# ESA CONTRACT REPORT

Contract Report to the European Space Agency

# Quality assessment of cloud screening, cloud analysis and cloud height assignment for atmospheric motion vectors

April 2010

*Authors: S. Di Michele*

**European Centre for Medium-Range Weather Forecasts**
**Europäisches Zentrum für mittelfristige Wettervorhersage**
**Centre européen pour les prévisions météorologiques à moyen terme**

**ECMWF**

Contract Report to the European Space Agency

# Quality assessment of cloud screening, cloud analysis and cloud height assignment for atmospheric motion vectors

*Authors: S. Di Michele*

WP-2300 report for ESA contract 1-5576/07/NL/CB:
Project QuARL - Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, Berkshire, UK

April 2010

# ABSTRACT

This report summarizes the studies performed using CALIPSO and CloudSat data to verify cloud variables in the context of the ECMWF assimilation system. In the first part of this report, cloud-top height derived from AIRS IR radiances has been compared with the same quantity from CALIPSO. Two different estimates from AIRS have been considered, one derived from the cloud detection scheme and the other used within the assimilation of IR cloudy radiances. Being AIRS (on board AQUA) and CALIPSO both members of the A-Train constellation, the required spatial and temporal coincidence is guaranteed. In the second part, cloud-top height from CALIPSO has been used to verify the wind height assignment for the atmospheric motion vectors derived from geostationary cloudy radiances. In the last part of the study CloudSat has been used to validate the assimilation of microwave radiances in cloudy regions. The intercomparison is possible when considering observations from the AMSR-E microwave radiometer (on board AQUA) which, as CloudSat, is part of the A-Train constellation. First, 1D-Var cloud and precipitation analyses (part of the 1D+4D assimilation scheme) from AMSR-E have been compared to co-located CloudSat estimates. Then, the so-called 'all-sky' 4D-Var assimilation has been considered: 4D-Var first guess and analysis have been compared with statistics of cloud and precipitation from CloudSat where AMSR-E observations were assimilated.
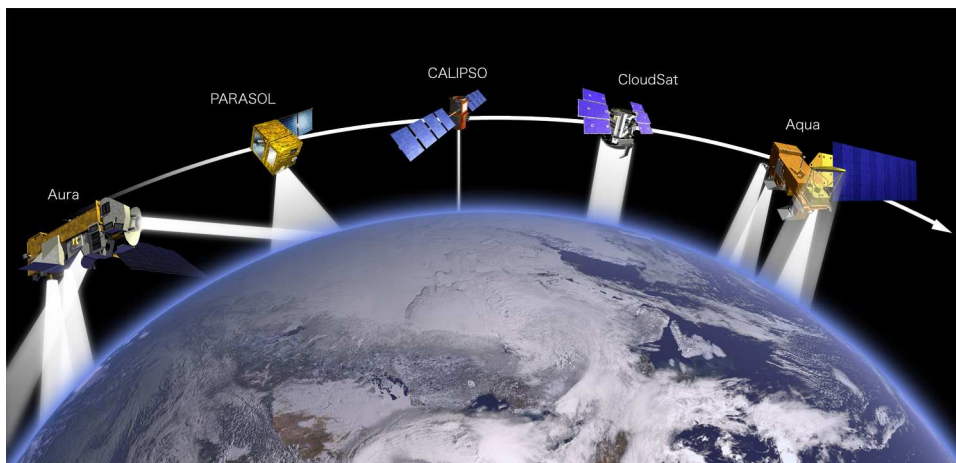
# Contents

*Figure 1.1: Schematic view of the A-Train constellation.*

# 1 Introduction

Knowledge of cloud properties is essential to understand their impact on the earth's radiation budget and on climate change. Satellite remote sensing of properties such as cloud amount and cloud-top height has the advantage of providing global coverage on a regular basis. Several active and passive satellite sensors with unprecedented observing capabilities are flying in a formation called the "A-Train". The constellation is composed by Aqua (Aumann *et al.*, 2003), Aura (Schoeberl *et al.*, 2006), the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) (Winker *et al.*, 2003), CloudSat (Stephens *et al.*, 2002), along with the Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar (PARASOL) (Fougnie *et al.*, 2007). These satellites are on sun-synchronous orbit and cross the equator a few minutes apart at around 1.30 PM local time, hence the nickname "afternoon constellation". Observations from a range of instruments on board the Aqua satellite, namely the Advanced Microwave Sounding Unit-A (AMSU-A), the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E, Kawanishi *et al.*, 2003) and the Atmospheric Infrared Sounder (AIRS, Aumann *et al.*, 2003) are operationally assimilated at the European Centre for Medium-Range Weather Forecast (ECMWF).

The synergy provided by the A-Train will be exploited in the first part of the study using CALIPSO to validate the information about cloud-top height (CTH) derived from AIRS. In the second part, although strictly not part of the A-Train, the height assignment (HA) for Atmospheric Motion Vectors (AMVs) derived from geostationary cloudy radiances will be also compared to CALIPSO. In the third and final part of the work, CloudSat products will be used to evaluate cloud variables in the analysis when coincident AMSR-E microwave (MW) radiances in cloudy areas are assimilated in the ECMWF forecasting system.

# 2    Verification of Cloud-Top height from AIRS using CALIPSO

The A-Train configuration will be exploited in this section taking advantage of the spatially and (almost) temporally coincident observations from AIRS (on board Aqua) and from the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) (on board CALIPSO). AIRS is a thermal infrared (IR) grating spectrometer with 2378 spectral channels ranging from 3.7 to 15.4 $\mu$m, scanning ±48.95°off-nadir, and with a footprint size of 13.5 km. The actual AIRS instantaneous angular field-of-view (FOV) is 1.1 × 0.6. Assuming a circular FOV of 1.1 yields a geolocated nadir footprint measuring 13.5 km in diameter. For off nadir scanning angles these footprints become elliptical, and at the maximum scanning angle of 49.5 the major (cross track) and minor (along track) axes are approximately 21 and 33 km, respectively. CALIOP is a nadir-viewing, three-channel lidar system (1064 nm and 532 nm parallel and perpendicular) with a 1 m receiving telescope (McCormick, 2004). CALIPSO provides cloud profiles on a global scale, and it is particularly useful for the detection of cloud-top height. The lidar vertical resolution is of 30 m. Due to constraints on CALIPSO's downlink bandwidth, this full resolution is only available from -0.5 km below sea level up to about 8.3 km above sea level. From 8.3 km to about 20.2 km the vertical resolution is of 60 m and above 20.2 km it is 180 m.

CALIPSO observations have been be used to investigate the accuracy of CTH derived from AIRS measurements. Two different methods for estimating a CTH value from AIRS are in use at ECMWF. The first is part of the cloud screening algorithm for advanced IR sounders and it is actually designed to give an upper bound to the actual CTH. The second is instead employed for the assimilation of cloud-affected IR radiances.

## 2.1    Cloud-Top height from CALIPSO

CALIPSO CTHs have been taken from the official Level 2 'layer product' (Vaughan *et al.*, 2009). This product provides the vertical location of cloud/aerosols boundaries for up to 10 possible layers, specifying spatial and optical characteristics of each feature found. It includes quantities such as layer base and top altitudes, integrated attenuated backscatter, layer-integrated volume depolarization ratio, and optical depth. It is available at three horizontal resolutions, namely 333 m (instrument native resolution), 1 km, and 5 km. In this study, the layer product giving the cloud detection at 1 km resolution has been used. The signal averaging performed for the 1 km product is a good compromise between better retrieval accuracies (due to increased signal to noise ratio) and the loss/smearing of features in heterogeneous scenes. Coincident CALIPSO-AIRS observations are identified searching for all possible CALIPSO shots falling within the FOV of each AIRS observation. The search is performed only for the AIRS scan pixels close to nadir, i.e. those with a possible overlap. AIRS FOVs are assumed circular with a nominal diameter size of 15 km and the elliptic distortions (increasingly) occurring across the off-nadir scan angles are neglected. This approximation leads to an enhanced computational efficiency when collocating the data.

Since the CALIPSO product consists of 1 km-spaced along-track samples, AIRS FOVs can be intersected by up to 15 lidar shots. Cloud-top height from CALIPSO is evaluated as the mean of those shots located within an AIRS FOV.

## 2.2    Verification of Cloud-Top height from cloud screening

IR data are frequently affected by clouds. Although efforts are underway to assimilate cloud information from IR sounders (e.g. McNally, 2009), numerical weather prediction (NWP) centres operationally use cloud-screened data, i.e. observations are processed before being assimilated to remove cloud-contaminated channels. At ECMWF, cloud screening checks are in place for IR sounders such as AIRS, the High Res-

olution Infrared Radiation Sounder (HIRS) or the Improved Atmospheric Sounder Interferometer (IASI). The A-Train constellation provides a unique opportunity for a global evaluation of the cloud detection and screening technique currently used.

### 2.2.1   Cloud-Top height from cloud screening

McNally and Watts (2003) developed a method to detect cloud in radiance data measured by high spectral resolution IR sounders. This method ensures that only channels whose weighting functions peak well above the cloud tops are used in the analysis. The presence of clouds and which channels are affected by cloud emission are determined evaluating observation-minus-model radiance departures and performing threshold checks using a digital filtering approach.

The methodology consists of comparing an observed spectrum with a spectrum computed from the best available background estimate of the cloud-free atmospheric state which, in the data assimilation context, is provided by a short-range NWP forecast. In order to isolate the cloud signal in the departure vector, the channels are first re-ordered according to a ranking that reflects their relative sensitivity to the presence of cloud. Each channel is assigned an altitude (or pressure) at which a black cloud must be placed in order to give an overcast radiance that differs from the clear-sky radiance by an amount in excess of a pre-defined threshold. The channels are then ordered on the basis of this altitude. As the radiance change in a given channel is dependent on the atmospheric state, ranking is performed at each sounding location, separately.

The ranking obtained with this approach is equivalent to ordering channels according to the lower tail of their respective weighting functions. In the ranked/band-partitioned radiance departures, the cloud signal is effectively organised into a smoothly varying function that increases monotonically (negatively over warm surfaces and positively over cold surfaces) with increasing channel cloud sensitivity.

With these characteristics in mind, an algorithm has been formulated to identify the point (i.e. the channel in the ranked space) where the cloud signal is first identifiable. It consists of a lowpass filter (that removes as much of the other noise sources, which are not organised by the ranking, as possible) followed by a search for the channel at which a monotonically growing departure can first be identified. Having found this channel, all channels ranked more sensitive are flagged as cloudy and those ranked less sensitive are flagged clear. CTH from AIRS screening can be defined as the altitude that a cloud must have to start affecting this channel.

### 2.2.2   Results

The CALIPSO-AIRS collocation procedure described above has been applied to outputs from the ECMWF assimilation system (CY35R3, operational from September 2009 to January 2010) over a period of 10 days (1-10 June 2009). The matching resulted in 24746 points where CTH measurements from CALIPSO are coincident with CTH from AIRS cloud detection.

A scatter plot comparing the values from AIRS (on the ordinate) with those derived from CALIPSO (on the abscissa) is given in Fig. 2.1. Situations over water or land surfaces are represented with different colours, but this distinction does not highlight any particular features and the two CTH products seem to be poorly correlated in both cases. It can be noted that CTH from AIRS assumes only discrete values. This is the consequence of the detection algorithm, whose altitudes (pressures) refer to the levels of the forecast model in use.

While for the optimal case AIRS and CALIPSO CTHs should be identical, situations where the AIRS CTH exceeds that derived from CALIPSO are still acceptable since the aim of the AIRS cloud detection algorithm

is the screening of cloud-contaminated channels. From Fig. 2.1, however, it seems that a substantial number of points exists where the AIRS CTH is smaller than that from CALIPSO. The spread (and the number of points with AIRS CTH lower than the one from CALIPSO) can be reduced by screening (i.e., excluding) situations for the CALIPSO CTH appears less reliable. For this purpose, cases where CALIPSO shots within the AIRS FOV were less than 10 (out of maximum 15) have been rejected. Also, the investigation has been restricted to situations over the ocean where, differently than over land, the detection algorithm works better since the radiometrically-warm background of water surfaces provides a good contrast.

The cases remaining after these restrictions are shown in Fig. 2.2. AIRS and CALIPSO appear now more correlated, however there is still a significant number of contradictory cases. We see that AIRS overestimates CALIPSO values most significantly in the lowest kilometres (or completely misses clouds detected by CALIPSO at these levels). As explained below, many of these cases occur for situations of broken cloud fields. Under these conditions only few of the CALIPSO shots within the AIRS FOV sense some cloud and therefore their mean is not representative. Different colours in Fig. 2.2 indicate the regions where observations are located. It is interesting to note that many of these situations occur in the Tropics: they can be linked to convective structure, which usually extend over horizontal domains less than a few kilometres. It makes sense to keep these cases out of the intercomparison both because the CALIPSO mean is not meaningful and because the AIRS measurements are affected by surface contamination.

Fig. 2.3 shows the scatter plot of the points retained after checking that all CALIPSO shots used in the matching do sense some cloud (and therefore are used to define the mean value). In this scatter plot, in the range of CALIPSO heights between 2 km and 8 km, the majority of points correctly lies along or above the diagonal. Also, for CALIPSO CTHs below 2 km, we note a cluster of points were AIRS values are too high. This feature is somehow expected since low clouds are difficult to detect using IR radiances.

However, as explained above, the greatest problem for cloud-screening applications is in situations where the AIRS algorithm produces too low estimates. According to Fig. 2.3 this mostly occurs at higher levels (above 8 km). To investigate this behaviour we considered additional information that CALIPSO provides. Beside the altitude of the cloud-top, a lidar also gives an indication about the opacity of the cloud: when the backscattered signal is fully attenuated we are in the presence of an optically-thick cloud (at the lidar wavelength). This information has been used in this comparison flagging as *opaque* those cases when none of the CALIPSO shots in the AIRS FOV reaches the ground. Occurrences where a cloud is detected for at least 10 shots in the AIRS FOV, but the rays reach the surface are flagged as *transparent*.

The result of this classification is shown in Fig. 2.3 where a different colour is used for the two categories. We note that most of the cases when CALIPSO CTH is above 8 km but AIRS assigns a lower altitude are flagged as transparent. Therefore the underestimation in these situations can be explained by the small optical depth: the detection algorithm is mislead by the low cloud emissivity. Given their height, above 10 km, they are likely to be cirrus clouds. Fig. 2.2, suggests that they are mainly located in the Tropical regions. Interestingly, the cases missed by AIRS (i.e. the points lying along the abscissa) also belong to the transparent class. The small optical depth of these clouds provides also an explanation for the failure of the AIRS detection.

As a final note, we must mention that the disagreement between the two products shown in Fig. 2.3 is not only due to differences in measurement principles. An additional reason can be found in the different viewing geometry of AIRS and CALIPSO. These instruments do not actually sense the same scene since the lidar does not perform any scanning, but it just observes a vertical two-dimensional scene. Although this scene is contained within the same volume of the atmosphere sensed by AIRS radiances, it is not guaranteed that it is representative of the AIRS FOV as it is only a sub portion. Although not explored in this study, the spatial representativeness could, to some extent, be verified by resorting to a coincident third observation, for example the Moderate Resolution Infrared Spectroradiometer (MODIS) flying on Aqua. Unfortunately,

there are no coincident observations with a comparable vertical resolution so that such a verification would have to be limited to features of the CALIPSO data which are not sensitive to the vertical location of the cloud (as, e.g., the total cloud cover).
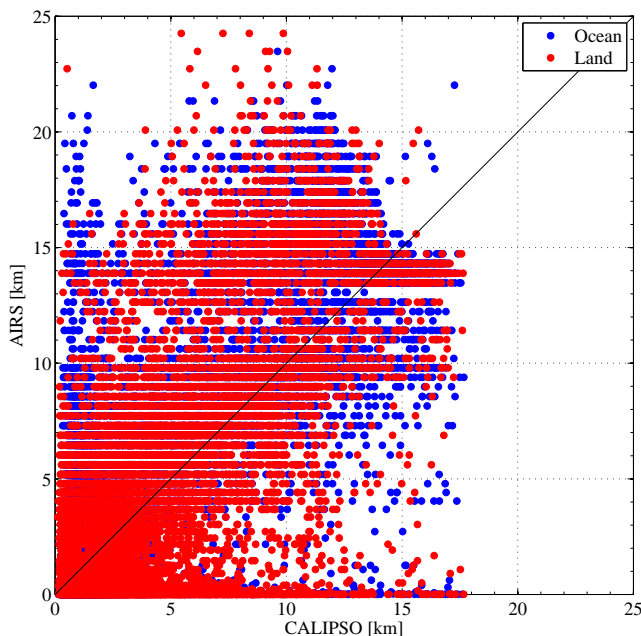


*Figure 2.1: CTH from AIRS cloud detection algorithm vs. CTH from coincident CALIPSO Level-2 Layer Product (averaged to AIRS resolution). Blue dots and red dots indicate AIRS observations over ocean and over land, respectively.*



*Figure 2.2: As Figure 2.1, but using only occurrences over ocean where more than 10 CALIPSO shots are found within the AIRS FOV. Black, blue, and red dots respectively indicate AIRS observations at South mid-latitudes, at North mid-latitudes, and over the Tropics.*

*Figure 2.3: As Figure 2.2, but using only occurrences where all CALIPSO shots within the AIRS FOV see a cloud. Black and red dots respectively indicate opaque and transparent occurrences. See text for explanation of their definition.*
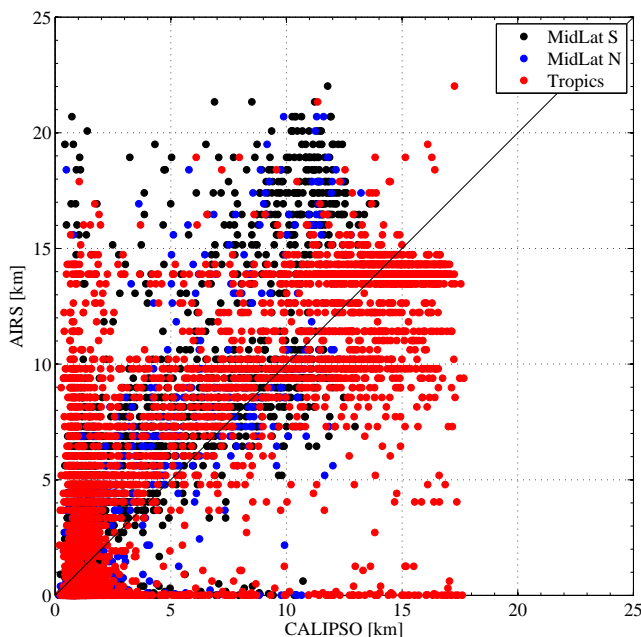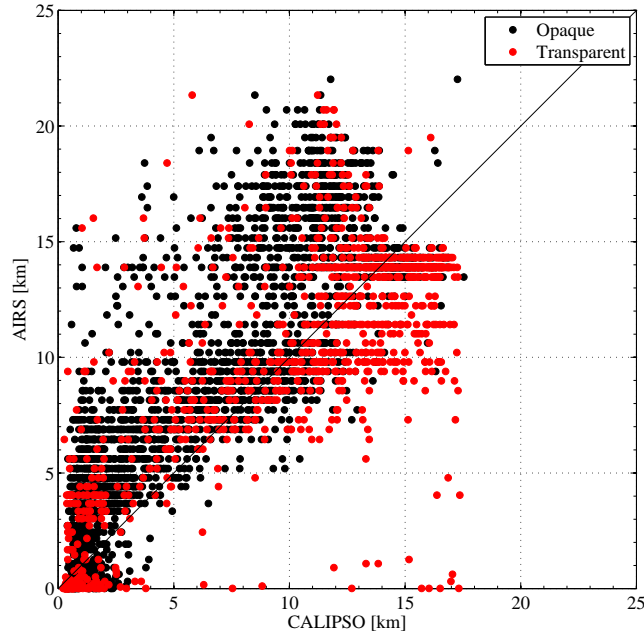
## 2.3 Verification of Cloud-Top height for assimilation of cloudy radiances

### 2.3.1 Cloud-Top height for assimilation of cloudy radiances

At ECMWF, a method has been developed to directly assimilate IR radiances in cloudy conditions within the 4D-Var analysis system (McNally, 2009). The 4D-Var analysis control vector has been extended to include parameters which describe cloud features and which are estimated simultaneously with temperature and humidity inside the main analysis cycle. The new cloud-related parameters are the cloud-top pressure ($C_P$) and the effective cloud fraction ($C_F$). The latter takes into account semi-transparency and represents the equivalent amount of opaque cloud in the instrument field of view.

The scheme only assimilates cloud-affected radiances in overcast conditions – thus avoiding many complications associated with the forward modelling and analysis of fractional cloud. Focusing on completely overcast scenes and using all available channels at these locations, this scheme departs from the more traditional approach of trying to use only weakly or moderately cloud-affected radiances. Background estimates of the additional analysis variables that describe clouds are not taken from the NWP model (since they do not provide accurate enough cloud information). Instead, a small subset of the available AIRS channels at a given location are used to define background estimates of cloud parameters. Following the method described in Eyre and Menzel (1989), values of $C_P$ and $C_F$ of an opaque cloud are found minimising the least squares fit to the observed radiances through the following cost function $J$:

$$J = \sum_{i=1}^{N} [R_{OBS} - R_{CLD}(C_P, C_F)]^2$$

where $R_{OBS}$ is the observed radiance and $R_{CLD}$ is the computed cloudy radiance, the latter being a combination of the computed clear-sky radiance $R_{CLR}$ and the corresponding overcast radiance $R_O$ ($C_P$):

$$R_{CLD}(C_P, C_F) = (1 - C_F)R_{CLR} + C_F R_O(C_P)$$

with inputs of temperature and humidity taken from the background model fields. Two channels, one around 13 $\mu$m and one around 11 $\mu$m, (i.e. AIRS channels 355 and 787) are used in the fitting process.

### 2.3.2  Results

CTHs derived via the cost function minimisation method of Eyre and Menzel (1989) were compared against CALIPSO following the same procedure employed for the verification of the AIRS cloud-screening heights discussed in Subsection 2.2. For the same dates as in subsection 2.2 (1-10 June 2009), cloud-top pressure and the effective cloud fraction obtained from AIRS observations using the above method were extracted from ECMWF operational analyses. Cases with coincident CALIPSO observations were then sub-selected. CTHs from CALIPSO were made comparable to the AIRS ones averaging the 1-km Level 2 Layer Product to AIRS resolution (15 km). As done in Section 2.2 for the cloud detection, some checks were performed to ensure the quality of CALIPSO averages and AIRS retrievals:

• AIRS cases matched but having less than 10 (out of maximum 15) CALIPSO shots within the FOV were rejected;

• AIRS cases matched but containing CALIPSO shots in clear sky were rejected (broken clouds);

• Again, only cases over ocean were considered;

• A screening for non-physical solutions, e.g. yielding levels which are too high or too low (below surface), was also needed (although not used in the assimilation, such cases are still contained in analysis output).

Fig. 2.4 shows CTH from AIRS against matched CALIPSO value for the cases where all the above conditions are verified. We note occurrences with reasonable agreement, but also situations with large discrepancies. For clouds below 5 km (as measured by CALIPSO) AIRS estimates typically overestimate the CTH. The same is true for very high clouds while at intermediate levels (between 5 and 12 km) cases where the AIRS CTH exceeds that of CALIPSO are not uncommon. These deviations (both over and underestimations) can be substantial and height differences of several kilometres are frequently observed. The opaque vs. transparent distinction introduced in the previous paragraph helps explaining some of the differences. For example, we note that the overestimation above 10 km and below 5 km usually occurs for cases of transparent clouds (red dots). As mentioned at the beginning of this section, cloud fraction is also retrieved by the algorithm together with CTH. Selecting instances with high values of cloud fraction should give better agreement with CALIPSO CTHs since partially filled AIRS FOV give rise to ambiguous (cloudy) radiances that in turn lead to errors in the retrieval. This effect is shown in Fig. 2.5 where, on the same scatter plot given in Fig. 2.4, situations with retrieved cloud fraction above 0.75 are plotted in red. As expected, the agreement of these cases with CALIPSO is larger when compared with Fig. 2.4. Interestingly, all these occurrences are also flagged as opaque according to CALIPSO.

## 2.4  Summary

The investigation above has shown that CALIPSO observations constitute a valuable tool for the validation of cloud-top height retrieved from IR radiances. The intercomparison revealed the need for screening cases with non-contiguous clouds in the radiometer FOV (broken) to take into account differences in resolution between the instruments. The study has also shown that an opaque/transparent classification highlights the limitations of deriving CTH from passive instruments. The verification of cloud-detection for hyperspectral IR sounders has revealed some difficulties with cirrus clouds.
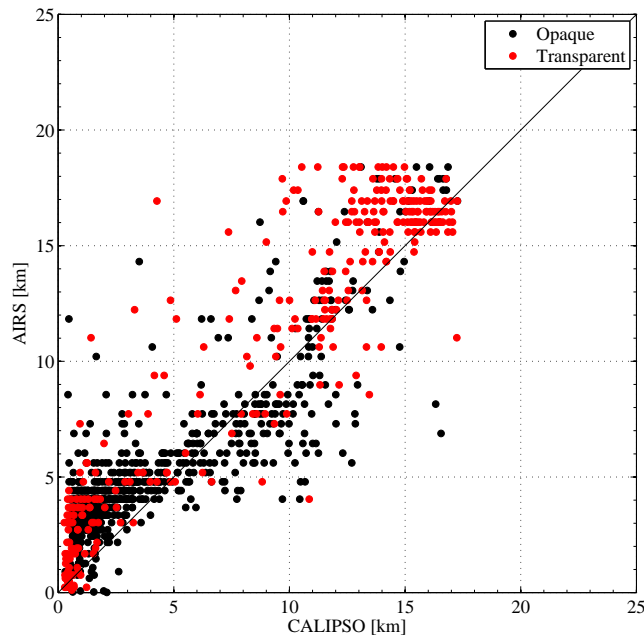
*Figure 2.4: CTH from AIRS retrieval algorithm of cloud top pressure and cloud fraction vs. CTH from coincident CALIPSO Level-2 Layer Product (averaged to AIRS resolution). Considered occurrences are the ones over ocean, with more than 10 CALIPSO shots in the AIRS FOV, and all these shots seeing cloud. Black and red dots respectively indicate opaque and transparent occurrences. See text for explanation of their definition.*
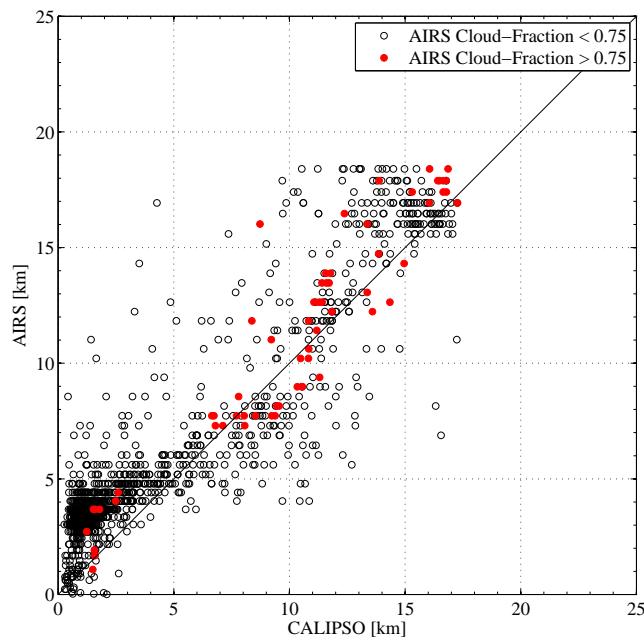


*Figure 2.5: As Figure 2.4, with red dots now indicating occurrences where AIRS-retrieved cloud fraction is higher than 0.75.*

The verification of cloud-top height for the assimilation of IR cloudy radiances has shown a reasonable agreement in overcast situations, but a larger scatter in conditions of partial beam filling.

# 3 Verification of AMV height assignment using CALIPSO

Atmospheric motion vectors derived from image sequences obtained from geostationary or polar orbiting satellites are an important part of ECMWF's operational data assimilation system. Currently, AMVs from 5 geostationary and 2 polar satellites are assimilated operationally. However, the monitoring of AMVs against short-range forecast information from within the assimilation or against collocated radiosonde observations often shows considerable biases or large random deviations in certain geographical regions. It is generally accepted that a large proportion of these biases or deviations can be attributed to the indirect measurement method of AMVs, i.e. the AMV processing and the interpretation of the AMVs as single-level wind observations.

Further analysis of the problems and the interpretation of the AMVs have been difficult for the lack of independent observations with sufficient coverage and detailed information on winds and clouds. With the launch of the CALIOP lidar embarked on the CALIPSO platform new insights in cloud altitude have been available. In this section, in a similar way to what has been done in Section 2, CALIPSO will be used to validate the height levels assigned to AMV winds when derived tracking cloud-top features.

## 3.1 Height assignment from Meteosat-9 AMVs

Height Assignment is currently the most challenging task in the generation of AMVs. Objective of this study is the comparison of AMV's HA with CTH observations from CALIPSO. In particular, we will investigate HAs derived from the Meteosat-9 geostationary satellite. Meteosat-9, together with Meteosat-8, constitutes the EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) Meteosat Second Generation (MSG) satellites. They are located over the Gulf of Guinea and thus they provide the primary service for Europe and Africa.

Main instrument on MSG satellites is the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI), which has twelve channels: three in the visible (0.4-1.1, 0.6, 0.8 $\mu$m), one in the near-IR channel (1.6 $\mu$m) and eight in the IR (3.9, 6.2, 7.3, 8.7, 9.7, 10.8, 12.0, 13.4 $\mu$m).

EUMETSAT currently generates AMVs from five SEVIRI channels: one visible channel (VIS-0.8, channel 2), one IR channel (IR-10.8, channel 9), two water vapour (WV) channels (WV-6.2 and WV-7.3, channels 5 and 6), and one high-resolution visible channel (HRV, channel 12). AMVs from each of these channels are based on the displacement of cloud targets are extracted, whereas the two WV channels derive clear-sky winds as well (De Smet, 2008).

When comparing CALIPSO with AIRS, the A-Train configuration has guaranteed the (almost) temporally coincidence of the two sets of observations. When considering an instrument in geostationary orbit like SEVIRI on board MSG a new *ad hoc* strategy must be used for the matching with CALIPSO. SEVIRI scans the Earth with three detectors in each spectral band. This enables a 3-km sampling distance between individual measurements (1 km for HRV) and a full disk coverage every 15 min. AMV products from the MSG satellites are derived hourly, at half-past reference time. Four (15 min-) disk images are used: two images before, one after, and one at the AMV reference time itself (EUMETSAT, 2009). We have chosen to match each AMV reference time with all CALIPSO observations taken within 30 min before and 15 min after this time. In this way, all cloud observations from CALIPSO corresponding to the four images (used to define the AMV HA) are taken into account.

EUMETSAT derives AMVs from the MSG satellites at a resolution of 80 km (24 pixels by 24 pixels), except for the HRV, where winds are derived with an average vector density of 32 km at sub-satellite point. The resolution of HA can be thought as the one of the tracer box used to estimate the AMV. In case of MSG, the

box is chosen of 15 pixels by 15 pixels in size, corresponding to a 48 km domain at the 3 km nominal MSG spatial resolution (EUMETSAT, 2009). As spatial matching, we therefore searched for all CALIPSO shots falling within a radius of 24 km from the AMV reference location. This radius reflects the need of overlap between CALIPSO shots and the AMV target.

The study was based on Meteosat-9 AMVs relative to a period of 10 days (1-10 June 2009). Fig. 3.1 shows the locations of AMVs successfully matched with CALIOP after applying to those data the process described above. Each panel refers to one of the four methods used to derive AMVs targeting clouds. AMVs from visible (top-left) are mainly over ocean, especially in the South Atlantic. Being available only on daytime, they are less in number that those produced by the other methods. AMVs from IR at 10.8 $\mu$m (top right) are the densest, working over a wider range of cloud types. The ones from WV channels at 6.2 $\mu$m (bottom-left) and 7.3 $\mu$m (bottom-left) are not present in South Atlantic, where low-level clouds are dominating (these clouds are tracked using the visible channel).

## 3.2 Results

HA for AMVs plotted against collocated CALIPSO CTHs are shown in Fig. 3.2. AMVs with Quality Indicator (QI) lower then 50 have been kept out in order to consider only the most reliable HAs. Situations with less than 35 CALIPSO shots (within the 24 km radius circle centred on the AMV location) have been also removed. Similarly to the AIRS intercomparison, the latter condition screens for situations where the mean of CTHs from CALIPSO shots is not representative of the AMV domain. Each panel of Fig. 3.2 contains HAs for one of the four AMV methods under consideration. The Visible channel (top-left) is only used to derive AMVs from low-level clouds, therefore HA values are always smaller than 2.5 km. IR channel at 10.8 $\mu$m (top-right) spans all heights, while the WV ones (bottom panels) are used for clouds above 5 km. Each of these scatter plots show a reasonable agreement between HAs and CALIPSO, but they also show the presence of several bad cases. The colour of the dots in Fig. 3.2 indicates if the AMV is over ocean (blue) or over land (red) surfaces. In all panels we note a larger dispersion for situations over land, indicating that HA is here more difficult.

In the analysis of CTH from AIRS it was shown that the agreement improves when considering cases where CALIPSO does not have clear-sky shots in the averaging domain. The same condition has been applied here to identify CALIPSO shots matched with AMVs that are observing *homogeneous* clouds. Fig. 3.3 shows the same cases of Fig. 3.2 plotting in red the ones flagged as homogeneous and in black the remaining (broken clouds). For each of the four AMV methods, HAs of homogeneous cases agree to CALIPSO better than the ones that do not satisfy the condition. For example, in the AMVs from IR (panel b) and in the AMVs from WV (bottom panels) the condition for homogeneity highlights some situations where HAs are above 5 km and much larger than CALIPSO. These bad cases are mainly flagged as non-contiguous. We also note that in the Visible method (panel a) and in the IR channel (panel b) those occurrences where HA is below 2.5 km, but CALIPSO CTH is much larger are not flagged as homogeneous. A possible explanation for this underestimation can be found in the contamination to the observed radiance by the (warm) surface background occurring in case of broken clouds (aliasing).

Homogeneous cases have been then further classified searching for optically-thin clouds. As done in the previous paragraph, a case is flagged as *opaque* if all CALIPSO shots within the radiometer FOV are fully attenuated by the cloud and *transparent* if all shots sense both a cloud layer and the surface. The result of this classification is shown in Fig. 3.4. Opaque cases (blue dots) show only a slightly larger correlation, although few occurrences are still present in both IR and WV AMVs where HAs are by far larger than CALIPSO CTH. Panels b), c), and d) show that clouds above 10 km are mainly classified as transparent (red dots). For these cases HA is usually lower than CALIPSO CTH, maybe because of the difficulty of SEVIRI to sense such thin clouds (likely cirrus).

Figure 3.5 shows the AMV locations using different colours according to the result of the homogeneous and opaque flagging. Most of the AMVs from the visible frequency range (panel a) do not satisfy the condition of homogeneity (grey dots). As for the AMVs from IR and WV (panels b, c, d) we note that the ones flagged as opaque (magenta dots) are present mostly at mid latitudes, while the ones flagged as transparent (cyan dots) do not show any geographical pattern.

Finally, it is important to note that the mean of (non zero) CALIPSO CTH shots (in the AMV domain) is not always the most appropriate quantity for comparing with the AMV HA. Restricting the analysis to homogeneous and opaque cases, Fig. 3.6 compares the AMV HAs not only against the mean values (blue dots) of the matched CALIPSO CTH but also against the the corresponding maximum values of the CALIPSO CTHs (red dots). Panels b), c) and d) show, as expected, that the cluster of points moves slightly to the right (off the diagonal) when the maximum is used instead of the mean. Interestingly, using the maximum (instead the mean) has its strongest impact on the outliers where the HA is much higher than the CALIPSO mean. Apparently, for most of the cases, these outliers correspond to situations where the CALIPSO mean differs significantly from its maximum and the CALIPSO maximum corresponds much better to the AMV HA.

This behaviour can be explained considering the way how the height of AMVs is defined. AMVs are derived identifying and tracking across images the cloud features (tracers) with stronger signal. HA is defined determining the height of the coldest (highest) portion of the cloud in the domain of such tracers. This height value can be much higher than the rest of the tracked cloud feature. In these cases, the use of the average for CALIPSO is particularly inappropriate since, differently from the HA, the mean *weights* all CTHs in the tracer domain, thus resulting in too small values.

## 3.3  Summary

The above intercomparison has shown an overall tendency of height assignment for AMVs to produce CTHs lower than CALIPSO, especially for high clouds. The study has revealed that the agreement is sensitive to the definition of the *averaged* CALIPSO CTH in the AMV domain. Similarly to the AIRS validation, we have found that a screening for spatial representativeness is needed after the co-location. Also, the opaque/transparent distinction of CALIPSO observations highlights those most difficult cases that cannot be easily retrieved from passive instruments. A cloud can be detected by a lidar, while it can exhibit such a small optical depth not to have any signature in visible or IR radiances.
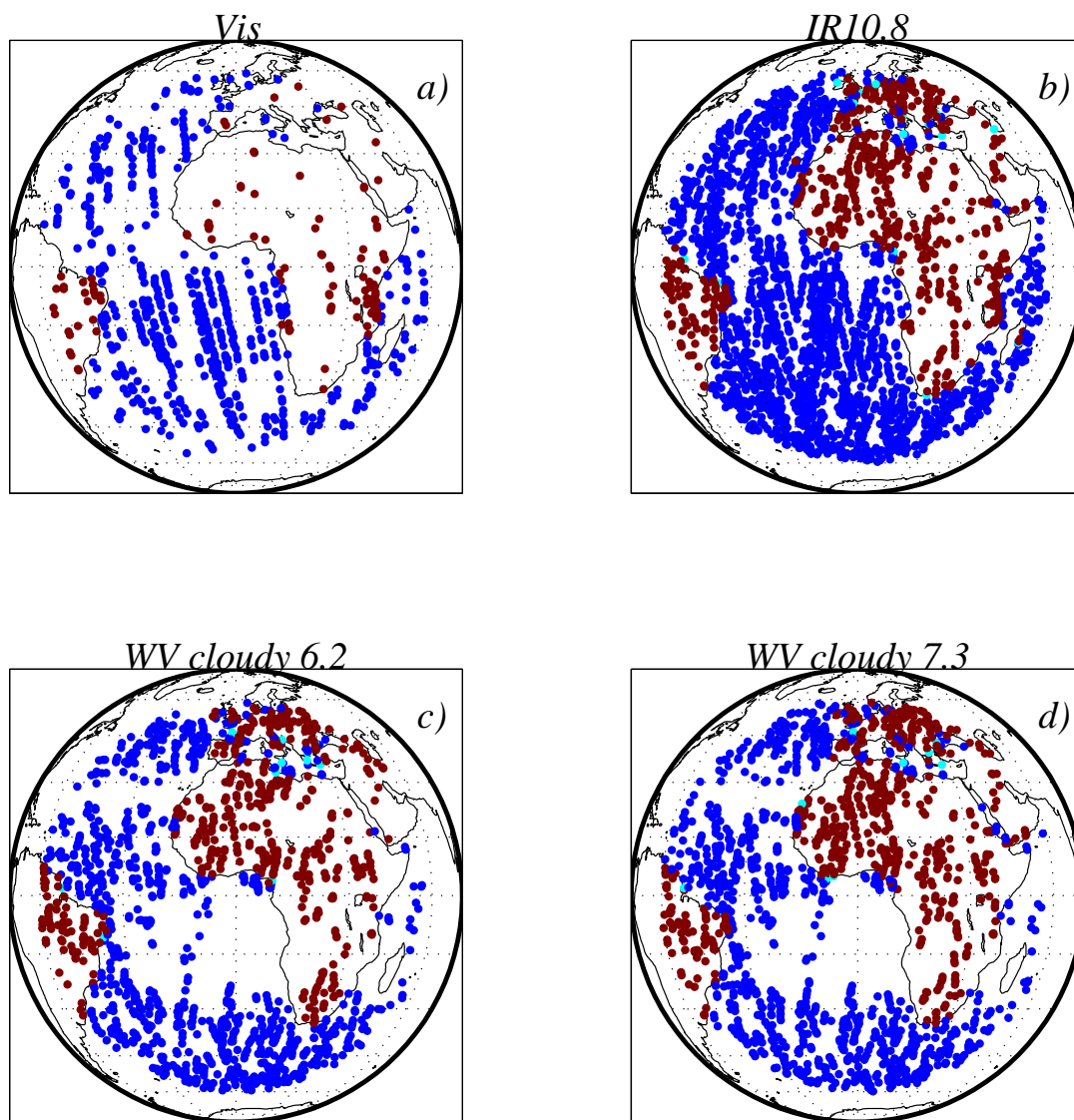
Figure 3.1: *Geographical position of Meteosat-9 AMVs collocated with CALIPSO observations for the studied period. Only AMVs with QI>50 are considered. Each panel contains AMVs generated by a different channel as indicated in the panel's title: a) Visible; b) IR at 10.8 μm; c) WV at 6.2 μm in clouds and d) WV at 7.3 μm in clouds. Blue and red dots indicate AMVs over ocean and over land, respectively.*
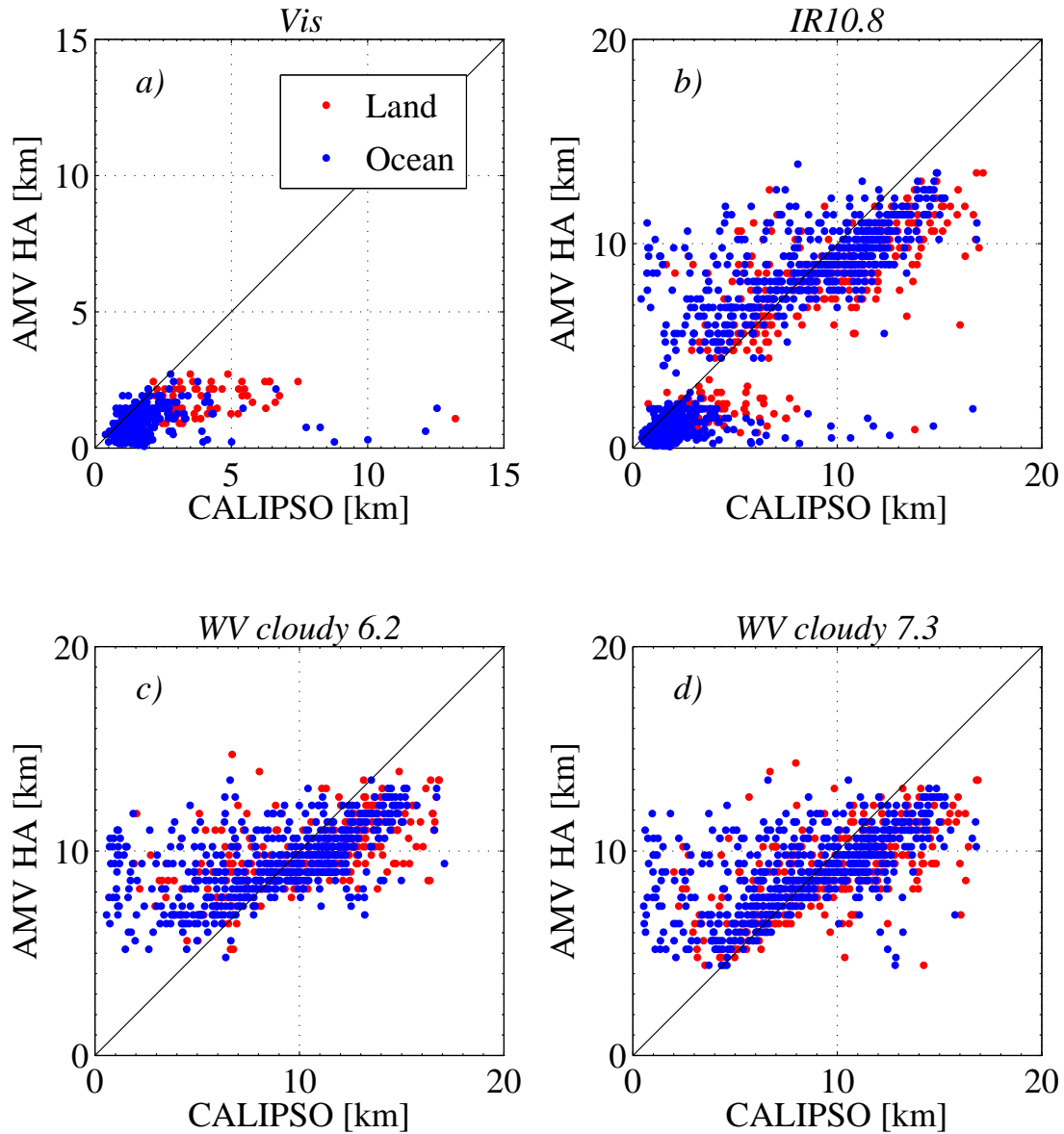
*Figure 3.2: CTH from Meteosat-9 AMV HA vs. CTH from coincident CALIPSO Level-2 Layer Product (averaged to AMV resolution). From all matched cases, only AMVs with QI>50 and more than 35 CALIPSO shots are considered. Each panel contains AMVs generated by a different channel as indicated in the panel's title: a) Visible; b) IR at 10.8 μm; c) WV at 6.2 μm in clouds and d) WV at 7.3 μm in clouds. In all panels, blue and red dots indicate AMVs over ocean and over land, respectively.*
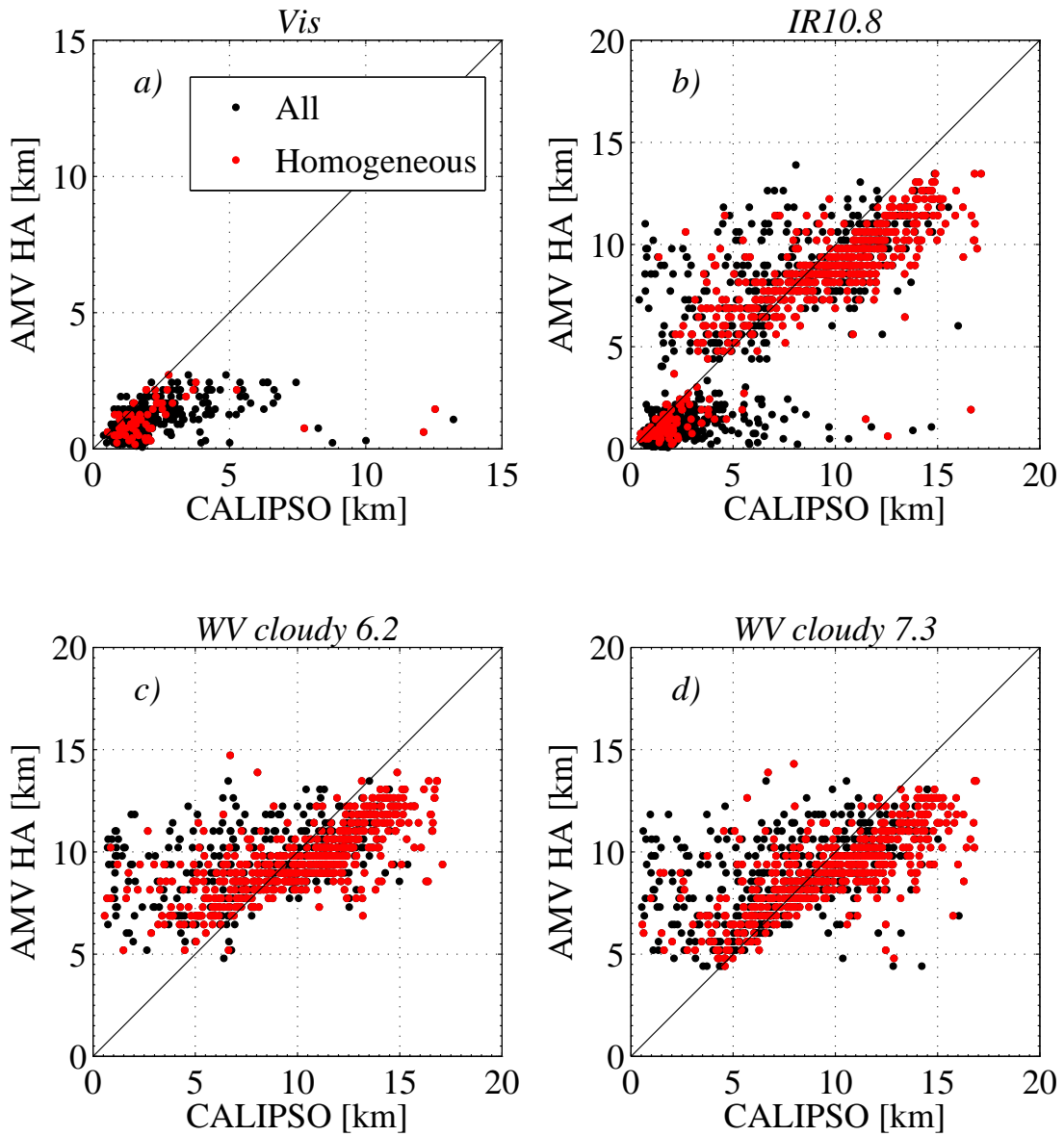
*Figure 3.3: As Figure 3.2, with red dots indicating occurrences where all CALIPSO shots see cloud (homogeneous cases). See text for further explanations.*
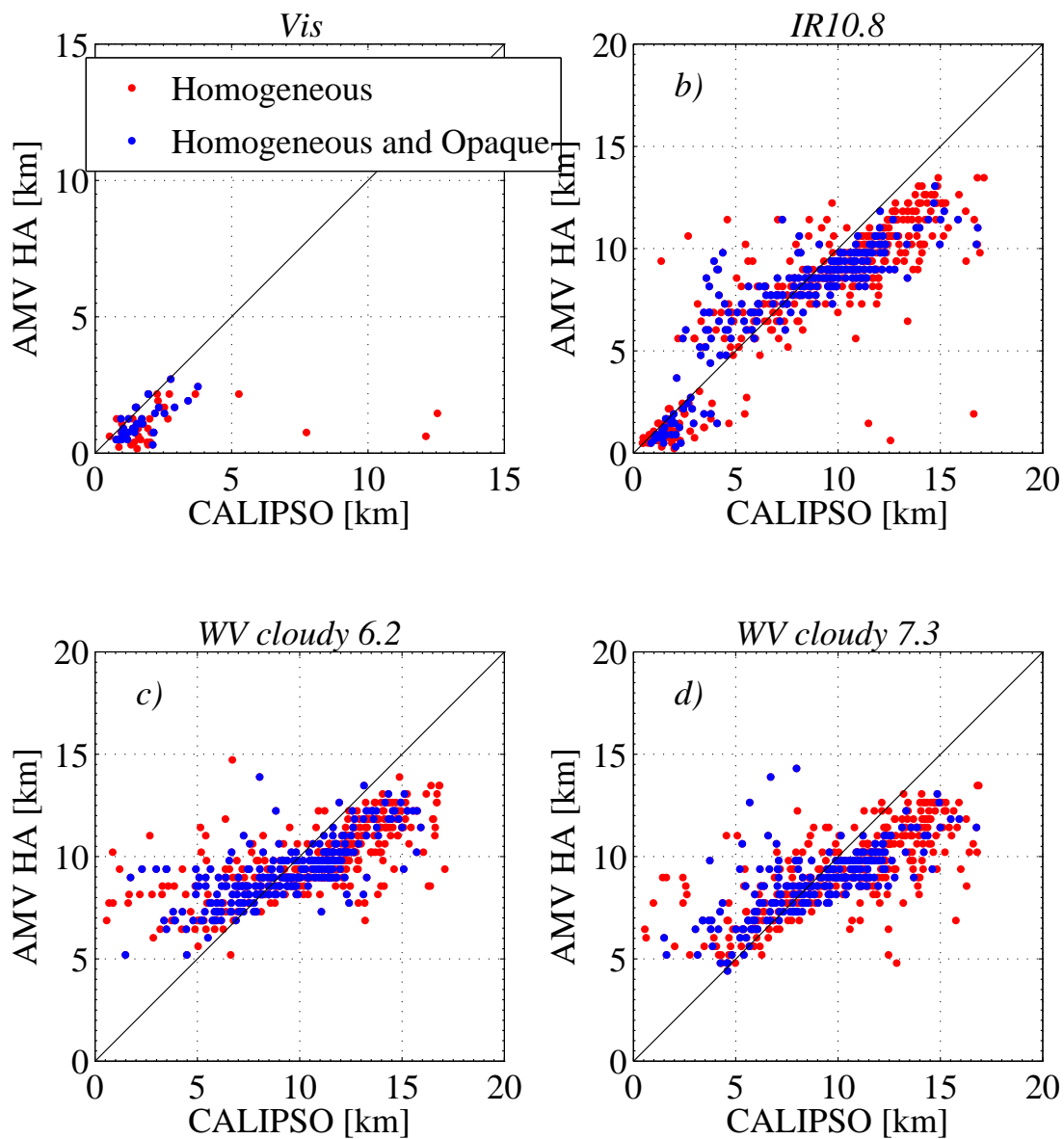
*Figure 3.4: As Figure 3.2, but only occurrences flagged as homogeneous are plotted. Blue dots indicate cases flagged as opaque according to CALIPSO. See text for further explanations.*
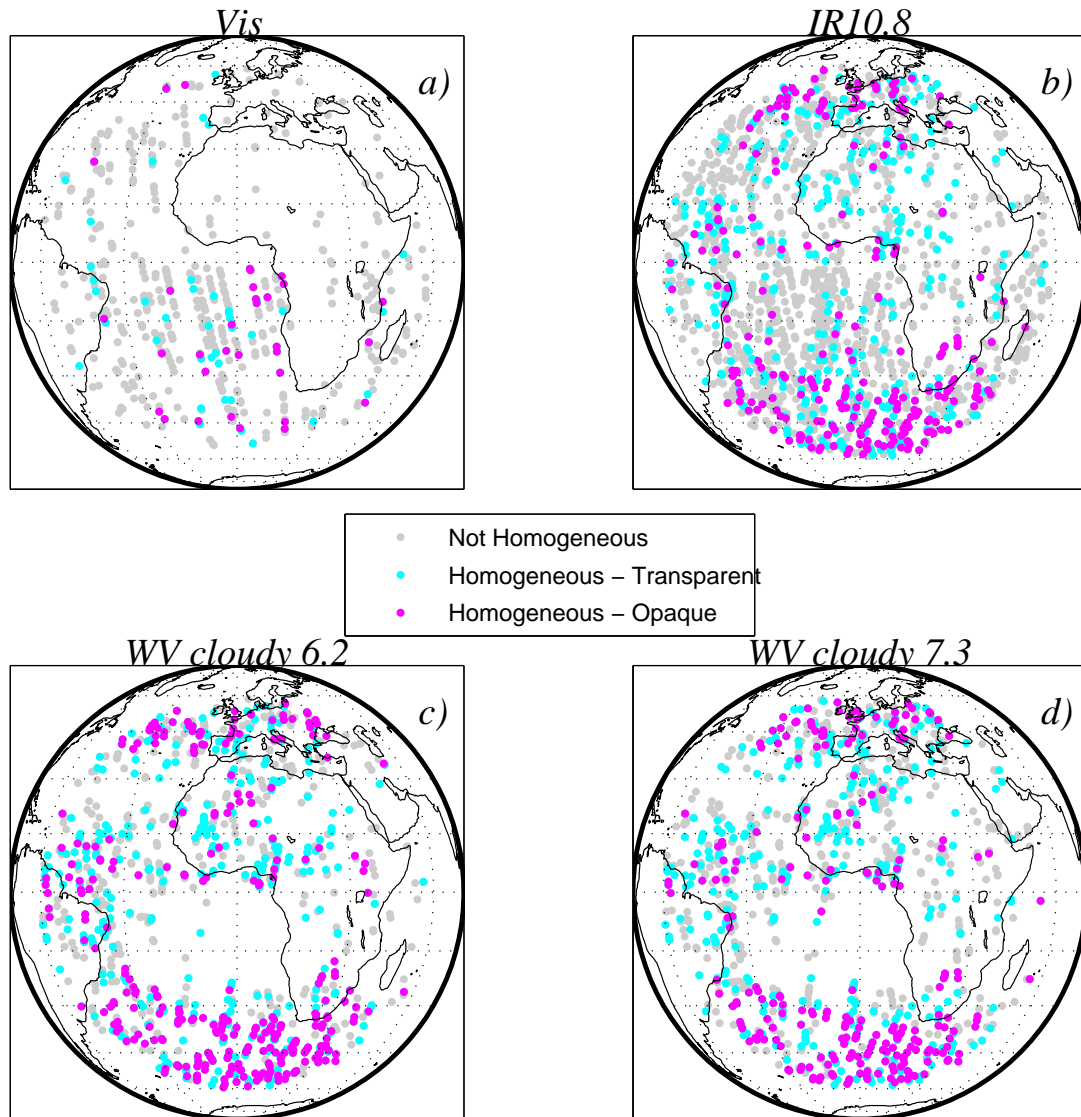
*Figure 3.5: Geographical position of Meteosat-9 AMVs collocated with CALIPSO observations for the studied period. Only AMVs with QI>50 and more than 35 CALIPSO shots are considered. Each panel contains AMVs generated by a different channel as indicated in the panel's title: a) Visible; b) IR at 10.8 μm; c) WV at 6.2 μm in clouds and d) WV at 7.3 μm in clouds. Dots colour code indicates flagging according to CALIPSO shots: grey is for non-homogeneous cases; cyan for homogeneous and transparent cases; magenta for homogeneous and opaque cases.*
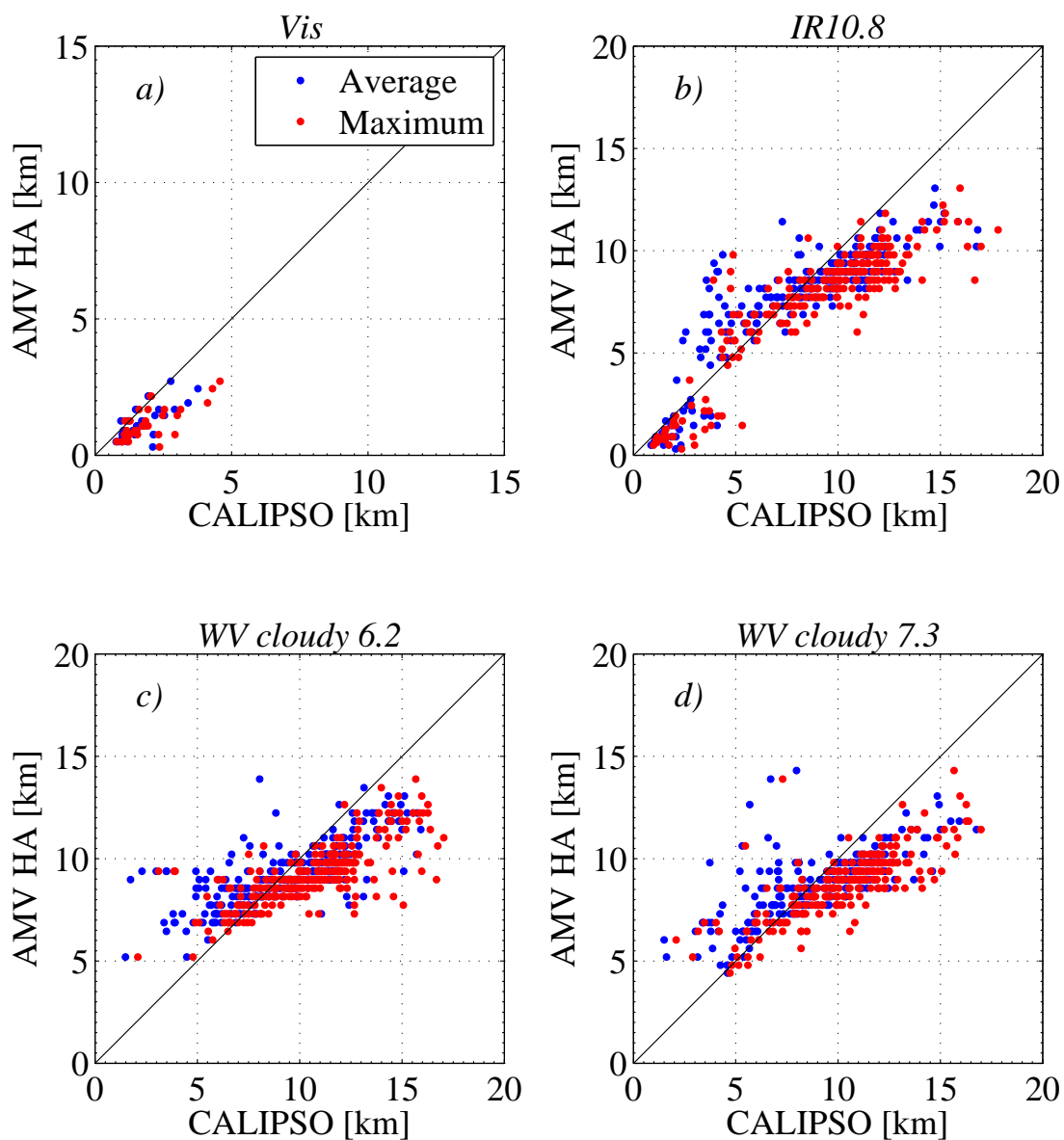
*Figure 3.6: CTH from Meteosat-9 AMV HA vs. CTH from coincident CALIPSO Level-2 Layer Product. Plotted are only the cases with AMV QI>50, more than 35 CALIPSO shots in AMV, flagged as homogeneous and opaque according to CALIPSO. Each panel contains AMVs generated by a different channel as indicated in the panel's title: a) Visible; b) IR at 10.8 um; c) WV at 6.2 um in clouds and d) WV at 7.3 um in clouds. In all panels, blue dots and red dots respectively indicate the use of the mean value or maximum value as CALIPSO CTH.*

# 4    Verification of cloud analysis from MW imagers using CloudSat

Cloud (and rain) affected observations from a number of microwave imagers are currently assimilated at ECMWF. Beside the Special Sensor Microwave/Imager (SSM/I, Hollinger *et al.*, 1990) on board the Defence Meteorological Space Program (DMSP) satellites, there are the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI, Kummerow *et al.*, 1998) and the AMSR-E on the Aqua platform. These instruments all use frequencies that are sensitive, besides to atmospheric water vapour and surface properties, to heavy cloud and precipitation. In particular AMSR-E, being part of the A-Train, allows the direct comparison with the CloudSat Profiling Radar (CPR). Aim of this section is to compare cloud variables from the ECMWF analyses when AMSR-E observations are assimilated against independent cloud products from CloudSat in order to evaluate the performance of assimilation system in cloudy conditions.

## 4.1    1D+4D-Var

Cloud- and precipitation-affected radiances from microwave imagers have been assimilated operationally at ECMWF from June 2005 to March 2009, using a two-step 1D+4D-Var procedure (Geer *et al.*, 2008). The first step consist of a 1D-Var retrieval at each observation point (Bauer *et al.*, 2006a). Only observations over ocean are used given the difficulties in modelling land emissivity at microwaves. The 1D-Var provides the best estimate of the state of a single column of the atmosphere at the observation point, combining the observed radiances with an a-priori temperature and moisture profile taken from the first guess forecast of the 4D-Var assimilation. A total column water vapour (TCWV) amount is calculated from each 1D-Var retrieval of the atmospheric state. These TCWV values are then assimilated as 'pseudo-observations' in the main 4D-Var analysis, causing changes in moisture, cloud, precipitation, and wind fields (Bauer *et al.*, 2006b). Our aim is to compare cloud variables of 1D-Var retrievals performed using AMSR-E observations with corresponding quantities derived from CloudSat.

### 4.1.1    Data

1D-Var retrievals have been extracted from a 1D-Var+4D-Var assimilation experiment based upon Cycle 35R1 (operational from September 2008 to March 2009) of the ECMWF assimilation and forecast model. The (operational) vertical resolution of 91 levels has been used, while a truncation at wavenumber 511 (T511) has been set for the spectral representation of the model. T511 roughly corresponds to a horizontal resolution of 40 km. The cloud and convection scheme within the 1D-Var specifies four hydrometeor categories: non-precipitating in liquid phase (cloud-liquid), non-precipitating in ice phase (cloud-ice), frozen precipitation (snow), and liquid precipitation (rain). Vertically-integrated equivalent water contents (kg/m$^2$) from each of the four hydrometeors and the surface precipitation rate (mm/h) are routinely saved for the first guess and analysis profiles in each 1D-Var retrieval.

AMSR-E carries channels at 6, 10, 19, 23, 37 and 89 GHz (all in double polarization). Channels at lower frequencies (6 and 10 GHz) are not included in the 1D-Var retrieval due to problems with the ocean emissivity model at low frequencies. The channels at 89 GHz are also not used because they are not spatially co-registered with the ones at lower frequency. Only AMSR-E channels 5 to 12, corresponding to frequencies 19 GHz, 23 GHz and 37 GHz, are then assimilated. These frequencies are sensitive to the emission from liquid cloud and precipitation, and only weakly to the frozen portion of clouds. Therefore, it makes sense to focus the comparison with CloudSat on liquid variables (cloud and precipitation).

Among the several standard products from the CloudSat mission, the 2B-Cloud Water Content - Radar Only product (2B-CWC-RO) provides vertical profiles of liquid or ice content converting radar reflectivities into

mass amount. The distinction between the phases is based on the height of the freezing level, obtained from co-located ECMWF temperature analysis. From CPR observations, also surface precipitation is derived and it is archived as the 2C-PRECIP-COLUMN CloudSat standard product. The precipitation retrieval algorithm makes use of the radar reflectivity closer to the Earth surface and the auxiliary information on path integrated attenuation (PIA) to determine rain occurrence and intensity. Since PIA must be determined from surface reflection characteristics, only oceanic and inland water scenes are considered.

### 4.1.2    Matching of AMSR-E with CloudSat

For the purpose of intercomparison, a procedure is required to collocate AMSR-E and CloudSat products. CloudSat Level-2 products are provided at the native resolution of the instrument, i.e. 1.3 km, which is the size of the instantaneous footprint of the radar. As resolution for the 1D-Var analysis, we have used the one specified for the cloud and convection scheme used within the 1D-Var instead of the one of the AMSR-E. This choice is due to the fact that the information about the AMSR-E spatial resolution is not used in the 1D-Var, and that the amount of cloud and precipitation in the 1D-Var depends on the resolution assigned to the model. As said above, in our experiment the model resolution has been set to T511, approximately corresponding to 40 km. Therefore, the matching has been achieved considering all CloudSat shots falling within ±20 km distance from each AMSR-E observation location and averaging them. The particular A-Train configuration guarantees that observations are (almost) temporally coincident.

The matched data sets obtained in this way have been subject to some quality control measures. Only cases for which the 1D-Var retrieval appeared successful (i.e. a successful minimisation) have been included and, also, in order to reduce the noisiness of the data, at least 20 CloudSat shots (out of a maximum possible of about 40) were required to fall into the region over which the average was taken.

An additional source of uncertainty results from the small footprint of the CloudSat radar. The corresponding sampling issue was not taken into account in this study, but the reader can refer to the second chapter of the QuARL report on WP 1000 for a statistical way to quantify the associated uncertainty. The collocation between AMSR-E 1D-Var and CloudSat has been done for the month of September 2007.

Fig. 4.1 (top panel) shows the geographical distribution of the number of co-located AMSR-E/CloudSat observations obtained using the matching procedure described above. The map gives the number of samples as counted within boxes with a horizontal extent of 5 x 7.5 (latitude x longitude). The size of the box represents a trade-off between the requirements of having enough samples in each box and that of having enough boxes for displaying geographical patters. Only cases where the (averaged) CloudSat 2B CWC RO provides a non-zero value of cloud liquid have been considered. This condition excludes situations where the cloud falls within the radiometer field of view but outside the radar track and therefore the radiometer sees some cloud but the radar does not.

The bottom panel of Fig. 4.1 shows the fractional amount of matched cases (within each grid box) which are cloudy according to CloudSat. This quantity can be seen as an indication of monthly cloud occurrence at a scale of 40 km according to CloudSat. The highest occurrences are in the Southern Hemisphere, in particular in the South East Pacific, South Atlantic and South Indian Ocean.

### 4.1.3    Intercomparison: mean differences

Fig. 4.2 shows mean values over the 5 by 7.5 grid boxes of precipitation (top panel), cloud liquid (middle panel) and ice (bottom panel) from CloudSat products averaged and collocated with AMSR-E. Only cases that have passed the checks on the quality of the 1D-Var retrievals and the number of matched CloudSat shots mentioned before have been considered.

These quantities can be directly compared with the corresponding grid-box averages of collocated first
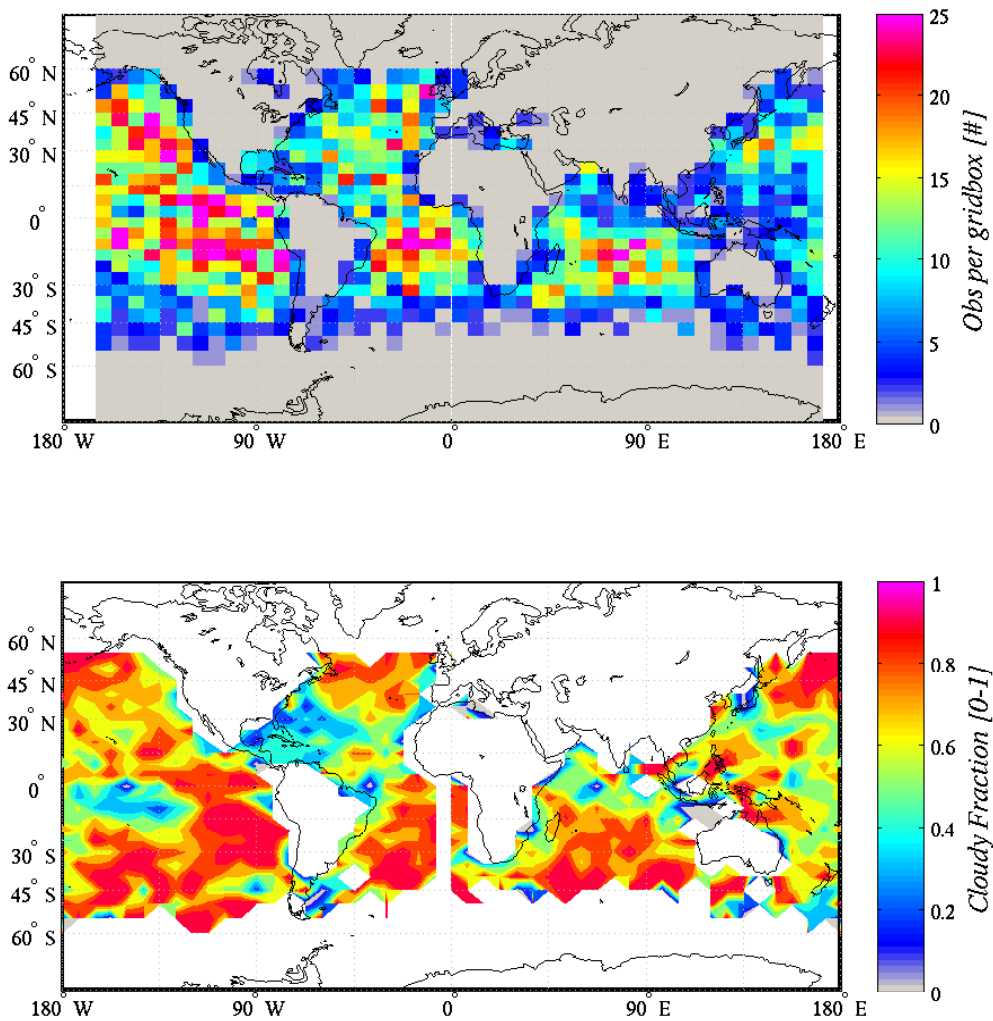
*Figure 4.1: Top panel: Number of AMSR-E 1D-Var occurrences matched with CloudSat contained in 5 by 7.5 boxes on a global grid. Bottom panel: Fraction of the matched occurrences being cloudy according to CloudSat.*

guess and 1D-Var analysis from AMSR-E. In Fig. 4.3, differences between 1D-Var first guess and CloudSat estimates are given for surface precipitation (top panel) and cloud liquid (bottom panel). For precipitation, we note that first guess values are almost everywhere larger than the ones from CloudSat. Differences are particularly large in the Tropics and in the Pacific Warm Pool. This behaviour is likely to be an artifact from the almost complete lack of precipitation that 2C-PRECIP-COLUMN exhibits in the Tropics which is most likely linked to the attenuation of the radar signal (radar signals at 94 GHz are usually completely attenuated if they hit a cloud whose rain rate is larger than few mm/h). Another possible reason for the 1D-Var values being larger is the high spatial variability of precipitation fields. Due to this variability, CloudSat can miss some of the rain because it is out of its track. However, in spite of these issues, this intercomparison is a valuable test to stress strengths and weaknesses of the precipitation retrieval from a passive instrument and from an active one.

When considering cloud liquid differences (bottom panel of Fig. 4.3) we note that CloudSat values are generally greater than the 1D-Var in the Southern Hemisphere. In particular, there are patterns of large differences between 0 and 45 South in the East Pacific, in the Atlantic, and over the Indian Ocean. Comparing with Fig. 4.2, we note that they correspond to large values in CloudSat estimates, which are not present in the 1D-Var. The map of cloud liquid differences shows that there are also areas where the 1D-Var first guess

is much larger than CloudSat estimates. Most of them have the same locations where the 1D-Var first guess of precipitation is much larger. They can be linked to the attenuation affecting CloudSat measurements in precipitation. In addition, there are also some areas (latitudes around 45 in the Northern Hemisphere) where the 1D-Var first guess is much larger than the values obtained from the CloudSat products.

Fig. 4.4 shows the differences between 1D-Var analysis and CloudSat estimates. Comparing the precipitation differences (top panel) with the corresponding values for the first guess shown in Fig. 4.3, one finds that the size of these differences is strongly reduced in large parts of the mid latitudes (in both hemispheres). This indicates that the 1D-Var procedure successfully decreases the error of the first guess and brings the analysis closer to CloudSat. In the Tropics, however, the big, generally positive differences are not changed considerably through the 1D-Var analysis, but a slight improvement in the Western Pacific and a degradation in the Eastern Atlantic are noticeable.

When considering cloud liquid (bottom panel) we note a remarkable improvement in the Northern Hemisphere: the pattern of (positive and negative) differences stays, but with a much-reduced amount. On the contrary, in the Southern Hemisphere there is not always an evident reduction.

We note an improvement in the Indian Ocean and the Central Pacific but the analysis seems to increase the differences everywhere else. We believe that the lack of agreement between model first guess and CloudSat cannot be corrected by AMSR-E 1D-Var because of the weak sensitivity of the radiometric channels (especially the ones at 19 GHz) to the type of clouds in these regions. The map of mean cloud-top height (Fig. 4.5) shows that they are mainly shallow clouds (cloud-top below 1km) composed of liquid phase only (as shown in Fig. 4.2, bottom panel). The analysis of the CloudSat measurements evidenced that they produce reflectivities smaller than -15 dBZ, suggesting that these clouds do not carry any drizzle or light precipitation. A similar result will be shown in the next section, when comparing CloudSat with the 'all-sky' 4D-Var assimilation of AMSR-E observations.
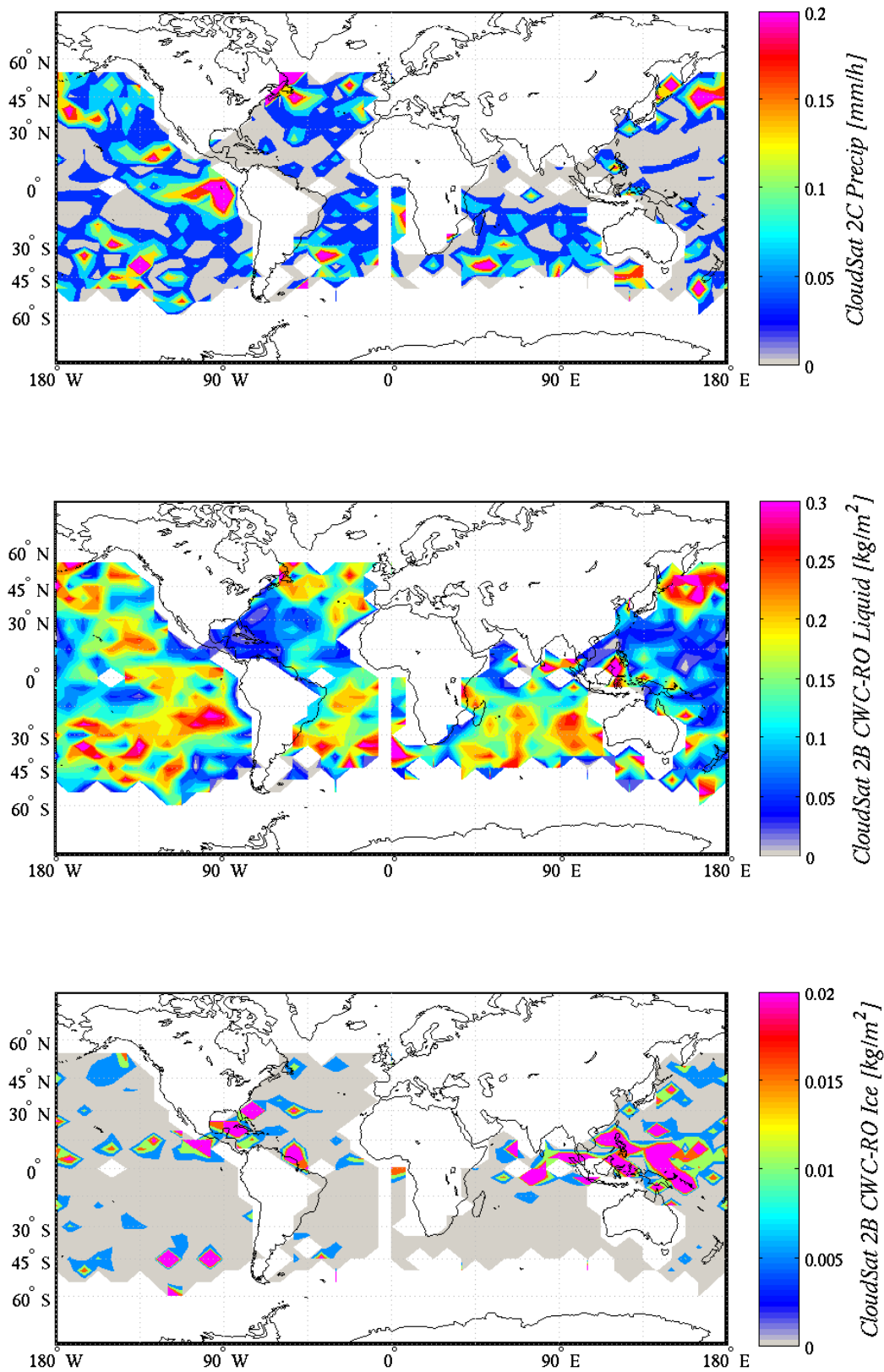
*Figure 4.2: Map of mean CloudSat estimates for the occurrences matched with AMSR-E 1D-Var. From top to bottom: 2C-PRECIP-COLUMN, 2B-CWC-RO Liquid, and 2B-CWC-RO Ice.*
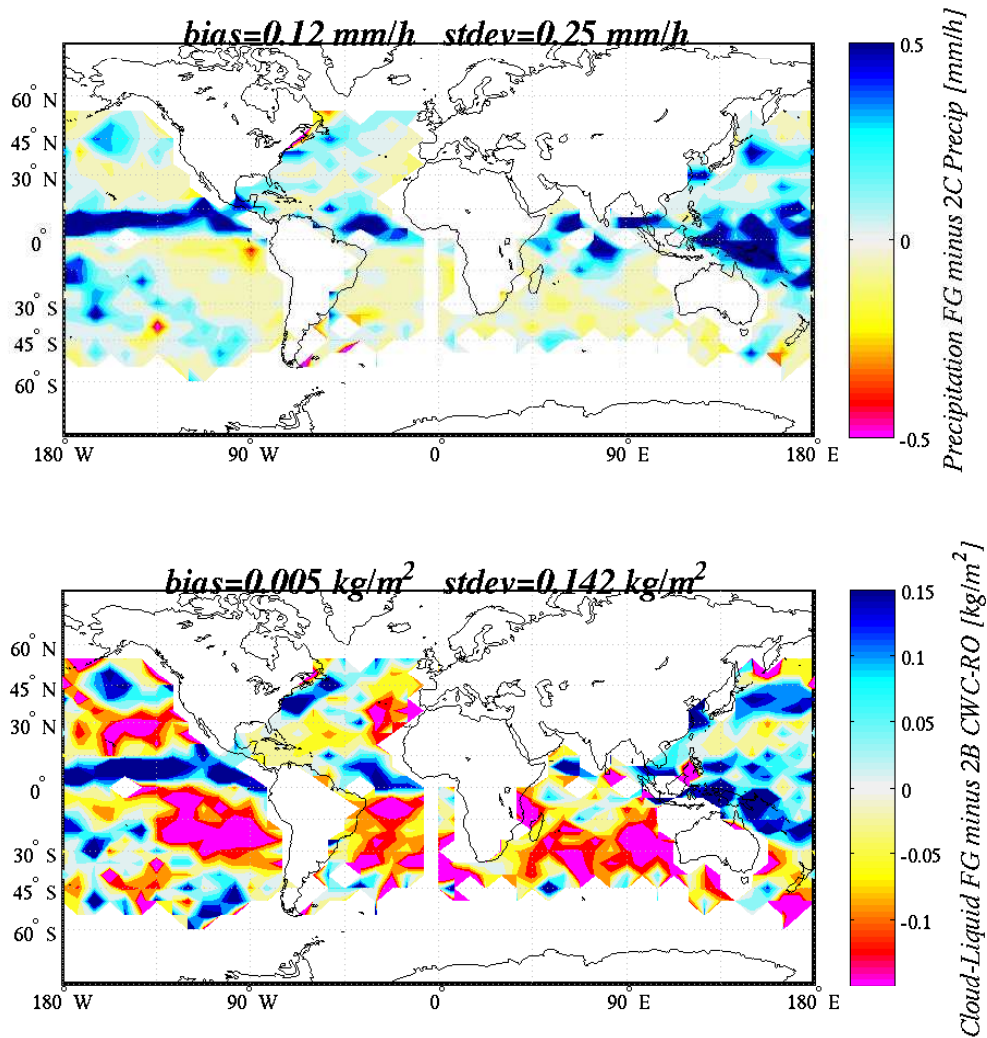
Figure 4.3: Map of mean differences between 1D-Var first guess and CloudSat estimate. Top panel: surface precipitation. Bottom Panel: cloud liquid.
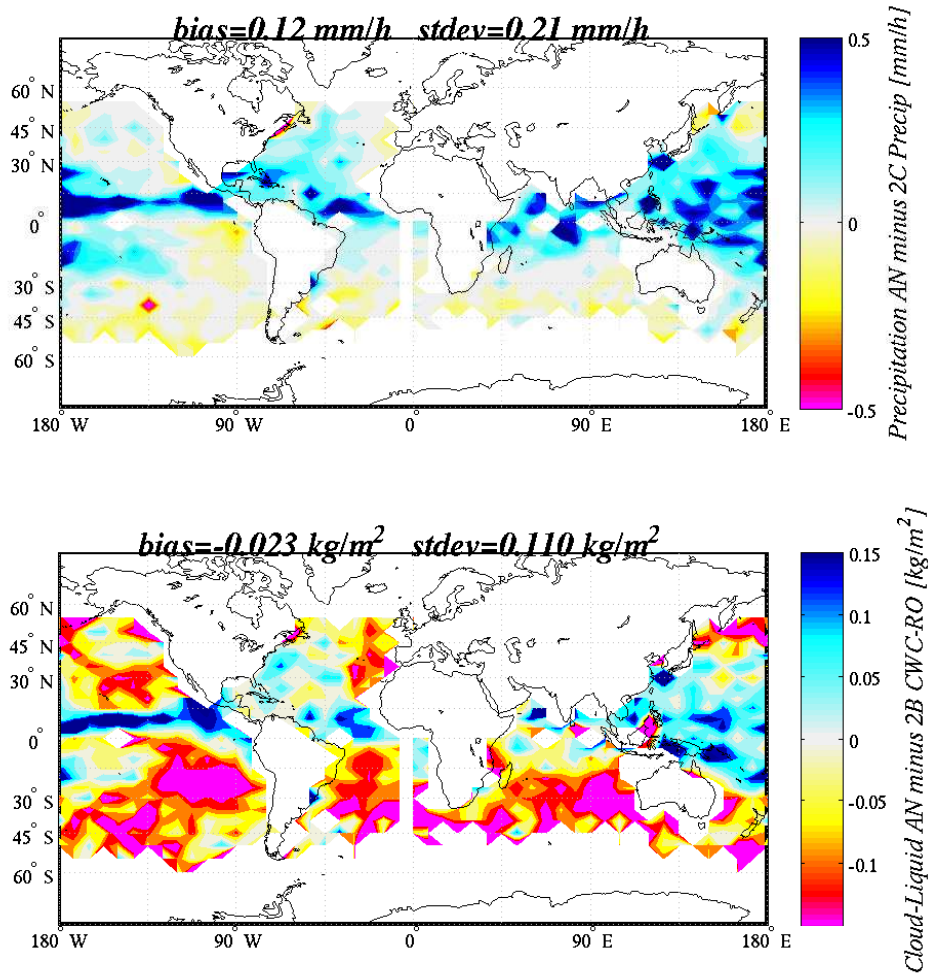
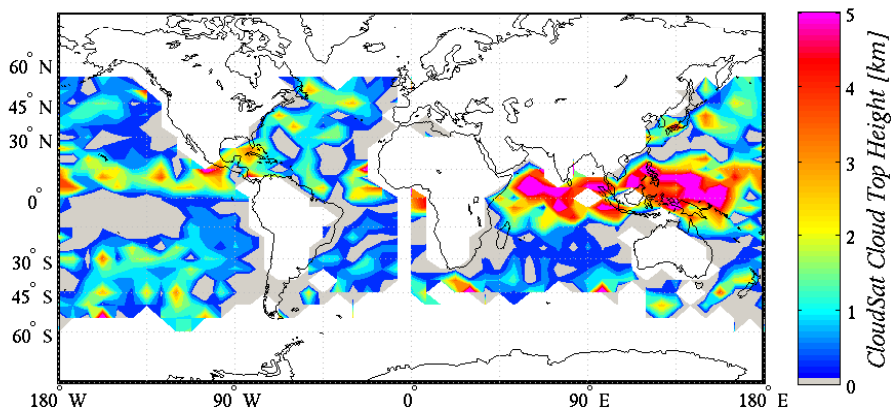*Figure 4.4: As Fig. 4.3, but for 1D-Var analysis.*



*Figure 4.5: Map of mean cloud-top height from CloudSat for the occurrences matched with AMSR-E 1D-Var.*

### 4.1.4   Intercomparison: scatter plots

In the previous paragraph the intercomparison was done in terms of grid-box monthly averages. We now want to directly compare AMSR-E 1D-Var analyses with the matched CloudSat estimates. The Cloud-Sat Clouds and Precipitation Radar (CPR) provides a vertical profiling, but lacks any scanning capability. Therefore, in spite of the along track average, the 1D-Var volumetric estimates must be verified against radar estimates consisting of along-track lines about 1 km wide. The representativeness of the mean value along this 'curtain' as the one within the whole radiometer FOV is then not always guaranteed. In order to alleviate possible inconsistencies between the two means, we introduced a further screening condition. Similarly to the approach followed for the verification with CALIPSO, CloudSat observations matched with AMSR-E which do not see 'broken scenes' (i.e. most of the radar profiles do not contain any cloud) have been removed. Although this condition does not always guarantee a homogeneous scene it excludes a large part of the situations for which inhomogeneity effects increase the uncertainty.

Fig. 4.6 compares the precipitation from 1D-Var against the 2C-PRECIP-COLUMN CloudSat product. The only considered situations are the ones where both 1D-Var and CloudSat have precipitation. In addition, only 'homogeneous' cases have been included, i.e. those (among the all matched in the month of our experiment) where less than 10 shots out of 20 (or more) used for evaluating the mean have some precipitation. The spread in these scatter plots is significant, both in the 1D-Var first guess (left panel) and analysis (right panel). However, when comparing the two panels we see the positive effect of AMSR-E observations: the dispersion (shown as ±3-sigma with red lines) is reduced and the correlation (shown by the linear regression with the blue line) is improved. In order to highlight the behaviour for smallest values, the same results are again displayed in Fig. 4.7 using logarithmic scales. The positive impact of the 1D-Var analysis of AMSR-E observations is evident also in this figure.

A similar comparison has been done for cloud liquid and it is shown in Fig. 4.8. Again, only 'homogeneous' situations have been considered, i.e. cases where less than 10 shots contain some cloud liquid were removed. We note that only few cases have the cloud liquid from model first guess larger than CloudSat, while for the rest the first guess is usually smaller than CloudSat 2B-CWC-RO, with values most of the times lower than 0.1 kg/m$^2$ (left panel). The result of the analysis (right panel) shows that the 1D-Var is able to move the model first guess closer to CloudSat estimates as it was for the precipitation. All points are closer to the diagonal, including the cluster of points where the first guess was lower than 0.1 kg/m$^2$. As evident in Fig. 4.9, 1D-Var analyses are closer to CloudSat estimates also for the lowest values of cloud liquid content. Table 4.1 provides some statistics of the 1D-Var vs CloudSat intercomparison. Consistently with scatterplots, the standard deviation of the analysis-minus-CloudSat differences is always lower than FG-minus-CloudSat differences for both precipitation and cloud liquid. Overall, the 1D-Var produces estimates closer to CloudSat than the FG in 70%.

| | Nr. of collocations | 1D-Var - CloudSat: FG (AN) mean | 1D-Var - CloudSat: FG (AN) st. dev. | 1D-Var AN closer to CloudSat than FG |
|---|---|---|---|---|
| Precipitation | 482 | -0.07 (-0.08) [mm/h] | 0.45 (0.33) [mm/h] | 70% |
| Cloud Liquid | 2078 | -0.106 (-0.136) [kg/m$^2$] | 0.301 (0.202) [kg/m$^2$] | 70% |

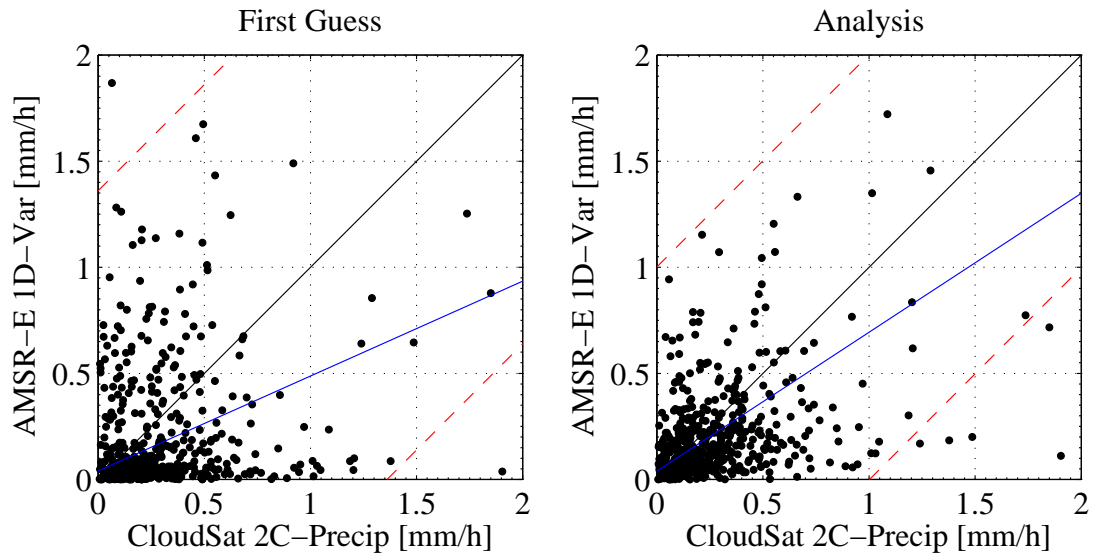*Table 4.1:  Comparison of 1D-Var statistics against CloudSat.*

*Figure 4.6: 1D-Var surface precipitation vs. CloudSat 2C-PRECIP-COLUMN. Left panel: First guess. Right panel: AMSR-E analysis.*
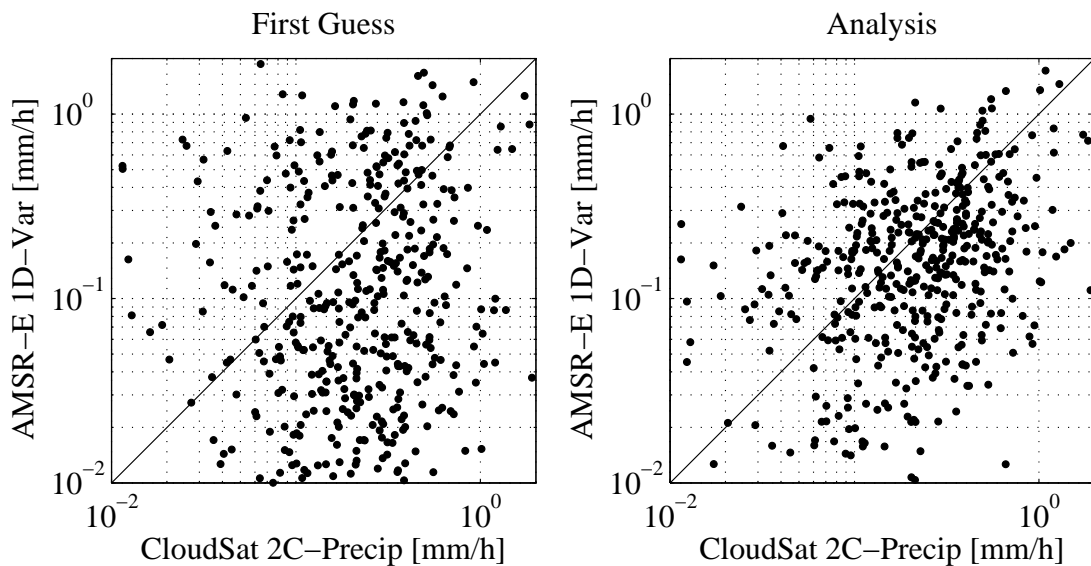


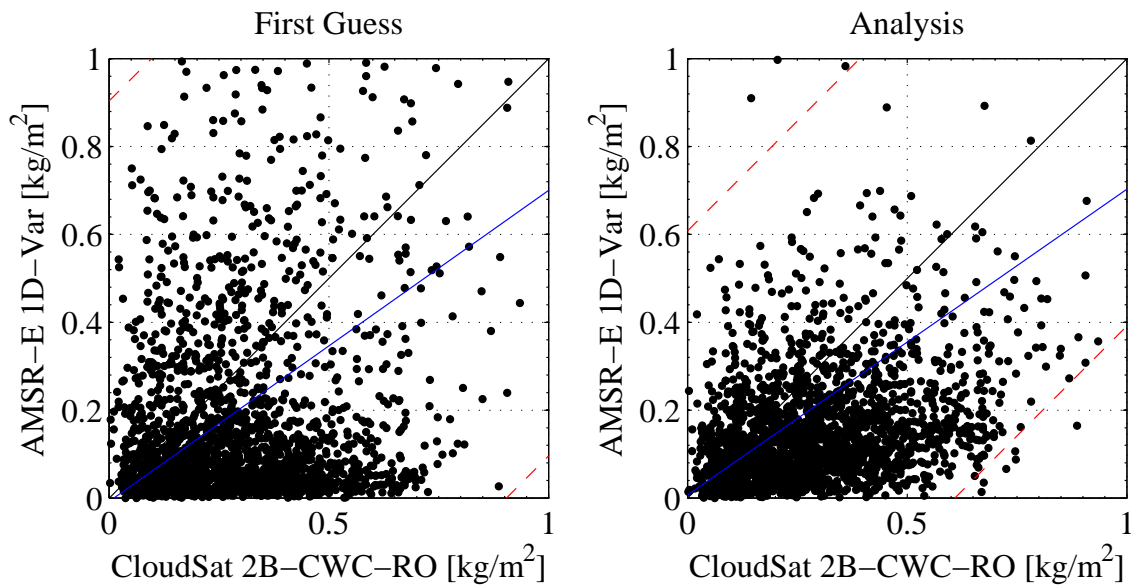*Figure 4.7: As Fig. 4.6, but using logarithmic axes.*

*Figure 4.8: 1D-Var cloud liquid vs. CloudSat 2B-CWC-RO Liquid. Left panel: First guess. Right panel: AMSR-E analysis.*
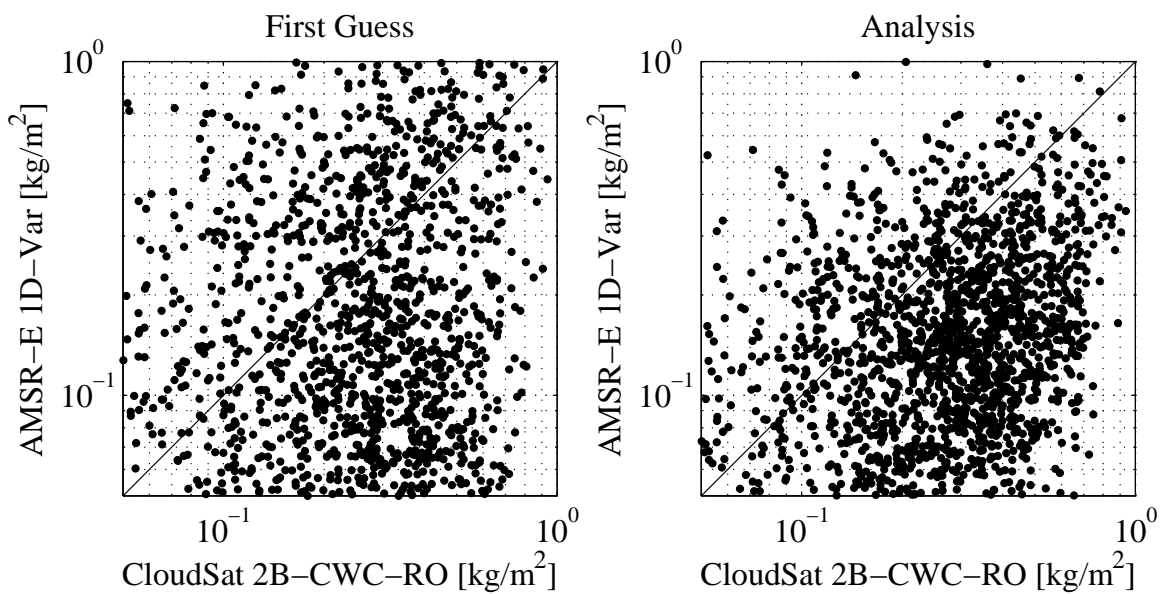


*Figure 4.9: As Fig.4.8, but using logarithmic axes.*

## 4.2 All-sky 4D-Var

Geer *et al.* (2008) noted a number of limitations of the 1D+4D-Var approach. First, the TCWV pseudo-observation was not bringing that much new information because most of the 4D-Var analysis humidity increment is inferred indirectly from other observations, such as temperature-sounding radiances. Second, the separate treatment of clear-sky and cloud/rain-affected observations created an unbalanced sampling. Cases with clear observations but a cloudy first guess went through the clear sky route, using clear-sky radiative transfer in the observation operator. This meant that (a) there was quite a large sampling bias in the 1D+4D-Var stream and (b) an opportunity was being missed to correct the erroneous first guess (FG) cloudy conditions when clear skies were observed. An 'all-sky' approach was proposed, where the treatment of clear and cloud-affected cases are put together in a single data stream with a single observation operator. The all-sky treatment unifies clear-sky and cloud-affected observations and has a single observation operator. All-sky assimilation of microwave imager radiances went operational with Cycle 35r2 in March 2009 (Bauer *et al.*, 2010; Geer *et al.*, 2010). In this new approach, radiance observations from SSM/I and AMSR-E are currently assimilated, either clear, cloudy or rainy. The A-Train configuration can be exploited in this case using CloudSat to investigate the impact of AMSR-E on the all-sky assimilation system.

### 4.2.1 Data and matching

The approach to compare all-sky 4D-Var with CloudSat closely resembles the one followed by the 1D-Var. Also in this case, for the same reasons explained for the 1D-Var, only over-ocean AMSR-E observations and only channels at 19 GHz, 23 GHz and 37 GHz are used in the assimilation. Estimates of cloud liquid from 2B-CWC-RO and precipitation from 2C-PRECIP-COLUMN, were matched with AMSR-E observations in the same manner done for the 1D-Var. We considered only the cases where AMSR-E observations were successfully assimilated (indicated by a reduction of the FG departure after the analysis).

The ECMWF 4D-Var uses an incremental formulation for the minimisation (Courtier *et al.*, 1994), based around a non-linear forecast 'outer-loop' update and a linearized 'inner loop' minimisation. The outer loop steps are run with the atmospheric model using a spectral resolution of T799 (corresponding to 25 km). There are three runs of the inner-loop, all at a reduced resolution, the first at T95, the second at T159, and the third at T255. Given this assimilation strategy, we have used the horizontal resolution of the last inner-loop (i.e. T255, corresponding to 70 km) as the one of the 4D-Var cloud variables.

Our investigation was based on an all-sky assimilation experiment performed for the month of July 2009. Fig. 4.10 (top panel) shows the number of matched samples that fall in 5 by 7.5 boxes on a uniform grid. In spite of using one month of data as for the 1D-Var, in this case there are more samples. This is mainly due to a larger number of AMSR-E observations being assimilated in the 4D-Var compared to the 1D-Var. We also note in the Southern Hemisphere the presence of samples at higher latitudes which had not been included in the 1D Var computations. This is due to different screening conditions on 'cold sectors' which for 4D-Var is more relaxed than the one used in the 1D-Var system. For the reasons explained in subsection 4.1.3, only the cases where CloudSat observations contain some cloud liquid have been considered. The bottom panel of Fig. 4.10 shows the fraction of samples in each box that satisfy this check.

Grid box mean values of CloudSat products are plotted in Fig. 4.11: surface precipitation (top panel), cloud liquid (middle panel) and cloud ice (bottom panel). Fig. 4.12 shows the corresponding values of AMSR-E mean observation at 19 GHz (top panel) and 36.5 GHz (bottom panel). Instead of the measured brightness temperatures ($T_B$), we have used a derived quantity, the normalised polarization difference (P) as defined by Petty (1994):

$$P = \frac{T_B^V - T_B^H}{T_{B,clr}^V - T_{B,clr}^H}$$

Where $T_B^V$ and $T_B^H$ indicate brightness temperatures with vertical and horizontal polarization, respectively. $T_{B,clr}^V$ and $T_{B,clr}^H$ indicate the corresponding brightness temperatures (of the same profile) without considering clouds and precipitation. In our case, the model clear sky profile which is closest to the first guess has been used for estimating $T_{B,clr}^{V,H}$.

Since emissions from the ocean surface are highly polarized, but atmospheric absorption is in general unpolarised, $T_B^V - T_B^H$ is a measure of the atmospheric opacity, with a completely opaque atmosphere giving no polarization difference. Hence, P generally varies between 0, which represents a profile with opaque cloud, and 1, which represents clear sky. P is a quantity more suitable for our purpose since it is proportional to the square of the optical depth of clouds and precipitation only, and does not include contributions from the surface and atmospheric gases. At microwave frequencies the ocean appears as a cold background where larger liquid drops are mainly responsible for $T_B$ enhancements at 19 GHz, and drops of smaller size (non-precipitating) do the same for 37 GHz $T_B$s. This effect is clearly shown in these plots.

### 4.2.2  Results

In the context of the 4D-Var, the analysis depends on all observations fed into the system within the 12-hour assimilation window. It is therefore not easy to disentangle the impact on the analysis of AMSR-E observations from the one of other observations. However, a comparison of AMSR-E first guess (analysis) departures with corresponding differences between model first guess (analysis) and CloudSat (in terms of water content) can still provide useful insights.

Fig. 4.13 shows the AMSR-E mean first guess departures at 19 GHz (top panel) and 36.5 GHz (bottom panel) in terms of P. The departures are defined as observation-minus-first guess, positive (negative) values are for cases where AMSR-E observations indicate that the amount of cloud/precipitation is higher (lower) than the model first guess. The map relative to 19 GHz shows a positive departure widespread on the globe: in the Northern Pacific, Indian Ocean, along the US East Coast and, only slightly in the Southern Atlantic. Given the strong correlation between 19 GHz radiances and rain amount, a positive departure is an indication that in the first guess there is too much precipitation. But it is important to mention that the same effect can also be due to partial beam filling, especially in the Tropics, where scales of precipitation can be smaller than the radiometer FOV. The same plot shows also two regions where the 19 GHz P first guess is smaller than observations, one in the Northern Atlantic and one in Southern Pacific (south of Australia/New Zealand). In these regions, the model first guess is likely to be drier than observations.

Mean AMSR-E 37 GHz P departures (bottom panel of Fig. 4.13) are almost everywhere negative. Situations where the first guess is smaller than the observations must be interpreted as a lack of cloud liquid, well correlated to the radiation emitted at this frequency. There are few localized spots where the opposite occurs. These areas are coincident with the strongest departures of the same sign in the map of 19 GHz, therefore indicating that the model may have too much cloud and precipitation in these locations. It is noteworthy, that the negative departures areas at 19 GHz mentioned before are also well correlated with the strongest negative ones at 37 GHz in the Northern Atlantic and in Southern Pacific.

Differences between model first guess and CloudSat are shown in Fig. 4.14. The map of precipitation differences (top panel) shows large areas in the Tropics where the first guess is larger than CloudSat. This is similar to what was found in the case of 1D-Var: it can be attributed again to CloudSat missing the intense precipitation typical in these regions, due to signal attenuation. The same explanation can be used for the 'blue' spots in the most Southern areas. The presence of rain in these areas is confirmed by the values of P at 19 GHz smaller than unity in Fig. 4.12.

Top panel of Fig. 4.14 shows that there are regions where the precipitation from the model first guess is lower than CloudSat estimates. This occurs everywhere along a belt around -45 and in a large area in Northern Atlantic. For cloud liquid, the corresponding differences are shown in the bottom panel of Fig. 4.14. Overall, CloudSat gives values higher than the model in the same regions where the precipitation is higher. This agreement in the sign of the two CloudSat products can be seen as an indication of robustness of our intercomparison since these products are independently derived. In the Northern Atlantic and in the South East Pacific we note two large areas where CloudSat puts large amounts of cloud liquid, much more than the model first guess. This result agrees with what was found in the 1D-Var case.

Patterns of AMSR-E first guess departures (Fig. 4.13) and FG-minus-CloudSat differences (Fig. 4.14) can be also qualitatively compared against each other each other having in mind that 19 GHz $T_B$s are well correlated with rain and that 37 GHz $T_B$s are well correlated with cloud liquid.

When considering the cloud liquid differences, we see a quite remarkable agreement with 37 GHz departures except for few areas (bottom panels of Fig. 4.13 and Fig. 4.14). The most evident discrepancy is in the South East Pacific, where a large positive cloud liquid bias does not have a correspondence in the AMSR-E observations. As in the 1D-Var intercomparison (where a similar condition was found), this can be explained by the particular structure of these clouds. Low-level clouds (stratocumulus) are often present in this region. AMSR-E has little sensitivity to them because of their small optical depth. This is confirmed by the mean P value (shown in Fig. 4.12), which is quite high at 37 GHz and very close to 1 at 19 GHz.

Maps of FG-minus-CloudSat precipitation difference and AMSR-E 19 GHz first guess departures (top panels in Fig. 4.14 and Fig. 4.13) do not show a very similar pattern. Differences between the two patterns are difficult to interpret because CloudSat misses moderate and intense precipitation regimes. Cases where the precipitation from the model first guess is larger than CloudSat can be the just a result of this effect. On the contrary, cases where CloudSat puts more precipitation than the model can be considered as genuine. One of these situations can be seen in Northern Atlantic (Fig. 4.14, top panel). In the same region, AMSR-E 19 GHz first guess departures (Fig. 4.13, top panel) also exhibit a negative mean, which indicates (consistently with CloudSat) a lack of precipitation in the model first guess.

The comparison between model and observations can be repeated considering the 4D-Var analysis instead of the model first guess. Mean differences between model analysis and CloudSat are given in Fig. 4.16. Generally, both, precipitation (top panel) and cloud liquid (bottom panel) fields are very similar to the corresponding values for the first guess plotted in Fig. 4.14. Noticeable differences between these figures are found only in few geographical regions and can have either sign (i.e., the seize of the residues can be increased or decreased by the 4D-Var analysis).

Similarly, the AMSR-E analysis residuals (in Fig. 4.15) are very similar to the AMSR-E first guess departures in Fig. 4.13. Again, differences (increases or decreases) are present only in few areas. The information about clouds contained in AMSR-E observation seems only weakly used by the 4D-Var analysis. This result is in line with what was found by Geer *et al.* (2010).

## 4.3 Summary

The investigation showed that CloudSat provides useful information about the impact of clouds- and precipitation-affected passive MW observations in the assimilation system. The 1D-Var verification demonstrated that CloudSat products can be used to validate retrieval of clouds/precipitation retrieval from passive MW instruments. However, a screening is needed to consider only CloudSat measurements representative within the radiometer FOV.

The intercomparison of CloudSat products with 4D-Var cloud variables showed a qualitative agreement with
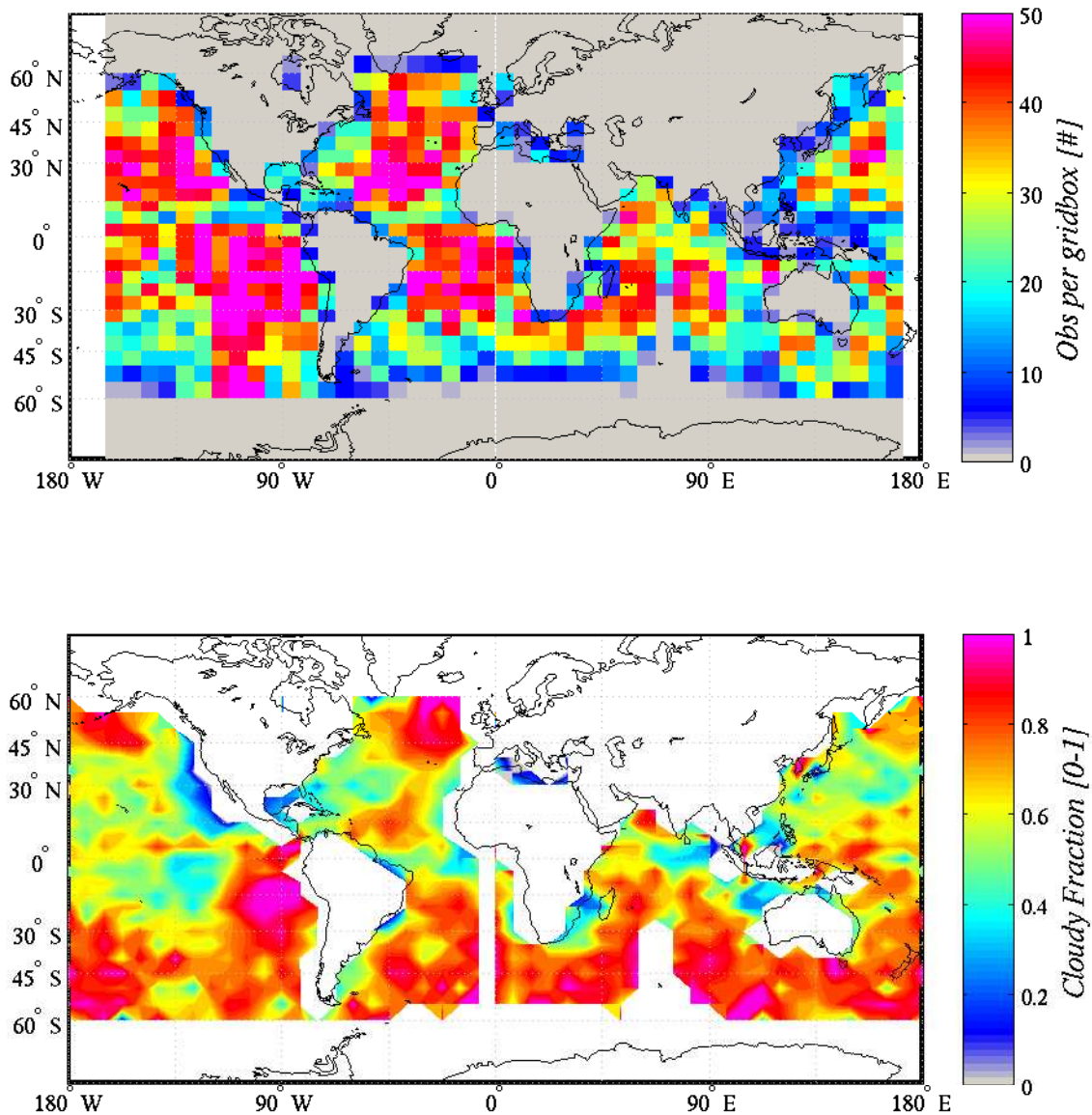
*Figure 4.10: Top panel: Number of AMSR-E 'all-sky' 4D-Var occurrences matched with CloudSat contained in 5 by 7.5 boxes on a global grid. Bottom panel: Fraction of the matched occurrences being cloudy according to CloudSat.*

AMSR-E first guess departures. These results show that CloudSat has the potential to be used as a diagnostic tool for the monitoring of the cloud- and rain-affected radiances assimilation.
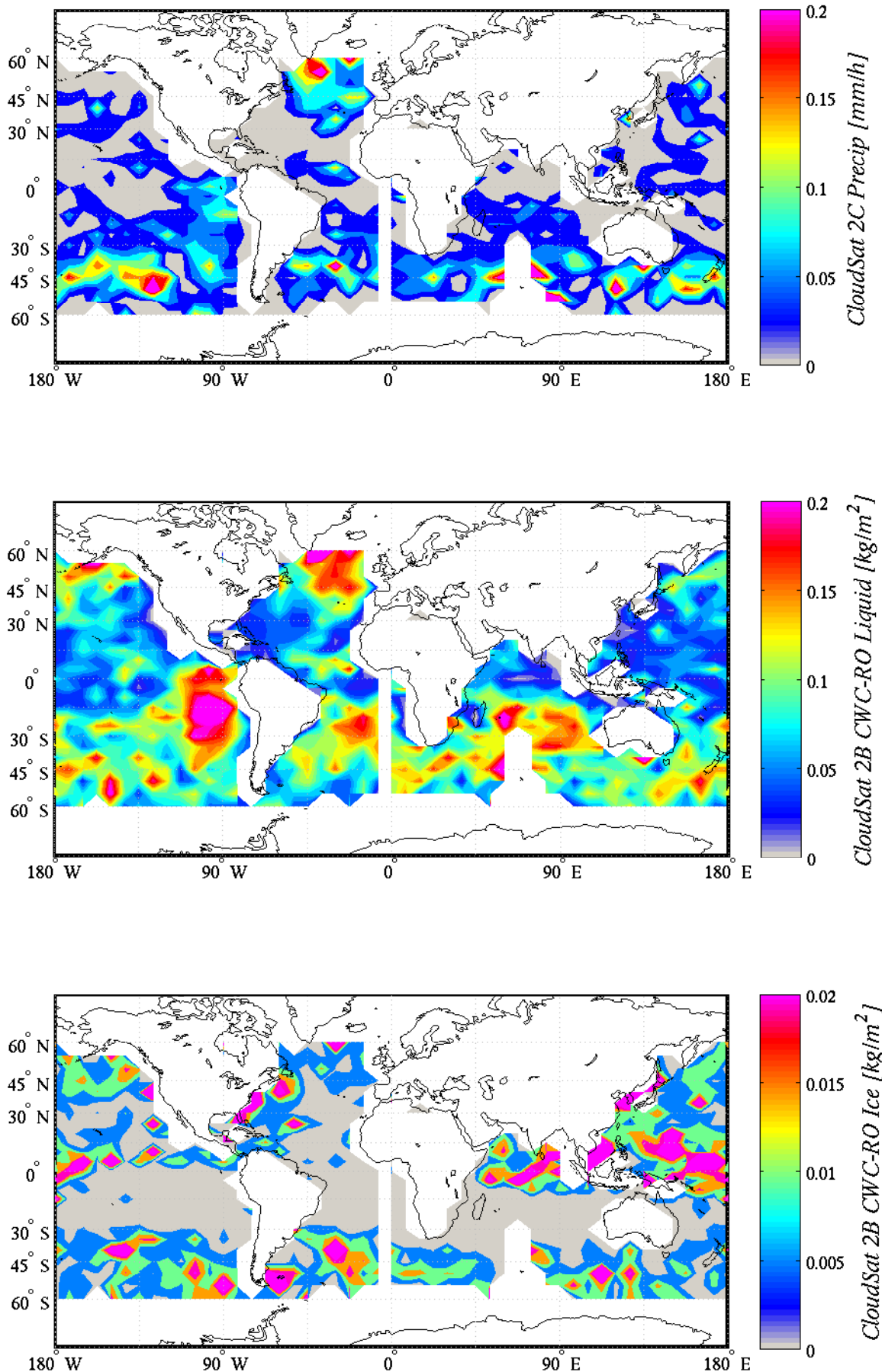
*Figure 4.11: Map of mean CloudSat estimates for the occurrences matched with AMSR-E 'all-sky' 4D-Var. From top to bottom: 2C-PRECIP-COLUMN, 2B-CWC-RO Liquid, and 2B-CWC-RO Ice.*
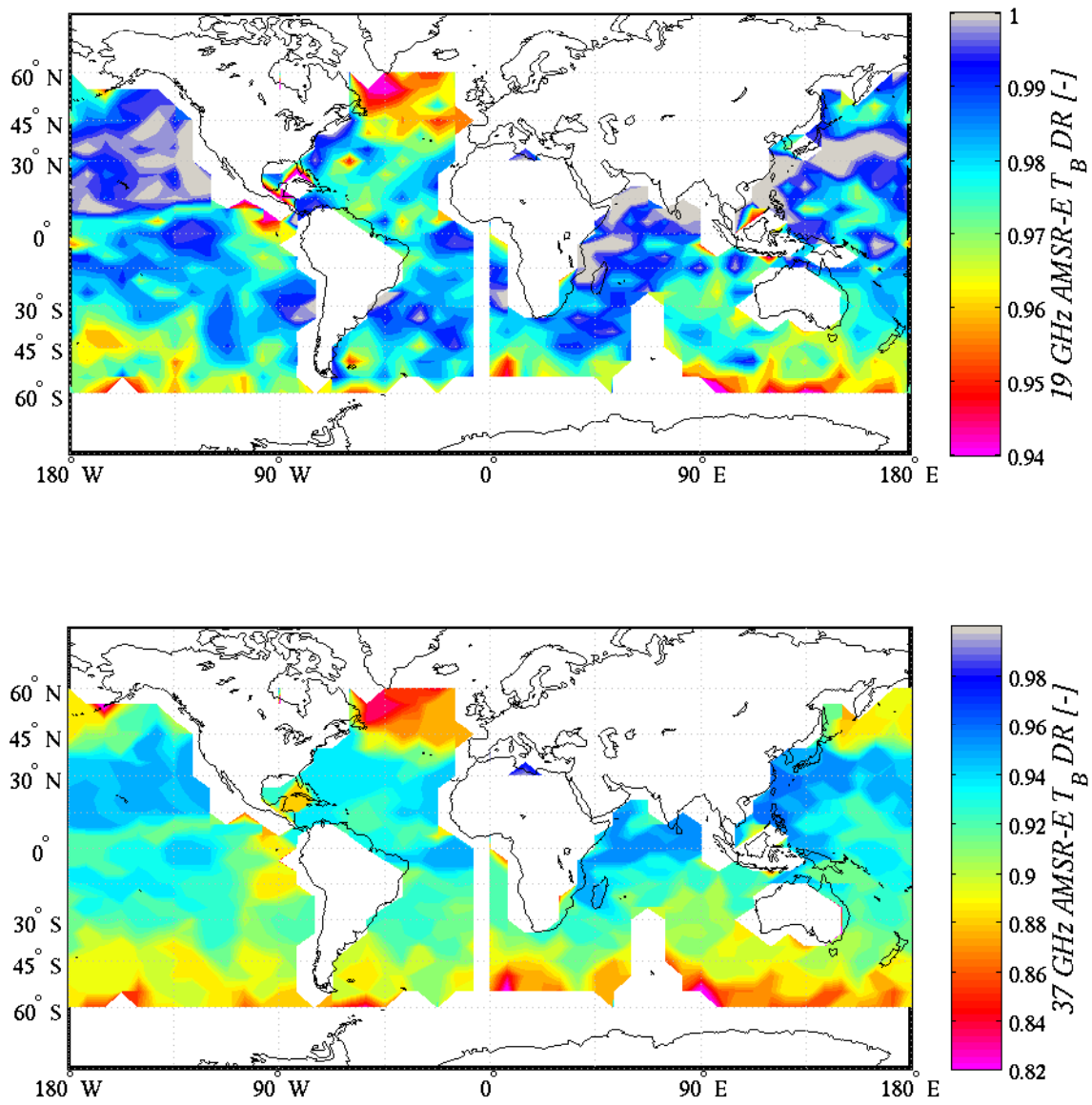
*Figure 4.12: Mean AMSR-E depolarization ratio for the occurrences matched with CloudSat. Top panel: 19 GHz. Bottom panel: 37 GHz.*
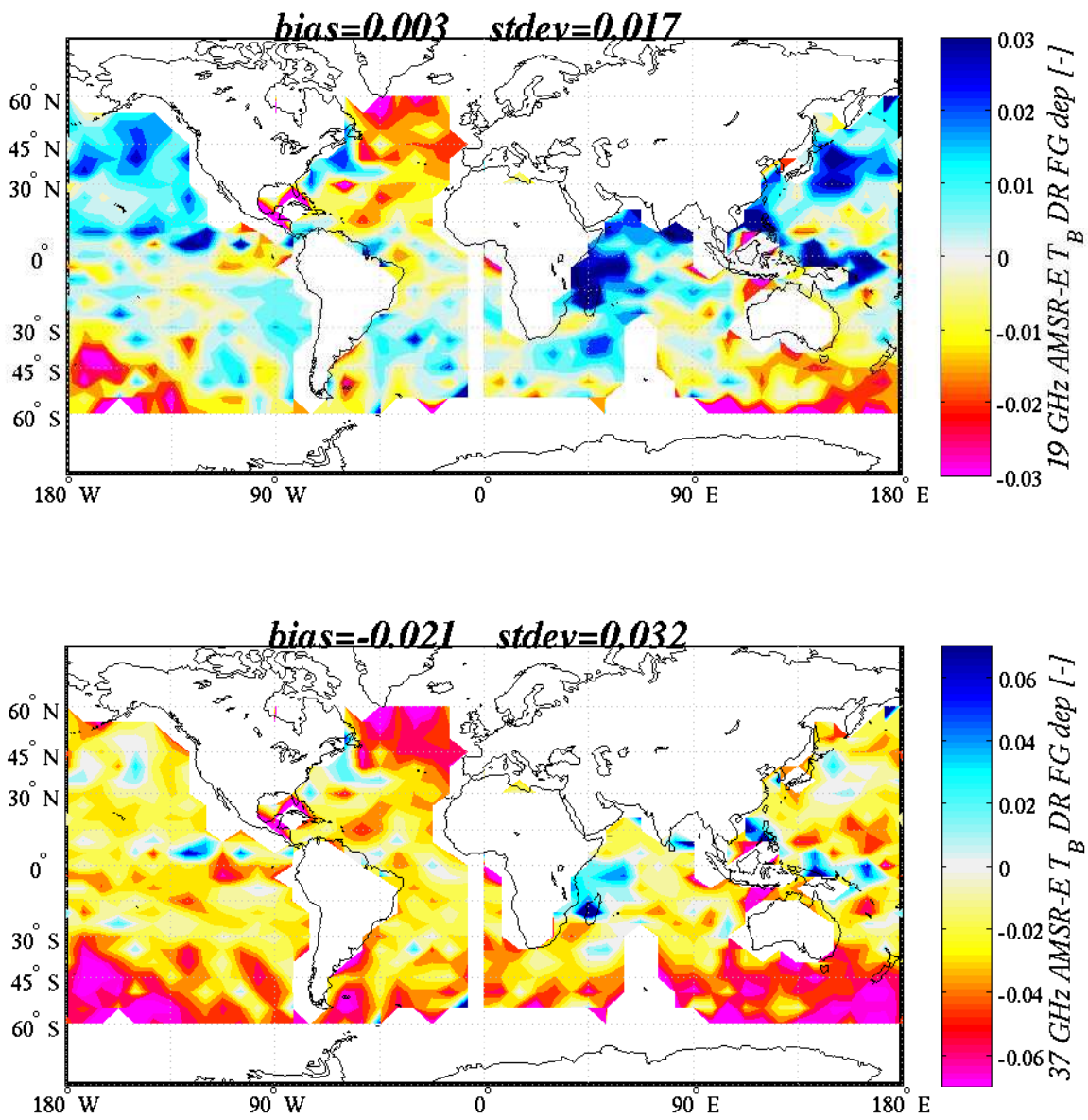
*Figure 4.13: Mean AMSR-E first guess departures (as depolarization ratio) for the occurrences matched with Cloud-Sat. Top panel: 19 GHz. Bottom panel: 37 GHz.*
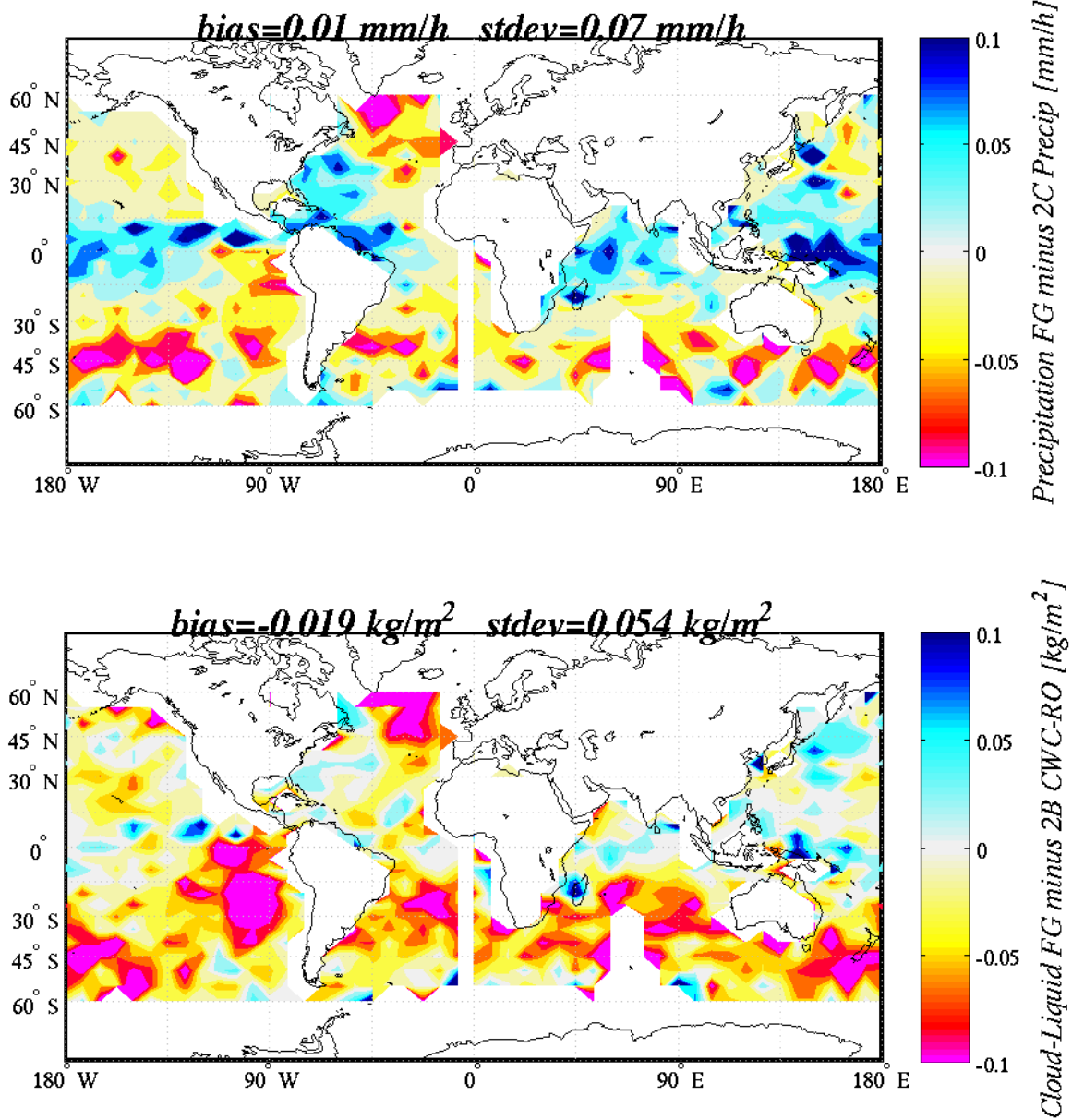
Figure 4.14: Map of mean differences between 4D-Var first guess and CloudSat estimate. Top panel: surface precipitation. Bottom Panel: cloud liquid.
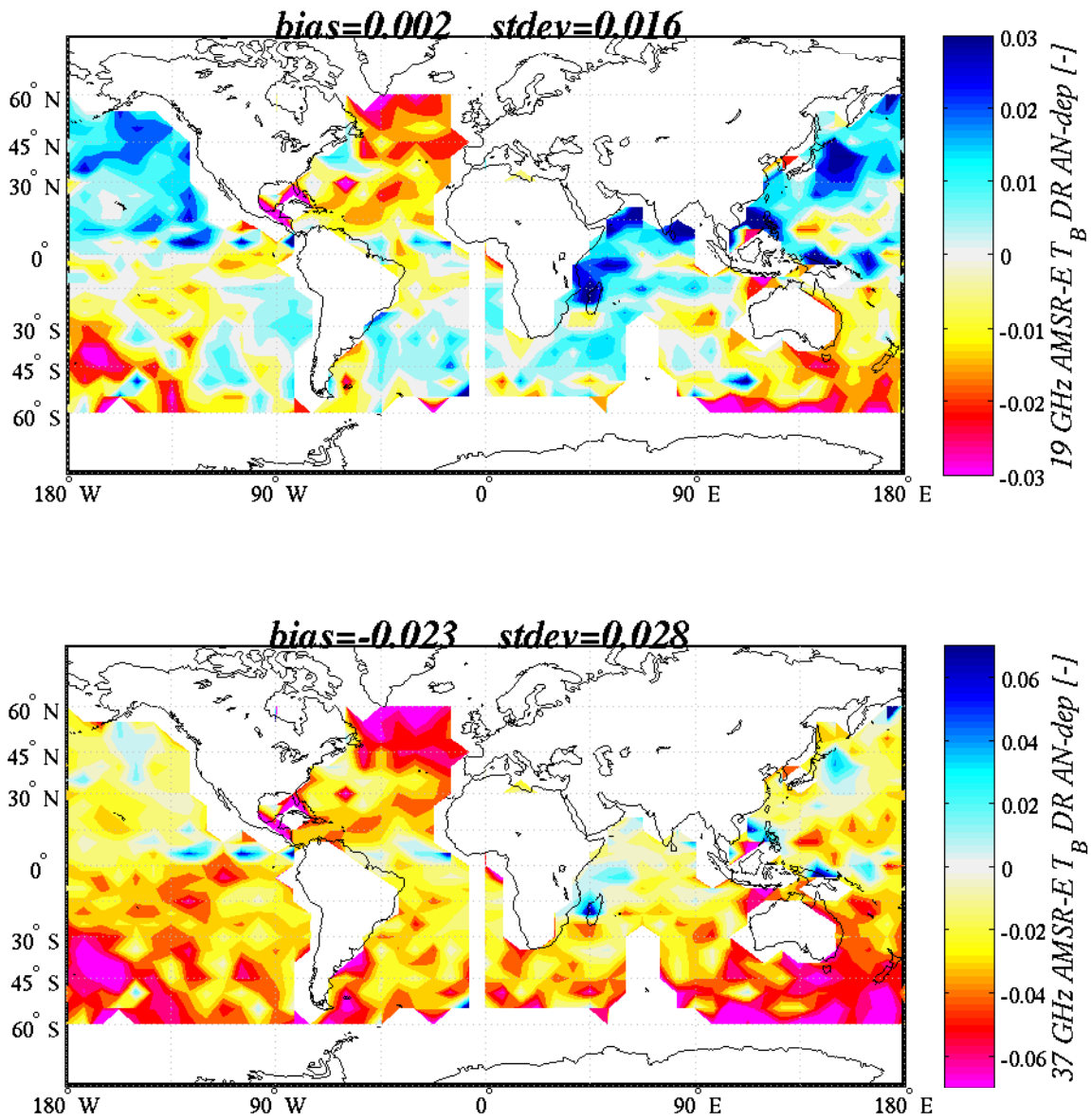
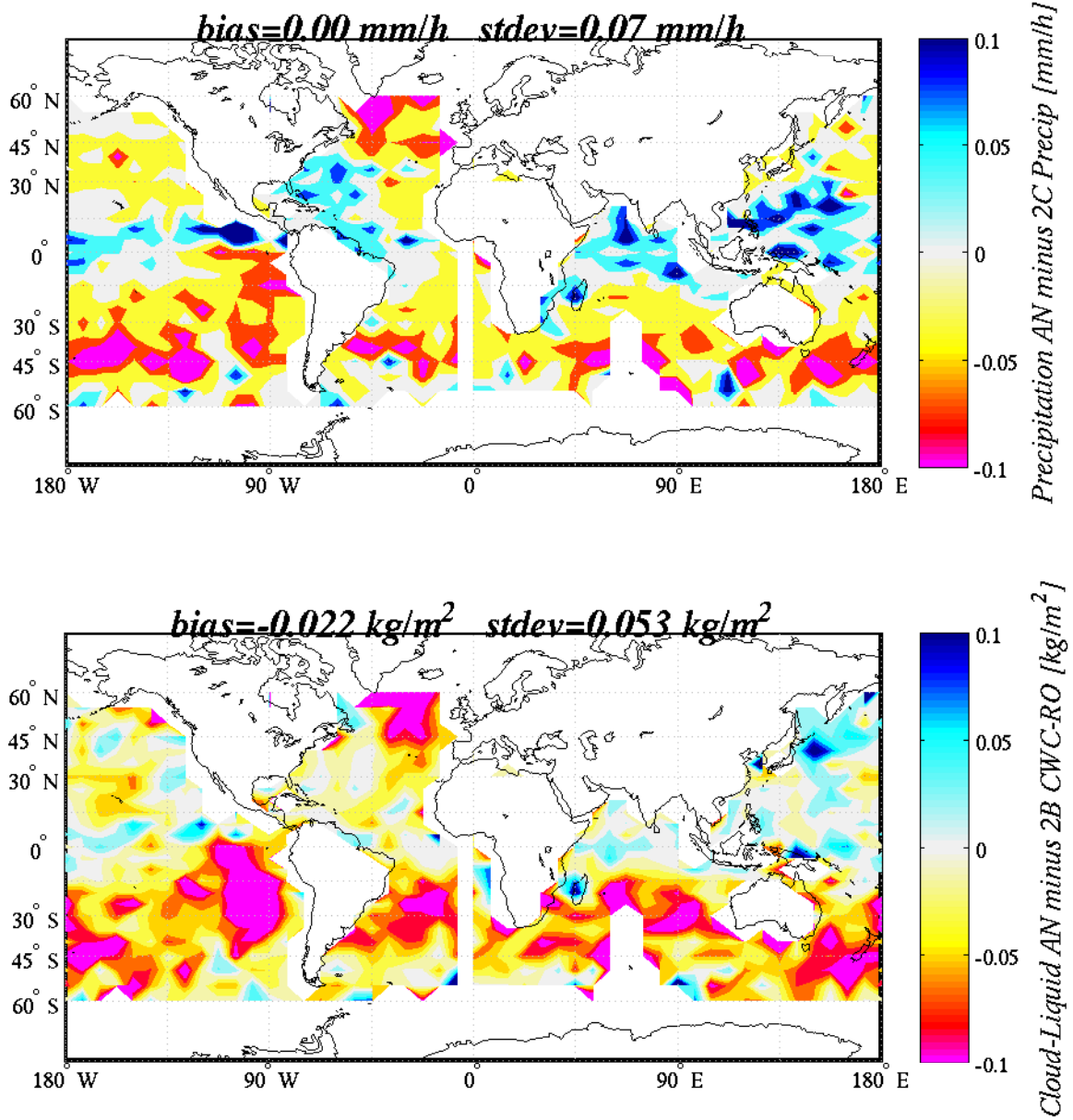*Figure 4.15: As Fig.4.13, but for 4D-Var analysis.*

*Figure 4.16: As Fig.4.14, but for 4D-Var analysis.*

# 5   Conclusions and perspectives

Cloud variables within the ECMWF assimilation system have been compared against coincident CALIPSO and CloudSat products. In Section 2, cloud-top height from CALIPSO has been used to verify the same quantity derived from AIRS IR radiances in the context of the cloud detection scheme and within the assimilation of IR cloudy radiances. The investigation has shown that CALIPSO observations constitute a valuable tool for the validation of cloud-top height retrieved from IR radiances. The intercomparison revealed the need for screening cases with non-contiguous clouds in the radiometer FOV (broken) to take into account differences in resolution between the instruments. The study has also shown that an opaque/transparent classification highlights the limitations of deriving CTH from passive instruments. The verification of cloud-detection for hyperspectral IR sounders has revealed some difficulties with cirrus clouds. The verification of cloud-top height for the assimilation of IR cloudy radiances has shown a reasonable agreement in overcast situations, but a larger scatter in conditions of partial beam filling.

In Section 3, height assignments for the atmospheric motion vectors have been compared against cloud-top height from co-located CALIPSO observations. The intercomparison has shown an overall tendency of height assignment for AMVs to produce CTHs lower than CALIPSO, especially for high clouds. This implies that the assimilated AMVs are associated with biased heights that, in turn, biases the wind analysis. The study has revealed that the agreement is sensitive to the definition of the averaged CALIPSO CTH in the AMV domain. Similarly to the AIRS validation, we have found that a screening for spatial representativeness is needed after the co-location. Also, the opaque/transparent distinction of CALIPSO observations highlights those most difficult cases that cannot be easily retrieved from passive instruments. A cloud can be detected by a lidar, while it can exhibit such a small optical depth that no signature in visible or IR radiances is observed.

Finally, the assimilation of microwave radiances in cloudy regions has been validated in Section 4 using CloudSat products of cloud and precipitation. The investigation showed that CloudSat provides useful information about the impact of clouds- and precipitation- affected passive MW observations in the assimilation system. The 1D-Var verification demonstrated that CloudSat products can be used to validate retrieval of clouds/precipitation from passive MW instruments. The intercomparison of the 4D-Var analysis revealed that differences between the first guess of the model cloud variables and CloudSat products have a geographical pattern similar to the one of AMSR-E first guess departures. This result shows that CloudSat has also the potential to be used as a diagnostic tool for the monitoring of the cloud- and rain-affected radiances assimilation. However, for both applications, a screening is needed to consider only CloudSat measurements representative within the radiometer FOV.

This study has clearly demonstrated that cloud radar and lidar observations can play a role as verification tools for the assimilation of cloud. The availability of similar observation from EarthCARE will therefore be beneficial to the future progresses of the ECMWF forecasting system. For this purpose, cloud products with comparable (or superior) quality to the ones of CALIPSO and CloudSat are required from this future mission.

# Acknowledgements

# 6  List of Acronyms

| | |
|---|---|
| 1D-/4D-Var | One-/Four-Dimensional Variational assimilation |
| AIRS | Atmospheric Infrared Sounder |
| AMSR-E | Advanced Microwave Scanning Radiometer for the Earth Observing System |
| AMSU-A | Advanced Microwave Sounding Unit-A |
| AMV | Atmospheric Motion Vector |
| CALIOP | Cloud-Aerosol Lidar with Orthogonal Polarization |
| CALIPSO | Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation |
| CPR | CloudSat Profiling Radar |
| CTH | Cloud Top Height |
| DMSP | Defence Meteorological Space Program |
| ECMWF | European Centre for Medium-Range Weather Forecast |
| EUMETSAT | European Organization for the Exploitation of Meteorological Satellites |
| FG | First Guess |
| FOV | Field Of View |
| HA | Height Assignment |
| HIRS | High Resolution Infrared Radiation Sounder |
| HRV | High Resolution Visible |
| IASI | Improved Atmospheric Sounder Interferometer |
| IR | Infrared |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| MSG | Meteosat Second Generation |
| MW | Microwave |
| NWP | Numerical Weather Prediction |
| PARASOL | Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a |
| PIA | Path Integrated Attenuation |
| QI | Quality Indicator |
| SEVIRI | Spinning Enhanced Visible and Infra-Red Imager |
| SSM/I | Special Sensor Microwave/Imager |
| $T_B$ | Brightness Temperature |
| TCWV | Total Column Water Vapour |
| TMI | TRMM Microwave Imager |
| TRMM | Tropical Rainfall Measuring Mission |
| WV | Water Vapour |

# References

Aumann, H. H., M. T. Chahine, C. Gautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkranz, W. L. Smith, D. H. Staelin, *et al.*, 2003: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems, *IEEE Transactions on Geoscience and Remote Sensing*, **41**(2), 253–264.

Bauer, P., A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radiances: Part I. Implementation, *Qurt. J. Roy. Meteor. Soc.*, **135**(submitted).

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2006a: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1D-Var, *Qurt. J. Roy. Meteor. Soc.*, **132**, 2277–2306.

Bauer, P., P. Lopez, A. Salmond, D.and Benedetti, S. Saarinen, and M. Bonazzola, 2006b: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var, *Qurt. J. Roy. Meteor. Soc.*, **132**, 2307–2332.

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach, *Qurt. J. Roy. Meteor. Soc.*, **120**(519), 1367–1388.

De Smet, A., 2008: AMV height assignment with Meteosat-9: Current status and future developments, in *Proceedings of the 9th International Winds Workshop, Annapolis. Available from EUMETSAT, Darmstadt, Germany, or online: http://www.eumetsat.int*.

EUMETSAT, 2009: Meteorological Products Extraction Facility (MPEF), Algorithms Specification Document. Doc. No. EUM.MSG.SPE.022.

Eyre, J. R. and W. P. Menzel, 1989: Retrieval of cloud parameters from satellite sounder data: A simulation study, *Journal of Applied Meteorology*, **28**(4), 267–275.

Fougnie, B., G. Bracco, B. Lafrance, C. Ruffel, O. Hagolle, and C. Tinel, 2007: PARASOL in-flight calibration and performance, *Applied Optics*, **46**(22), 5435–5451.

Geer, A. J., P. Bauer, and P. Lopez, 2008: Lessons learnt from the operational 1D+ 4D-Var assimilation of rain-and cloud-affected SSM/I observations at ECMWF, *Qurt. J. Roy. Meteor. Soc.*, **134**(635).

Geer, A. J., P. Bauer, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances: Part II. Assessment, *Qurt. J. Roy. Meteor. Soc.*, **135**(submitted).

Hollinger, J. P., J. L. Peirce, and G. A. Poe, 1990: SSM/I instrument evaluation, *IEEE Transactions on Geoscience and Remote Sensing*, **28**(5), 781–790.

Kawanishi, T., T. Sezai, Y. Ito, K. Imaoka, T. Takeshima, Y. Ishido, A. Shibata, M. Miura, H. Inahata, and R. W. Spencer, 2003: The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies, *IEEE Transactions on Geoscience and Remote Sensing*, **41**(2), 184–194.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The tropical rainfall measuring mission (TRMM) sensor package, *Journal of Atmospheric and Oceanic Technology*, **15**(3), 809–817.

McCormick, M. P., 2004: Space lidar for Earth and planetary missions, in *22nd Internation Laser Radar Conference (ILRC 2004)*, volume **561**, p. 933.

McNally, A. P., 2009: The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var, *Qurt. J. Roy. Meteor. Soc.*, **135**(642).

McNally, A. P. and P. D. Watts, 2003: A cloud detection algorithm for high-spectral-resolution infrared sounders, *Qurt. J. Roy. Meteor. Soc.*, **129**(595), 3411–3424.

Petty, G. W., 1994: Physical retrievals of over-ocean rain rate from multichannel microwave imagery. Part I: Theoretical characteristics of normalized polarization and scattering indices, *Meteorology and Atmospheric Physics*, **54**(1), 79–99.

Schoeberl, M. R., A. R. Douglass, E. Hilsenrath, P. K. Bhartia, R. Beer, J. W. Waters, M. R. Gunson, L. Froidevaux, J. C. Gille, J. J. Barnett, *et al.*, 2006: Overview of the EOS Aura mission, *IEEE Transactions on Geoscience and Remote Sensing*, **44**(5), 1066–1074.

Stephens, G. L., D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, *et al.*, 2002: The Cloudsat mission and the A-Train, *Bull. Amer. Meteor. Soc.*, **83**, 1771–1790.

Vaughan, M. A., K. A. Powell, R. E. Kuehn, S. A. Young, D. M. Winker, C. A. Hostetler, W. H. Hunt, Z. Liu, M. J. McGill, and B. J. Getzewich, 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *Journal of Atmospheric and Oceanic Technology*, **26**(10), 2034–2050.

Winker, D. M., J. Pelon, and M. P. McCormick, 2003: The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds, in *Proc. SPIE*, volume **4893**, p. 11.