VALIDATION OF SAFNWC / MSG CLOUD PRODUCTS WITH ONE YEAR OF SEVIRI DATA

M.Derrien[1], H.Le Gléau[1], Jean-François Daloze[2], Martial Haeffelin[2]

[1]Météo-France / DP / Centre de Météorologie Spatiale. BP 50747. 22307 Lannion. France
[2]Laboratoire de Météorologie Dynamique. Ecole Polytechnique. 91128 Palaiseau. France

ABSTRACT

Within the SAF in support to Nowcasting and Very Short Range Forecasting (SAF NWC), Météo-France has developed a software to extract cloud parameters (cloud mask and types, cloud top temperature and height) from MSG SEVIRI imagery. These modules are part of the SAFNWC/MSG software package whose first operational version (v1.0) has been available to users since June 2004. One year later, users could get an improved version (v1.2) tuned using SEVIRI images on a longer period.

This paper focuses on the v1.2 validation results. We especially present the validation of the cloud mask with surface observations available in SYNOP and the cloud top height with radar and lidar measurements, and radio-sounding data (only for low clouds).

1. INTRODUCTION

Within the SAF in support to Nowcasting and Very Short Range Forecasting (SAF NWC), Météo-France has developed a software to extract cloud parameters (cloud mask and types, cloud top temperature and height) from MSG SEVIRI imagery. These modules are part of the SAFNWC/MSG software package whose first operational version (v1.0) has been available to users since June 2004. One year later, users could get an improved version (v1.2) tuned using SEVIRI images on a longer period. Overviews of the algorithms have already been presented in previous Eumetsat Meteorological Satellite Conferences and at SAFNWC users' workshops. A detailed description is available in the SAFNWC User Manual (scientific part) that can be downloaded from the SAFNWC help desk (http://nwcsaf.inm.es) or from www.meteorologie.eu.org/safnwc.

This v1.2 version has been validated from one year of SEVIRI well calibrated data by using collocated SYNOP observation, manual nephanalyst reports, radio-soundings, ground-based radar and lidar measurements. This paper intends to summarize some aspects of these validation activities focusing on the cloud mask and cloud top height. A report presenting full validation results can be downloaded from www.meteorologie.eu.org/safnwc.

The paper first presents the comparison over Europe and north Africa of the SAFNWC/MSG cloud mask with surface-based cloud cover available in SYNOP. Contingency tables and statistical scores have been computed to assess the cloud detection efficiency for different illumination and geographical areas. These scores have also been compared to those obtained from the MPEF cloud mask using the same validation procedure.

The paper then addresses the cloud top height validation. SAFNWC/MSG cloud top heights and collocated ground-based radar and lidar measurements regularly gathered during one year at the SIRTA site (located at LMD near Paris) have been faced each other to produce error statistics stratified according the retrieval method. Low cloud top heights have been automatically estimated from radio-soundings scattered all over Europe and compared to collocated SAFNWC/MSG cloud top pressure during one year, allowing to analyse the impact on the error of the presence and strength of low level thermal inversions.

2. CLOUD MASK (CMA) VALIDATION

2.1 CMa validation database

A database has been automatically built, gathering collocated MSG-1/SEVIRI data and surface observations. The surface data used are the hourly weather observations, coded by the observers into the World Meteorological Organisation (WMO) synoptic code (SYNOP), from a list of 302 selected land stations over the European and Northern Africa area (see figure 1; stations suspected to be automated have been rejected from an initial list of 535 stations). The corresponding satellite data and ancillary data (satellite and solar angles, time information, atlas value, the collocated and nearest in time meteorological information extracted from the French NWP model ARPEGE forecast fields) are collected in a 5x5 pixel (IR resolution) box centred on the land station location. The database gathers 708797 collocated observations from 1 November 2003 until 28th February 2005.

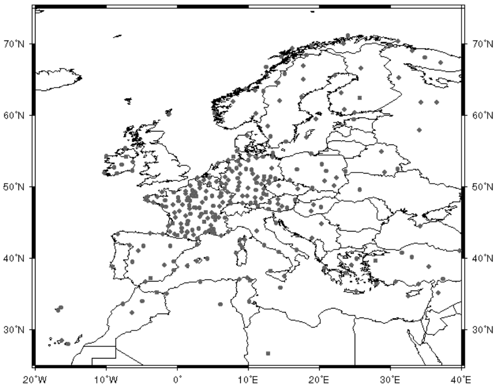

Figure 1 - Location of the 302 selected land stations over the European and African area, used for the validation of the MSG/SEVIRI cloud mask.

From the SYNOP data set, ground-based total and partial cloud covers from low, medium and high clouds are available. Satellite cloud coverage is estimated from SAFNWC/MSG cloud mask (CMa) applied over an area of 5 by 5 SEVIRI pixels centred on the SYNOP observation, counting every pixel detected as cloud contaminated as totally covered by clouds. To simulate the surface observations from the satellite pixels, no attempt has been made to take into account the complexity of the observation, and all the twenty-five pixels inside the satellite data target are used for the evaluation. The technique used to assess the CMa cloud detection is to compare its overall cloudiness with that given by the SYNOP which is considered as the ground truth. We are aware that the difference between the two overall cloudiness can be partly due to errors linked to the comparison method itself (different areas viewed by the observer and by the satellite, parallax error impacting the clouds' position on the satellite image (especially at high viewing angles)), but also to the subjectivity of the human observation especially in night-time conditions.

The database is stratified according to illumination (day, night, twilight), and latitude (latitude higher (Nordic) or lower (mid-latitude) than 55 degrees). In all these comparisons, CMa is retrieved using NWP fields forecast by the French model ARPEGE four times per day (0h, 6h, 12h and 18h) on a 1.5 degree horizontal resolution grid.

2.2 CMa validation results

The distribution of the satellite mean CMa cloud cover error expressed in octas (defined as the difference between CMa and SYNOP cloudiness) is plotted as a function of SYNOP observed cloudiness on figure 2 for mid-latitude (left) and Nordic (right) conditions.

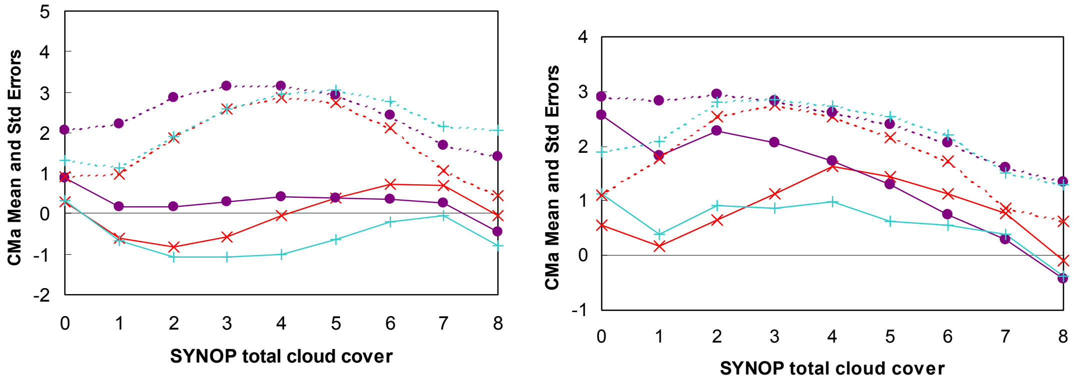

Figure 2 – CMa errors (defined as the difference between CMa and SYNOP cloud cover) as a function of cloud cover (in octa) visually observed in meteorological station (SYNOP measurements) for continental mid-latitude (left) and Nordic (right) regions. Red (x), blue (+) and purple (o) symbols correspond respectively to daytime, twilight and night-time conditions. Solid and dotted lines corresponds to mean and standard deviation errors.

The ability of CMa to detect fully cloudy and cloud free events can also be quantified by building contingency tables and computing statistical scores as explained below. A target (5x5 pixels) is considered as:

- « observed cloudy » when the total cloudiness observed from the surface is strictly larger to 5 octas,
- « observed clear » when the total cloudiness observed from the surface is strictly inferior to 3 octas (even if the ground is covered by snow)
- « detected cloudy » when strictly more than 16 of its 25 pixels are masked cloudy by CMa,
- « detected clear » when strictly less than 8 of its 25 pixels are masked cloudy by CMa.

According to these definitions contingency tables are built (see table 1 for convention) and several statistical scores are computed :

- PC= 100x(n_a+n_d)/($n_a+n_b+n_c+n_d$) is the percentage of correct CMa results.
- 1-POD$_{cloud}$=100x(n_b)/(n_a+n_b) reflects the failures among cloudy targets (CMa underestimation of clouds). (POD is the Percentage Of Detection).
- 1-POD$_{clear}$=100x(n_c)/(n_c+n_d) reflects the failures among clear targets (CMa overestimation of clouds).

	Detected Cloudy	Detected Clear
Observed Cloudy	n_a	n_b
Observed Clear	n_c	n_d

Table 1 - Contingency tables conventions

	PC (%)	1-POD$_{cloud}$ (%)	1-POD$_{clear}$ (%)
Midlatitude	95.02	4.37	6.14
Midlatitude day	97.68	2.40	2.16
Midlatitude night	92.93	4.89	10.66
Midlatitude twilight	91.97	10.14	3.14
Nordic	93.24	2.77	22.02
Nordic day	97.01	1.23	9.76
Nordic night	89.99	3.45	33.34
Nordic twilight	94.40	3.57	14.58

Table 2 - CMa performance in the detection of fully cloudy and cloud-free events over continental mid-latitude regions estimated from collocated surface and satellite observations (from 1/11/2004 to 28/02/2005).

Figure 2 and table 2 summarize the CMa performance. The mid-latitude subset shows a rather good agreement as illustrated on figure 2 (left). The slight underestimation of the cloud cover in twilight conditions, observed in the figure 2 and confirmed by a (1-POD$_{cloud}$) of 10.14%, is probably linked to the rather frequent low clouds non-detection when the sun is very low over the horizon (due to the 3.9μm channel contamination by the sun). On the other hand, it is likely that the cloud cover overestimation in night-time conditions, observed in Figure 2 (for SYNOP cover of 0) and confirmed by a (1-POD$_{clear}$) of 10.66%, is partly (if not mainly) due to human observation errors in darkness.

Concerning latitude impact, a poorer agreement in Nordic conditions was expected, because of more extreme climatic conditions, lower solar energy and especially high viewing angles and large pixels size. Surprisingly, the (1-POD$_{cloud}$) (table 2) indicates that the ability to detect overcast situations in Nordic conditions is similar to that in mid-latitude regions and even significantly better in case twilight conditions. This could be explained by the strong efficiency of the 8.7μm when combined to the 10.8μm to detect low clouds at high viewing angles. Nevertheless, a trend to overestimate the cloud cover in all illumination conditions, not observed in mid-latitude except at night, is highlighted by high (1-POD$_{clear}$) values and is also observable in figure 2 (right). This trend can be explained by two reasons: the surface of fractional clouds seen by satellite increases with viewing angles thus making their detection easier, and a pixel detected by CMa algorithm is always treated as overcast in our comparison even if it is only slightly contaminated by clouds. For completely cloud free situations (SYNOP cover of 0 in figure 2), this overestimation is especially present at night-time and to a lower extent in twilight conditions. This could be related to human observation errors in darkness (as in mid-latitude regions) but also to cold snowy surfaces confused with clouds.

In order to compare the SAFNWC/MSG cloud mask (CMa) and the MPEF cloud mask, the previous procedure is applied to both cloud masks on the same list of stations for the limited time period (from 25th August 2004 to 28th February 2005). The two schemes are not compared before 25th August 2004 because the visible channels were not yet used by the MPEF scheme leading to much lower MPEF quality. The statistics shown in table 3 indicate that the SAFNWC/MSG cloud mask has a better quality (better PC score, better (1-POD$_{clear}$) and (1-POD$_{cloud}$) scores) in all illumination conditions (daytime, night-time and twilight).

	PC (%)	**1-POD$_{cloud}$ (%)**	**1-POD$_{clear}$ (%)**
SAFNWC	95.03	4.97	4.98
SAFNWC day	98.06	2.23	1.36
SAFNWC night	93.96	4.79	8.34
SAFNWC twilight	91.31	11.30	2.07
MPEF	90.21	11.57	6.22
MPEF Day	93.20	9.26	1.76
MPEF night	89.28	11.00	10.21
MPEF twilight	86.18	18.0	3.19

Table 3 – SAFNWC/CMa and MPEF cloud mask performance in the detection of fully cloudy and cloud-free events over continental regions estimated from collocated surface and satellite observations (from 25/08/2004 to 28/02/2005).

3. CLOUD TOP TEMPERATURE AND HEIGHT (CTTH) VALIDATION

In both validation studies, CTTH is retrieved with SAFNWC/MSG software using NWP fields forecast by the French model ARPEGE four times per day (0h, 6h, 12h and 18h) on a 1.5 degree horizontal resolution grid. Temperature and humidity are available on twenty pressure levels (10, 20, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 850, 900, 925, 950, 1000).

3.1 Low clouds Cloud Top Pressure validation with radio-soundings

Radio-sounding measurements, gathered all over Europe (see figure 3) during one year (March 2004-February 2005), are used to evaluate low cloud top pressure errors. The cloud top pressure is estimated from each radio-soundings using empirical Pone’s rules: stable and unstable layers are first looked for from surface up to the tropopause ; cloud layers are then derived depending on the stability and saturation of these layers. Only cases corresponding to one cloud layer in the radio-sounding and to large spatially homogeneous low cloud layers in the satellite imagery are retained. SEVIRI retrieved cloud top pressures are averaged over 9 by 9 SEVIRI pixels areas and are compared to those obtained from corresponding radio-sounding. Daytime (around 12hUTC) and night-time (around 0hUTC) observations are analysed separately.

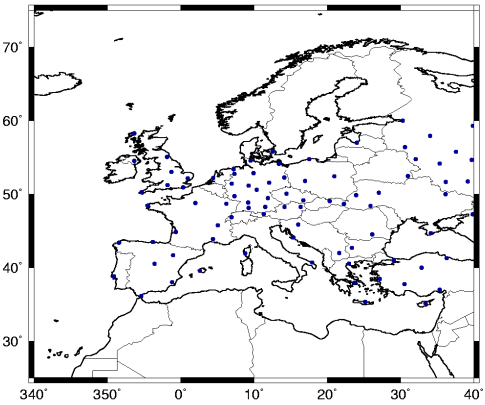

Figure 3 - Location of the selected radio-sounding stations used for the validation of the MSG/SEVIRI cloud top pressure.

Table 4 shows that on average, low clouds top pressures are underestimated by 15hPa at night-time and by 30hPa at daytime, but a large scatter is observed (70hPa standard deviation). Figure 4 suggests that clouds top pressures are generally underestimated by more than 50hPa for cloud pressures lower than 800hPa whereas both underestimation and overestimation are observed for cloud pressures larger than 800hPa. This can be linked to the presence and strength of a thermal inversion in the radio-sounding, as shown in figure 5. The method to retrieve the cloud top pressure of low clouds is dependent on the presence of a thermal inversion in the forecast NWP fields. If a very strong inversion is observed in the radio-sounding, it is likely that this inversion is forecast by the NWP model even if its strength is not well simulated. In this case, the SEVIRI retrieved cloud top pressure is usually set below the inversion, and figure 5 indicates that a cloud pressure overestimation ranging around 50hPa is then observed. On the reverse, if no strong inversion is observed in the radio-sounding, there is a chance that no thermal inversion is simulated in the NWP fields and therefore a chance that the cloud top is set much to high in the atmosphere, thus leading to cloud pressure underestimation.

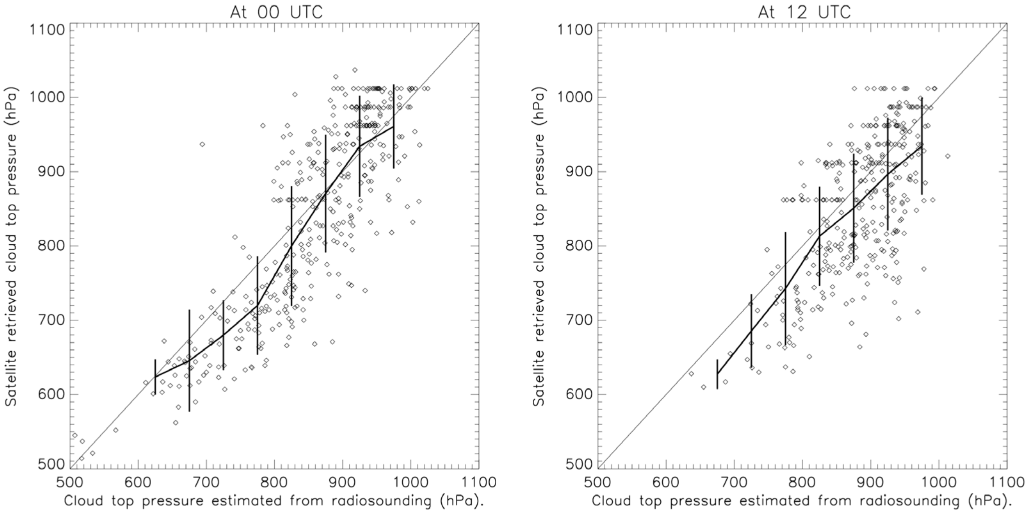

Figure 4 - Comparison between cloud top pressures retrieved from satellite and estimated from radio-soundings. The solid lines corresponds to mean and standard deviation computed in 50hPa intervals.

Time of day	Bias (Satellite-radio-sounding)	Standard deviation	Number of cases
0hUTC	-15.89 hPa	70.67 hPa	382
12hUTC	-30.06 hPa	71.77 hPa	344

Table 4 - Comparison between cloud top pressures retrieved from satellite and estimated from radio-soundings: bias, standard deviation and number of cases at 0hUTC and 12hUTC.

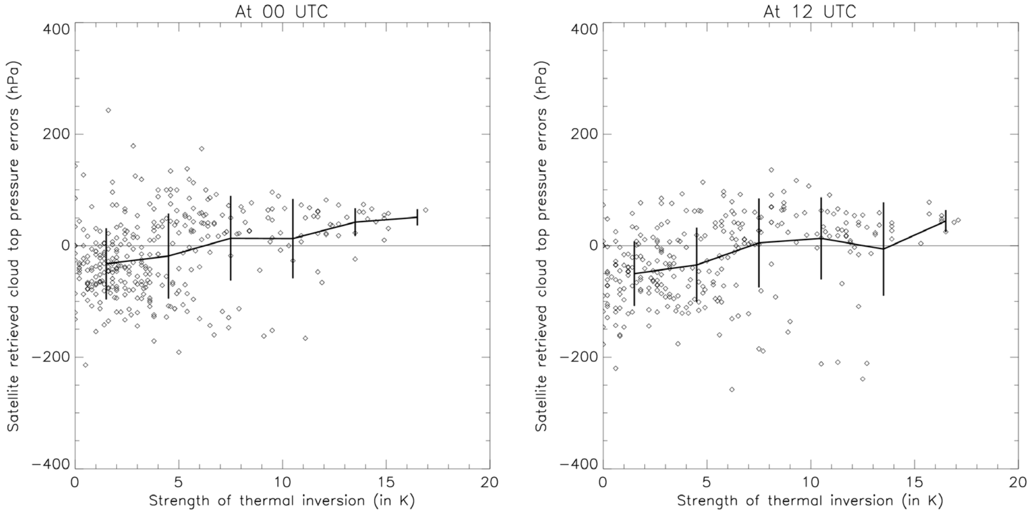

Figure 5 - Satellite retrieved cloud top pressure error as a function of the thermal inversion strength. The error stands for the difference of cloud top pressures derived from satellite and estimated from the collocated radio-sounding. The thermal inversion strength (observed in the radio-sounding) is defined as the difference between the cloud top temperature (estimated from radio-sounding) and the maximum air temperature observed above the cloud. The solid lines corresponds to bias and standard deviation computed in 2.5K intervals.

3.2 Cloud top Height (CTH) validation with radar and lidar measurements

Measurements of ground-based radar and lidar located near Paris gathered during one year (September 2003-October 2004) are used to evaluate cloud top height errors. The ground-based measurements used in this study are provided by SIRTA (Site Instrumental de Recherche par Télédétection Atmosphérique), an atmospheric observatory for cloud and aerosol research located near in Palaiseau near Paris and operated by the Institut Pierre Simon Laplace (IPSL).

The lidar instrument called LNA: Lidar Nuage Aerosol is a Nd-Yag pulsed lidar emitting at 532 and 1064 nm and linearly polarized that allows the detection of aerosol and cloud layer with visible optical thickness ranging from 0.05 to 3. The backscattered signal is sampled with a vertical resolution of 15 meters (up to 15km) with a nominal temporal resolution of 10s. The lidar is not appropriate to estimate cloud top height (CTH) for optically thick water clouds or in case thin ice clouds overlying a continuous layer of optically thick water clouds.

The cloud radar called RASTA (Radar Aéroporté et Sol de Télédétection Atmosphérique) is a vertically-pointing single beam 95GHz Doppler radar with a range resolution of 60 meters and the temporal resolution is 1 s. This instrument is devoted to the investigation of cloud processes, through the documentation of the microphysical, radiative and dynamical properties of all type of non-precipitating clouds. The radar is not appropriate to retrieve CTH for optically thin clouds above 10 km or fair weather cumulus clouds at the top of the boundary layer.

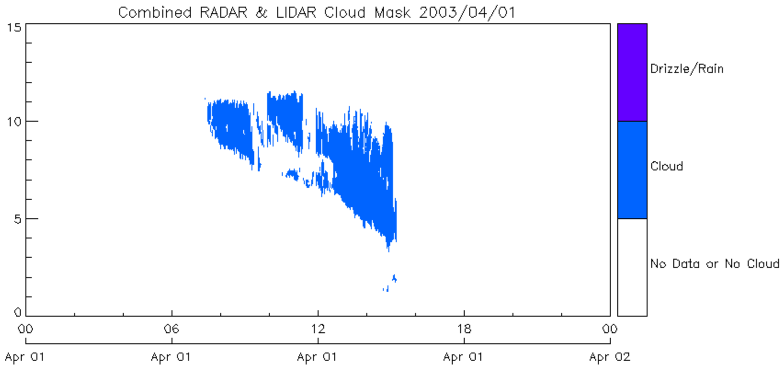

Figure 6 - Cloud mask derived from radar-lidar synergy (violet: drizzle or rain, blue: cloud, white: no data or no cloud). The temporal resolution is 30 s.

Products retrieved from the synergy between the radar and lidar are called the RALI data in this study. The radar and lidar measurements are averaged over a 30 s timeframe and then are independently analysed to retrieve cloud mask product (see figure 6). Even the combination of both instruments does not allow the CTH retrieval in case multi-layer situation with continuous layer of optically thick water clouds underlying a layer of optically thin ice clouds at a high altitude.

Radar and lidar data have been used to validate SAFNWC/MSG CTH for respectively opaque and semi-transparent clouds. The combination of radar/lidar is more reliable but restricts too much the dataset. The following procedure has been applied to build up the validation dataset:

- CTH derived from lidar/radar observations and retrieved from SEVIRI are respectively temporally averaged over 30 minutes and spatially averaged over 5x3 pixels.
- To minimize inherent discrepancies in ground and spatial observational scales, the comparison is restricted to the SEVIRI scenes showing a large homogenous cloud coverage (low/middle or high/semi-transparent cloud fractions must be larger than 80 % within an area of 11x7 pixels and 50% within an area of 5x3 pixels).
- When more than one cloud layer is detected by the radar or the lidar, the SEVIRI CTH is systematically compared with the nearest layer of ground-based instrument dataset.
- Suspect cases have been manually analysed by displaying the SEVIRI images and the full lidar and radar measurements. They often corresponds to multilayer clouds where the upper semi-transparent cloud layer was not automatically detected from the radar/lidar measurements. These cases have been removed from the study. There certainly still remains such cases, therefore part of the comparison errors certainly comes from the ground-based radar/lidar measurements themselves.

Cloud type	Mean (km)	STD(km)	Median (km)	Number
Low opaque clouds	0.32	1.03	0.26	1133
Medium or high opaque clouds	-0.40	1.17	-0.46	1030
Semi-transparent clouds (intercept method)	-1.08	1.09	-1.26	314
Semi-transparent clouds (radiance ratioing method)	-0.06	1.18	-0.10	211

Table 5 - Statistical scores for (CTH_SEVIRI-CTH_radar/lidar) for various cloud types. Negative mean or median values correspond to SEVIRI CTH underestimation.

Results are illustrated on figure 7 and summarized in table 5 for low opaque clouds, medium and high opaque clouds, semi-transparent clouds using either intercept or radiance ratioing methods:

- <u>low opaque clouds</u>: on average, SEVIRI CTH shows an overestimation of 320m as compared to radar/lidar retrieved CTH with a standard deviation of 1030m. This average overestimation was also found with radio-soundings but with a lower standard deviation (“only” 71hPa, i.e. around 700m in the lower troposphere). The study indicated a slight diurnal cycle with SEVIRI CTH overestimation at night or at low solar elevation and SEVIRI CTH underestimation at high solar elevation.

- <u>medium/high opaque clouds</u>: on average, SEVIRI CTH shows an underestimation of 400m as compared to radar/lidar retrieved CTH with a standard deviation of 1170m. Figure 7 shows that this underestimation is a function of cloud height: stronger estimation for clouds at 6-8km and practically no bias at 4-5km. No dependency with illumination has been found.
- <u>semi-transparent clouds (intercept method)</u>: on average, SEVIRI CTH shows an underestimation of 1080m as compared to lidar retrieved CTH with a standard deviation of 1090m. The study indicated a stronger underestimation when channel 7.3µm is used (274 cases). No dependency with illumination has been found.
- <u>semi-transparent clouds (radiance ratioing method)</u>: on average, SEVIRI CTH shows an underestimation of 60m as compared to lidar retrieved CTH with a standard deviation of 1180m. It must be noted that this method is only applied to nearly opaque clouds (effective cloudiness larger than 0.8). No dependency with illumination has been found.

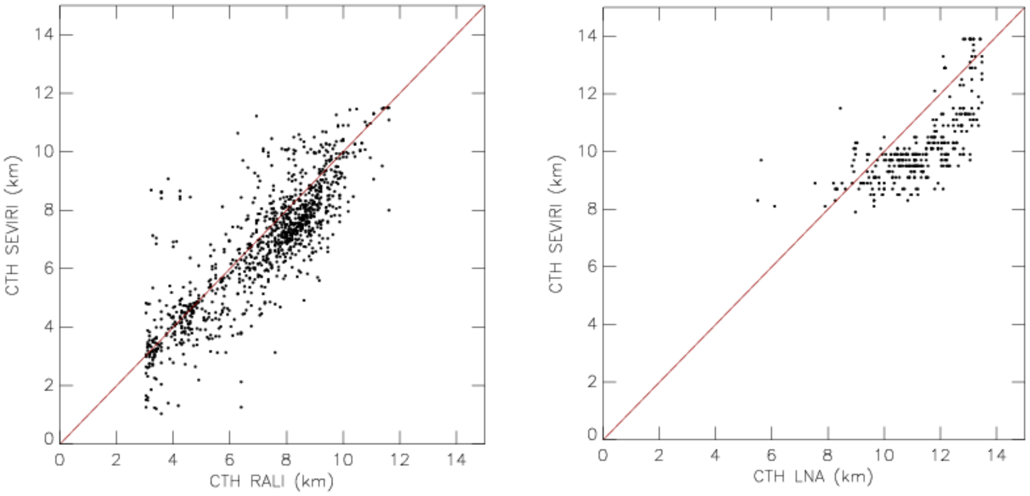

Figure 7 - Example of comparison of CTH retrieved from SEVIRI and from Radar/lidar (RALI) or lidar (LNA) for medium/high opaque clouds (left); for semi-transparent clouds (right).

4. SUMMARY AND CONCLUSION

The SAFNWC/MSG cloud mask (CMa) has been compared with SYNOP observations over Europe and north Africa. This comparison shows an overall good quality (the best results being obtained at daytime and in mid-latitude regions). The main problem (low cloud non-detection in twilight conditions) is emphasized by this study. This study also objectively showed that the quality of SAFNWC/MSG cloud mask is significantly better than that of MPEF.

The quality of SAFNWC/MSG cloud top height (CTTH) has been estimated by the use of radio-sounding (over Europe) and ground-based lidar and radar measurements over Paris. The main conclusions are a general underestimation of cloud top height for opaque medium of high clouds (in general a negative bias of 400m) and semi-transparent (a negative bias of 1.17km if retrieved using intercept method) and a lower quality for low opaque clouds (especially in case the presence of a thermal inversion) characterised by a relatively small cloud top height overestimation (320m) associated with a rather high standard deviation (1.03km). Additionally, it must kept in mind that no CTTH is obtained for clouds classified as fractional and for some very thin cirrus.

Some development activities will be planned to account for these validation results. Moreover, the strategy for building up validation database will be slightly improved. The SYNOP database will be improved: a more homogeneous repartition of the station will be looked for, the satellite content will be modified by keeping the HRV data and two successive SEVIRI slots to allow the testing of algorithm including temporal analysis. We will also try to use radar/lidar measurements in other locations. More over, a new opportunity to validate cloud parameters will appear in 2005 with satellite based radar (rainsat) and lidar (calypso).