in Fig. 1) may lead to bias in the statistics, in particular if operation is linked to weather conditions as it is the case for the lidar at the Palaiseau site lidar as shown in Protat *et al.* (2006). Another effect that has to be accounted is the instrumental sensitivity. Indeed the all the radar of the Cloudnet project have not the same sensitivity and the Chilbolton and Palaiseau radars known a loss of power of their emission tube as demonstrated in Hogan *et al.* (2003), therefore their sensitivities do not remain constant during the whole project. These effects are taken into account for the statistics computation by computing for a given model the corresponding reflectivity using the empirical IWC-reflectivity relationships of Liu and Illingworth (2000). When the model computed reflectivity felt under the detection threshold of the radar, the corresponding grid-box is not included in the statistics, leading to the green dotted line in Fig. 1. By the end the green dotted line (model) must be compared to the solid blue line.

The ECMWF model (first column in Fig. 1) exhibits a very good agreement for mid-level clouds (between 3 and 7 km) over the three sites and an over-estimation of the high-level (above 7 km) and low-level cloud (below 3 km) occurrences. Although the overestimation is about the same over all sites for the low-level cloud occurrence (about 10%), the overestimation of the high-level cloud occurrence is much larger at Palaiseau (up to 20-25%) than at Chilbolton and Cabauw (5-10%).

The ARPEGE model (second column in Fig. 1) exhibits a radically different behaviour between the two cloud schemes used during the Cloudnet period. The diagnostic cloud scheme labelled arpege1 produces a very strong and systematic overestimation. This scheme also produces large occurrences of high-level clouds which are classified as "detectable" by the instruments, since there is no difference between the total profile and the profile with instrumental effects included. The modified ARPEGE diagnostic cloud scheme (labelled arpege2) significantly improves the frequency of occurrence but a systematic underestimation of about 5% appears up to 8 km altitude. It appears to produce the best overall estimate of cloud occurrence for all sites. This is somewhat surprising, since ARPEGE is the only model which does not treat clouds with prognostic equations.

The RACMO model (third column in Fig. 1) includes the same cloud scheme as ECMWF. It is nevertheless clearly characterized by a much larger and systematic over-estimation of cloud occurrence at all levels than ECMWF, except for high-level cloud occurrence (once the instrumental effects are included) for which it provides a better cloud occurrence than ECMWF overall.

The Met Office model (fourth column in Fig. 1) produces the best cloud occurrences of all models for the low-level clouds, overestimations similar to those observed in RACMO of the mid-level cloud occurrence, and the largest overestimations among models for the high-level cloud occurrence.

3. MODEL CLOUD FRACTION

Cloud fraction is an important parameter since it is a crucial input to the radiation scheme and it has now become a prognostic variable of the cloud scheme held in some NWP models (ECMWF and RACMO for instance). Once data are re-organised at the model scale, we compute the mean cloud fraction profiles and Contoured Frenquency by Altitude Diagrams (CFAD). The rationale for using CFADs in addition to the mean profiles considered in earlier studies (Mace *et al.* 1998 or Hogan *et al.* 2001) is to investigate how the model is distributing the cloud fraction values between 0 and 1. As for cloud occurrence we document the statistics of cloud fraction using the two years of the Cloudnet observations and compare with the four models. The results are shown in Fig. 2 only for the Cabauw site but they are valid for the two others. In this figure each line represents a model (ECMWF, arpege1, arpege2, RACMO and Met Office from top to bottom). For each line the first column shows the distributions obtained using the whole model sample of cloudy grid-boxes. The third column corresponds to the distributions obtained from the whole observation sample. Finally second and fourth columns show the distribution for model and observations respectively but for the sub-sample where models and observations agree on a cloud occurrence (the amount of points included in the statistics is decreased by about one third, see black number in the top left corner).

The first important thing to notice from Fig. 2 is that the mean profile (red solid line) does not correspond to the most representative class of cloud fraction. This shows that the evaluation of cloud fraction in models is difficult to achieve using mean profiles, since an accurate representation of the mean cloud fraction profile by a model could clearly be due to a wrong cloud fraction distribution.

By comparing column 1 and 2 on one hand and 3 and 4 on the other hand, it appears that the model and observation sub-samples on which they agree on a cloud occurrence are representative of the whole data set and the comparison of column 2 and 4 can be undertaken with confidence regarding its statistically significance.

The observed cloud fraction (column 3 or 4 in Fig. 2) is characterised by a bi-model distribution for low level clouds (clouds below

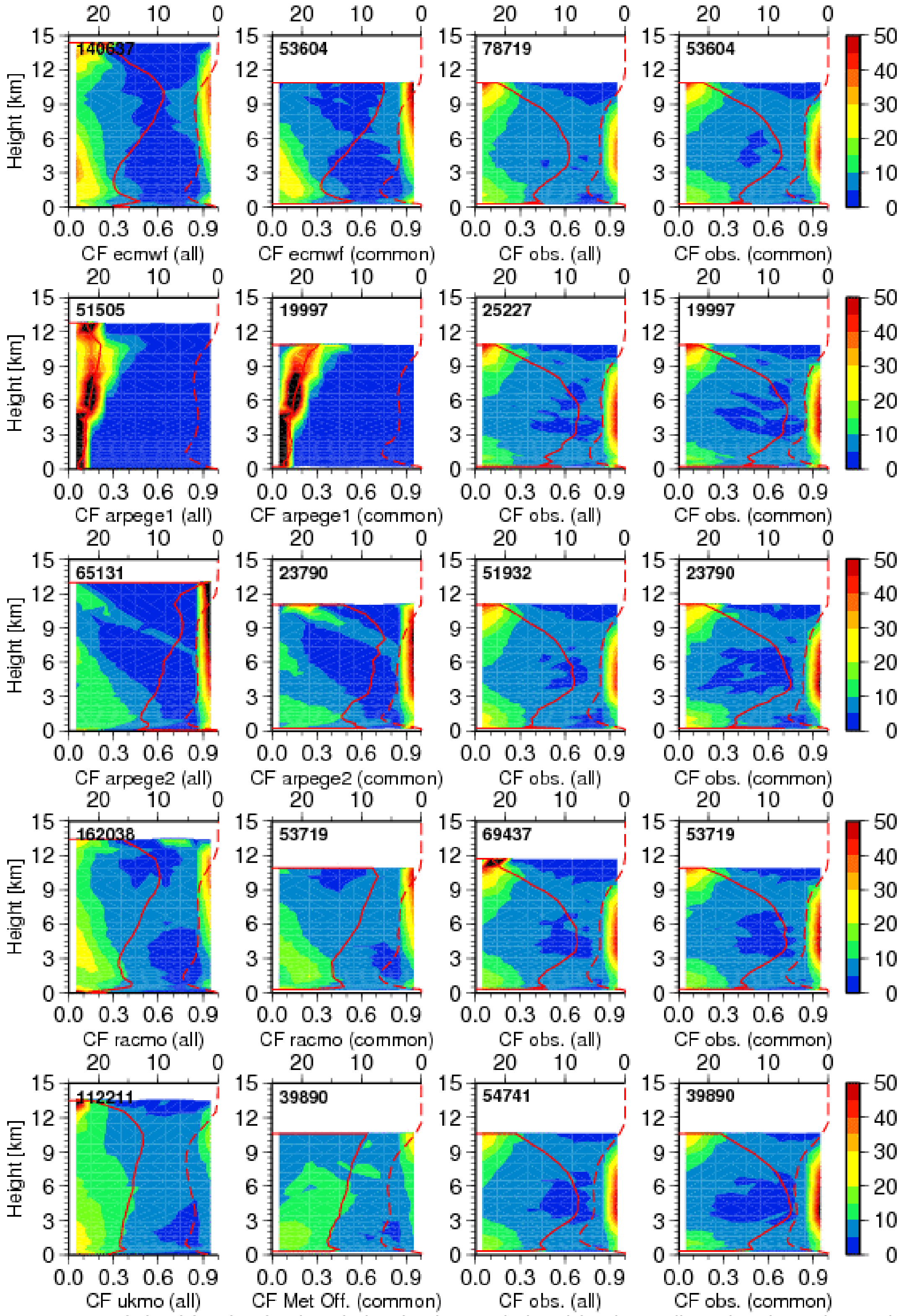

Figure 2: CFADs of cloud fraction for the whole Cloudnet period and for data collected at the Cabauw site. The red line (lower axis) shows the mean value. The red dashed line (upper axis) shows the amount (in percentage) of cloudy grid boxes at each level. The black number in the upper left corner of each panel corresponds to the total number of points used for statistics computations.

Each line is dedicated to a model (or model version), from top to bottom : ECMWF, arpege1, arpege2, RACMO and Met Office. The first column shows the distributions obtained from the models. The second column is the same but for the grid boxes for which model and observations agree on a cloud occurrence. The third column is the distribution obtained from all observations and the fourth shows the same distribution but when there is an agreement in occurrence between models and observations.

3 km) with about the same amount of clouds with low cloud fraction (less than 0.3) and high cloud fraction (larger than 0.85). At mid-levels (clouds between 3 and 7 km altitude) clouds are essentially characterized by large cloud fraction values (larger than 0.8). In contrast the high-altitude clouds (from 7 to 12 km) are characterized from 7 km up to 9-10 km by a bimodal distribution (very small and very large cloud fraction values), and above 10 km by very small cloud fractions (less than 0.2).

The models are not characterised by the same cloud fraction distributions. The ECMWF and RACMO (which use the same cloud scheme) present strong similarities, and therefore show the same skills and discrepancies with respect to the observations. The core of small cloud fractions observed for low-level clouds extends too much upward (up to 8 km instead of 3 km in the observations), which means that the ECMWF and RACMO models produce too many broken clouds at mid-levels. This signature is also reflected by the smaller amount of clouds characterized by high cloud fractions between 3 and 7 km (30% compared to the 50% observed, roughly) in these models. These models also clearly have difficulties to produce high-level clouds with small cloud fractions (high-level clouds are essentially characterized by very high cloud fractions in Fig. 2), which is not in agreement with the observations.

The Met Office model is characterized by a very different cloud fraction distribution, with much more small cloud fractions overall (between 0 and 0.4), and a much more homogeneous distribution of cloud fractions between 0 and 1 than in the observations. The most striking feature is the lack of mid-level clouds with a high cloud fraction, and high-level clouds with a small cloud fraction.

The impact of the change in cloud scheme in the ARPEGE model is obvious, as was also the case of the cloud occurrence profiles. As seen in Fig. 2, clouds with height cloud fraction values were not generated by the first ARPEGE cloud scheme at any height. The second scheme is clearly much better, with a good representation of the bimodal distribution of low-level clouds, and of the high cloud fractions in mid-level clouds. However, as it was the case of the other models it is less accurate in describing the distribution of cloud fraction in high-level clouds, with too many high-level clouds characterized by a high cloud fraction, although there is some indication of bimodality in this model for the high-level clouds, which claims for further improvements / tuning of this scheme.

4. ICE WATER CONTENT

The methodology as for cloud fraction is used (distributions and mean profiles are computed) in order to evaluate the ability of the model in representing clouds with the right IWC. The reason for looking in more details to this parameter is that it is the second parameter generally held in the model prognostic cloud scheme. There is no remote sensing instrumentation able to provide a direct measurement of IWC profiles. Therefore, Morcrette (2002) for instance took the option of computing the radar reflectivity from the ECMWF model. In their paper Bouniol *et al.* (2007b) show that by comparing statistically different methods (with different complexity degree) for the Cloudnet data set that they produce rather similar distributions. A good intermediate is to use the retrieval method of Delanoë *et al.* (2007) that retrieves the IWC from the radar reflectivity and the mean Doppler velocity without *a priori* assumption on the more representative particle habit for a given cloud.

Figure 3 displays the CFADs of IWC obtained from the model time series and from the observations at the Cabauw site. For each distribution the mean IWC profiles have also been computed (solid red lines), and the model profile of the second column is plotted as a dashed line in the panels of the fourth column for straight comparisons with the observations. Comparison of column 1 and 2 and 3 and 4, as for cloud fraction, shows that the sub-sample for which model and observations agreed on a cloud occurrence is reasonably representative of the whole sample at all heights. Some differences are observed at the top and bottom of the profiles, but in this region the number of points included in the analysis is much reduced anyway.

The observed IWC distribution (column 3 and 4 in Fig. 3) is fairly skewed, with a narrower distribution of the high values and a wider distribution of the small values. The comparison of the observed and model IWC distributions on the second and fourth panels of Fig. 3 clearly shows that the models tend to well reproduce this skewed distribution. The IWC distribution of the models is nevertheless generally narrower than the distribution obtained from the observations.

The comparison of the solid and dashed profiles on the fourth columns of Fig. 3 clearly shows that the shape of the mean IWC profiles is pretty well reproduced by all models. The ECMWF and arpege1 cloud schemes both tend to overestimate IWC above 4~km height, and slightly underestimate below. In the case

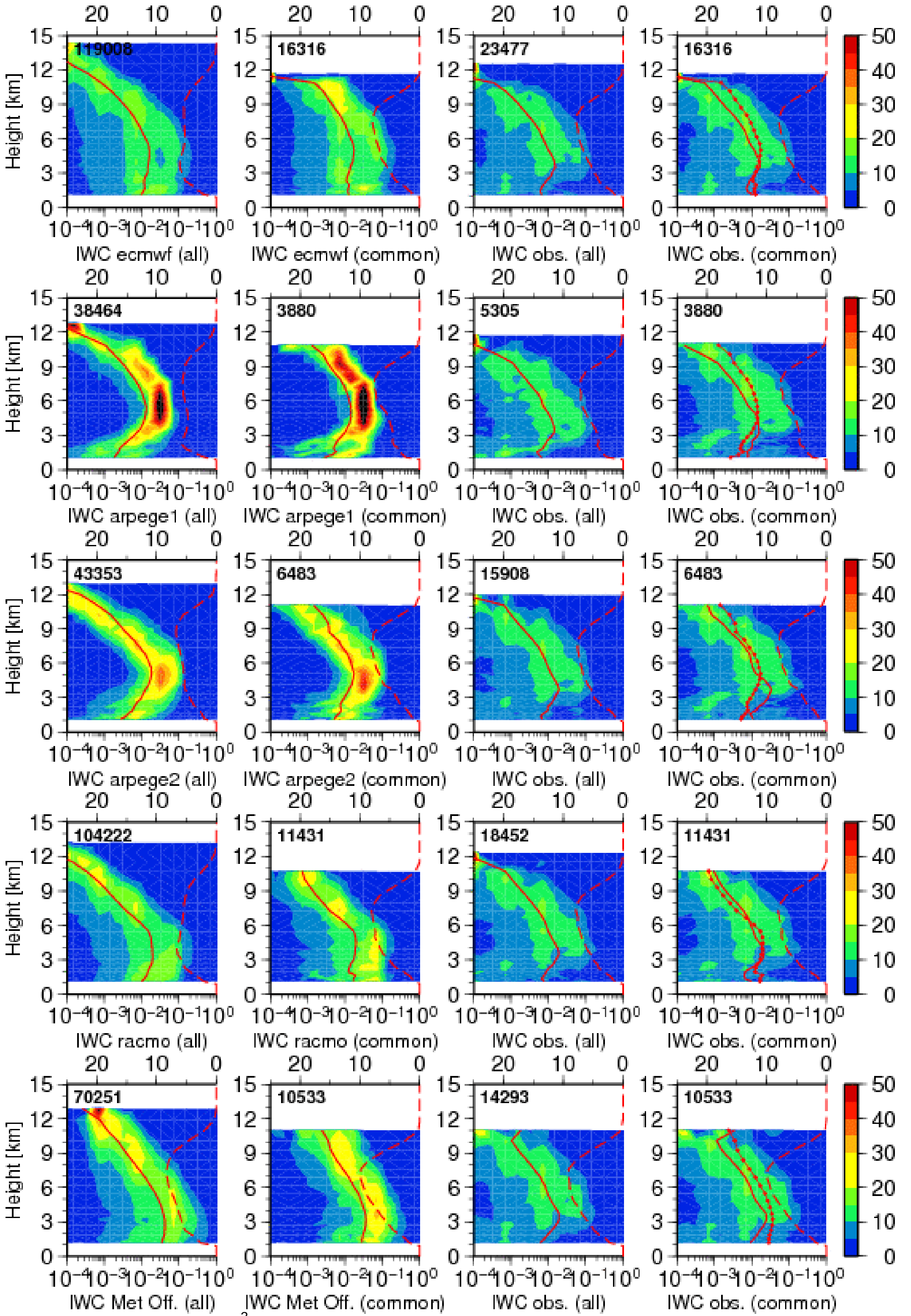

Figure 3: CFADs of IWC in g m^{-3} for the whole Cloudnet period and for data collected at the Cabauw site. The red line (lower axis) shows the mean value, the red dashed line (upper axis) show the amount (in percentage) of cloudy grid boxes at each level. The black number in the upper left corner of each panel corresponds to the total number of points used for statistics computations.

Each line is dedicated to a model (or model version) : ECMWF, arpege1, arpege2, RACMO and Met Office from top to bottom. The first column shows the distributions obtained from the models. The second column is the same but for the grid boxes for which model and observations agree on a cloud occurrence. The third column is the distribution obtained from the data and the fourth shows the same distribution but when there is an agreement on occurrence between models and observations. Superimposed in the fourth column (diamond lines) is the mean profile obtained from the model sub-sample in the second column for sake of direct comparison.

of ECMWF, this overestimation is the result of a too strong production of large IWCs, while in the case of arpege1 it is merely the result of an overall way too narrow distribution of IWC values. The second ARPEGE cloud scheme tends to much better reproduce the observed IWC profile than the first cloud scheme above 4 km height, but strongly underestimate the observed profile below 4 km height. The Met Office cloud scheme produces a systematic overestimation of IWC at all heights, with however a very good representation of the shape of the profile. This result suggests that the Met Office cloud scheme is good but probably needs some general tuning in order to produce smaller IWCs (there is a significant lack of small IWCs in this model). Finally, the RACMO model reproduces almost perfectly the observed IWC profile, as a result of a relatively good representation of the observed IWC distribution.

A more detailed analysis is performed in Bouniol et al. (2007b) showing that all models tend to overestimate the small IWCs and to underestimate the high IWCs, which in other words means that the model schemes all tend to generate a range of IWC values smaller than observed resulting in a concordant mean profile but a narrower distribution.

5. CONCLUSION

To summarise the main results of this evaluation, regarding the high-level clouds all the models tend to overestimate the frequency of cloud occurrence (even when the instrumental sensitivity is taken into account) and all schemes except the first version of the ARPEGE model fail to produce the low cloud fraction values observed at these levels. The IWC is generally overestimated, except for RACMO. So the picture is as follows: there are too many high-level clouds in models, not broken enough, and with too large IWCs. These clouds are considered as radiatively important because of their feedback on weather and climate and therefore their relatively inaccurate representation in models may not be negligible when computing fluxes with the radiation scheme.

Regarding mid-level clouds their frequency of occurrence is generally overestimated, except for the second version of the ARPEGE model scheme. They also generally appear too much broken in models. In contrast, the IWCs in mid-level clouds are generally well reproduced.

Finally the accuracy of low-level clouds occurrence is very different from one model to another: RACMO, ECMWF and arpege1 tend to overestimate the occurrence of such clouds, while the Met Office is pretty accurate and arpege2 underestimates their occurrence. On the other hand, only the arpege2 model is able to generate the observed strongly bimodal distribution of cloud fraction of the low-levels. All the other models tend to generate broken clouds. Another characteristics common to all models is their tendency to underestimate the width of the IWC distribution as compared to the observed one.

From this overall comparison it is interesting to see that the arpege2 model, although it is still using a diagnostic cloud scheme and a coarser resolution than RACMO and Met Office, produces overall the "best" representation of clouds .A further evaluation of the parametrisation would require a comparison of observed and model shortwave and longwave fluxes, as done by Morcrette (2002). However, it is to be noted that these fluxes computations in the models depend strongly (but not only) on both cloud fraction, IWC and cloud occurrence. As a result compensating effects may well produce a good agreement between model and observations but for wrong reasons.

Finally it appears clearly that the use of long time series is a perfect framework to test the improvement resulting from a change in the parametrisation as it has been highlighted for the two versions of the ARPEGE model. The use of the A-Train data set will be also very valuable in the future to investigate the cloud properties in regions where ground based observations are generally sparse but where clouds have a large influence on the climate system such as in the tropical belt (including Africa and South-America).

6. REFERENCES

Bouniol D., Protat A., Delanoë J., Pelon J., Donovan D.P., Piriou J.-M., Bouyssel F., Tompkins A., Wilson D., Morille Y., Haeffelin M., O'Connor E., Hogan R.J., Illingworth A.J.. 2007a: Using continuous ground-based radar and lidar measurements for evaluating the representation of clouds in four operational models. Part I: Cloud occurrence. *Quart. J. Roy. Meteor. Soc.*, submitted June 2007.

Bouniol D., Protat A., Delanoë J., Pelon J., Donovan D.P., Piriou J.-M., Bouyssel F., Tompkins A., Wilson D., Morille Y., Haeffelin M., O'Connor E., Hogan R.J., Illingworth A.J.. 2007b: Using continuous ground-based radar and lidar measurements for evaluating the representation of clouds in four operational models. Part II: Cloud fraction and ice water content. *Quart. J. Roy. Meteor. Soc.*, submitted June 2007.

Delanoë J., Protat A., Bouniol D., Heymsfield A., Bansemer A., Brown P., 2007: The characterization of ice clouds properties from Doppler radar measurements. *J. Appl. Meteor. Clim.*, in press.

Hogan R.J., Bouniol D., Ladd D.N., O'Connor E.J., Illingworth A.J., 2003: Absolute calibration of 94/95 GHz radar using rain. *J. Atmos. Oceanic Technol.*, **20**(4), 572-580.

Hogan R.J., Jakob C., Illingworth A.J., 2001: Comparison of ECMWF Winter-Season Cloud Fraction with Radar-Derived Values. *J. Appl. Meteor.,* **40**, 513-525.

Illingworth A.J., Hogan R.J., O'Connor E.J., Bouniol D., Brooks M.E., Delanoë J., Donovan D.P., Gaussiat N., Goddard J.W.F., Haeffelin M., Klein Baltink H., Krasnov O.A., Pelon J., Piriou J.-M., Protat A., Russchenberg H.W.J., Seifert A., Tompkins A., van Zadelhoff G.-J., Vinit F., Willen U., Wilson D.R., Wrench C.L., 2007: Cloudnet - continuous evaluation of cloud profiles in seven operational modeles using ground-based observations. *Bull. Am. Meteorol. Soc.*, DOI:10.1175/BAMS-88-6-xxx.

Mace G., Jakob C., Moran K.P., 1998: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. Geophys. res. Lett., 25, 1645-1648.

Morcrette J.-J., 2002: Assessment of the ECMWF Model Cloudiness and Surface Radiation Fields at the ARM SGP site. *Mon. Weather Rev.*, **130**, 257-277.

Protat A., Armstrong A., Haeffelin M., Morille Y., Pelon J., Delanoë J., Bouniol D., 2006: Impact of conditional sampling and instrumental limitations on the statistics of cloud properties derived from cloud radar and lidar at SIRTA. *Geophys. Res. Lett.*, **33**, L11805, doi: 10.1029/2005GL025340.