MAIA AVHRR CLOUD MASK AND CLASSIFICATION

L. Lavanant

Météo-France. CMS

Lannion, France

MAIA AVHRR CLOUD MASK AND CLASSIFICATION

L. Lavanant

Météo-France. CMS

Lannion, France

1. GENERALITIES

The MAIA scheme is a library independent of the main program and can be invoked by different applications, provided the input arguments of the processed situation are correctly filled. Its aim is to determine if the input situation is clear or cloudy and to classify the cloud.

Depending of the main program, the input situation can be an AVHRR observation at full resolution for LAC applications or for cloud characterization inside a sounder fov. Or, the input situation can be output data from a previous processing; ex: cluster analysis outputs. In that case, the AVHRR channels and local variances are mean values computed from the AVHRR pixels inside the cluster. For convenience, the input situation is often referenced as AVHRR pixel or observation.

The cloud detection algorithm is a succession of thresholds tests applied to every AVHRR situation to various combinations of the AVHRR channels. It follows the scinetific algorithm developed in **[1]**. A situation is said to be cloudy if one test is not satisfied (so a pixel is said to be 'clear' if all tests are satisfied).

The series of tests applied depend on:

- the surface type (land, sea or coast)
- the solar zenith angle which determines the period of the day (day, twilight or night) and if there is or not specular reflection during the daytime (sunglint)

The surface type, altitude and solar zenith angle are input arguments of the subroutine and are defined in the calling program. Surface and altitude are optional and if not available, the MAIA scheme makes their determination from the situation position using provided datasets.

Daytime period is defined by solar zenithal angle less than 83 degrees, night-time period by solar zenith angle higher than 90 degrees and twilight if the angle is between 83 and 90 degrees.

Depending on the surface type, daytime period and specular reflection, different subroutines are invoked (ex: testsd.f90 for sea and daytime conditions) with different series of tests and threshold values. Successions of tests for each case are described below in tables.

The tests are done on single channels (11µm brightness temperature, visible reflectance), on combination of channels, in brightness temperature, for 11-12µm (t4-t5), 11-3.7µm (t4-t3), 3.7-12µm (t3-t5), on spatial local variances of channels 1,2,4,3-4 computed on a 3*3 box centered on each AVHRR pixel.

Applied thresholds in the tests could be constant values (cst_...) initialized within the module file mod_maia.f90 or computed values (s_a4 threshold, s_a1 threshold, s_34 etc.).

The constant values were determined by a long experience over the Europe acquisition area. However, most of the thresholds are now computed to allow the use of the cloud mask in other climates (ex: wet tropical).

The computed thresholds depend on the specific conditions of the situation through:

- the measurement conditions (solar and local zenith angles)
- the environmental conditions obtained from external information at the pixel position

- the satellite through tables computed off-line : these tables which describe the variation of some thresholds (ex: t4-t5) with the twvc and the secant of the zenithal angle, depend on the characteristics of the channels and have to be recomputed for each new satellite.

The external environmental conditions, determined at the situation position from provided global and monthly climatologies or from a forecast. are:

- the surface temperature from a climatology over sea or the air surface temperature from the forecast over land
- the visible reflectance from a climatology over land and daytime
- the total water vapor content (twvc) from a previous processing (ex from a regression with mapped AMSU-A channels over sea), the forecast over land or finally when nothing else is available.

Three sets of monthly global climatologies are provided with the software for the sea surface temperature over sea (atlas_sst_mm.dat), the visible reflectance over land (atlas_alb_mm.dat) and the specific humidity (atlas_cwv_mm.dat), where mm is the month to take into account.

Two formats for reading forecasts are available : GRIB (WMO standard format) and ASCII (described in annexe 1). The extracted forecast values at the pixel position are the 2 meters air temperature (K*10) over land, the total water vapor content (g/cm^2*100), the surface pressure, the air temperature at three standard levels and the geometric altitude. The twvc is computed inside the software from the forecast temperature and humidity profiles, when the twvc is not directly available in the forecast file.

Over sea, the forecast is only used to determine the twvc when it is not available in input argument. Over land, the forecast is only used to determine the air surface temperature and the twvc. If the forecast is not available, the software determines the twvc from climatology and for land conditions uses the observed T11mm instead of the air surface temperature to compute the IR threshold. To resume, MAIA does not go on exit if the forecast is not available, is still accurate over oceans but the accuracy is highly degradated over land for some surface conditions (cold surface temperature during winter) or cloud type (warm clouds).

When a situation is flagged cloudy, a further process is done to determine its cloud type. The input AVHRR channels vector goes through a classification tests sequence governed by its illumination (day, night, dawn), with the same philosophy for computing the thresholds.

Ten cloud categories are defined :

- five opaque cloud classes according to their altitude: very low, low, medium, high and very high
- three semi-transparent classes according to their thickness: thick, mean and thin
- one class of semi-transparent clouds above lower clouds
- one fractional clouds class

The sotware is written in fortran90.

2. CLOUD MASK INPUT/OUTPUT ARGUMENTS

2.1 INPUT

The input data are the following:

- debug value (0- no prints, 1, 2)
- new_bg: logical which allows the software to determine the environmental conditions (background) at a different spatial resolution than the pixel
- geometry definition (latitude, longitude, local zenith and azimuth angles, solar zenith and azimuth angles)
- surface altitude (optional). miss data= -9
- surface type : sea, mixed, land (optional) miss data= -9999.
- observed twvc (optional) miss data= -9999.
- albedo/brightness temperatures of the 5 AVHRR channels representative of the situation (full LAC data, cluster analysis outputs..).
- local variances (standard deviation of channels 1, 2, 4, 4-3 and local max values with neighbours for channels 1 and 4) representative of the situation
- name of a set_up file containing the adress and name of all input files

2.2 OUTPUT

A chart of 9 parameters (maia_par) :

1. clear /cloudy/snow/ice flag
2. skin surface temperature from AVHRR split-window (K*100) for clear pixels
3. background surface temperature from climatological SST (sea) or forecast T2m (land) (K*100)
4. CWV used: from input aurgument, forecast or climatology (degK*100)
5. surface altitude (m)
6. surface type
7. cloud type
8. black body flag
9. top cloud temperature (degK*100) for black body clouds

2.3 SET-UP FILE

The name of this file is an input argument of the maia.f90 subroutine. The file contains the adress and the names of all the files read by the library. It is read by the maia_setup.f90 subroutine at the first call of the library with the format:

read(70,'(a13,1x,a80)') maia_id, file_name

An example is given for Noaa14:

```
TH45sea_cold     t108t120_ocean_+3:+3_noaa14.dta
TH45sea_warm     t108t120_ocean_+0:+0_noaa14.dta
TH4sea_cold      tsurt108_ocean_+0:+0_noaa14.dta
TH43sea_cold     t108t038_ocean_+3:+3_noaa14.dta
TH34sea_cold     t038t108_ocean_+0:+0_noaa14.dta
TH34sea_warm     t038t108_ocean_-3:-3_noaa14.dta
TH43_ln          t108t038_veget_+3:+3_noaa14.dta
TH45_ld          t108t120_veget_-10:-10_noaa14.dta
TH4_ln           tsurt108_veget_+5:+5_noaa14.dta
TH4_ld           tsurt108_veget_-5:-5_noaa14.dta
TH43_land        t108t038_veget_+5:+5_noaa14.dta
```

TH43_desert	t108t038_desert_+5:+5_noaa14.dta
TH34_land	t038t108_veget_+3:+3_noaa14.dta
TH34_desert	t038t108_veget_+0:+0_noaa14.dta
TH34_cloud	t038t108_nuage_+0:+0_noaa14.dta
TH45_cloud	t108t120_nuage_+0:+0_noaa14.dta
NOISE	avhrr_noise_noaa14.txt
ALBEDO	atlas_albedo_06.dat
SST	atlas_sst_06.dat
CWV	atlas_wv_06.dat
FORECAST	maia2_forecast_file
TSAVHRR	tskin_noaa14.txt
MAPBITLS	mapbitls.dat
MAPTOPOG	maptopog.dat

3. EXTERNAL GEOGRAPHICAL INFORMATION

3.1 SEA SURFACE TEMPERATURE ATLAS

Global and monthly climatological data-sets of SST at a resolution of 0.15 * 0.15 degrees are provided with the scheme. Each of the 12 binary files contains the minimum of the monthly SST based on NOAA-07, NOAA-09 and NOAA-11 AVHRR imagery on the grid. Unit of SST is Celsius*100. Information over land exists in the files but is not used. There size is of 5.8GB each. They are of the form atlas_sst_mm.dat, mm=month.

The information of the file for the concerned month is fully read by the maia_clim_sst.f90 subroutine invoked at the first call of MAIA. The sst_clim value at each IASI spot is the nearest value in latitude, longitude of the spot position (maia_init.f90 routine).

3.2 SURFACE VISIBLE REFLECTIVITY ATLAS

Global and monthly climatological data-sets of visible reflectivity (AVHRR channel 1) at a resolution of 0.15 * 0.15 degrees are provided with the scheme. Each of the 12 binary files contains the mean of the monthly visible reflectivity on the grid over land only. Unit of visible reflectivity is %*100. More information on the data sets are given in **[2].** The size of the files is of 5.8GB each. They are of the form atlas_alb_mm.dat.

The information of the file for the concerned month is fully read by the maia_clim_alb.f90 subroutine. The alb_clim value at each spot is the nearest value in latitude, longitude of the spot position (maia_init.f90 routine).

3.3 SPECIFIC HUMIDITY ATLAS

Global and monthly climatological data-sets of specific humidity atlas and the corresponding surface pressure at a resolution of 2.5*5.0 degrees (lat, lon) are provided with the scheme. Each of the 12 binary files contains the mean of the montly specific humidity on 11 atmospheric levels and surface pressure on the grid over land and sea. The specific humidity is in g/kg*100 and the surface pressure in mb. More information on the data sets are given in **[3].** The size of the files is of 1.3GB each. They are of the form atlas_cwv_mm.dat.

The information of the file for the concerned month is fully read by the maia_clim_cwv.f90 subroutineinvoked at the first call of MAIA. The cwv_clim value is the interpolation of the specific humidity profiles at the 4 nearest grid nodes to the pixel position (maia_init.f90 routine). The interpolated profile is then converted to total water vapor content (maia_twvc.f90).

3.4 SURFACE TYPE AND ELEVATION ATLAS

Two datasets: mapbitls.dat for the surface type (flag sea/land) and maptopog.dat for the surface elevation (in meter) from the AAPP package **[4]** are provided with the MAIA scheme. They are at a 1/6*1/6 degrees resolution. If one of this information is not given by the calling program, they allow to continue running MAIA but ofcourse users are highly recommended to determine the surface type with a better resolution atlas (ex: IGBP atlas at the AVHRR pixel resolution). If the main program provides the surface type but not the surface elevation, only the elevation is computed inside the MAIA routine.

3.5 FORECAST

When available, the forecast file is read by the scheme, in ASCII (with lec_previ_ascii.f90) or GRIB (with lec_previ_grib.f90) format. The extracted fields are :

- the geometric altitude at the nodes of the forecast field, necessary for interpolating the forecast air surface temperature to the pixel position.
- the surface pressure
- the air surface temperature (2m altitude).
- the forecast profile (T, H, P) on atmospheric levels. The level pressures are fixed in the lec_previ.f90 routine. The field of total water vapor content is computed from the forecast profile and the surface pressure (lec_previ.f90).

T2m, air temperatures at 500, 700, 850hPa and TWVC values are interpolated at the input position from the values of the 4 nearest grid nodes (maia_init.f90 routine). The routine maia_init.f90 makes use of the surface elevation of the input situation to determine the air surface temperature at its position.

4. DETERMINATION OF ENVIRONMENTAL CONDITIONS

4.1 FIRST CALL OF MAIA

At the first call of the MAIA routine, the software reads:

- the set_up file which name is given as input argument
- the 3 climatological global atlas of SST, visible reflectivity and specific humidity. Fields are put in memory. Names of the files are given in the set_up file
- the forecast fields of T2m, TWVC, surface pressure and temperature and humidity profiles are put in memory. Name of the forecast file is given in the set_up file
- the coefficients to compute the skin surface temperature for the processed satellite

The logical “first” is then put to .false.

4.2 NEW ENVIRONMENTAL CONDITIONS

In order to save computation time, it is possible to determine the environment conditions at a different spatial resolution than the input observation. For example, the SST atlas is at a coarse resolution (0.15*0.15 degrees) compared to the AVHRR LAC resolution and the interpolation in the atlas for each AVHRR pixel will not improve significantly the result. It is the role of the input ‘new_bg’ logical argument to force the determination of the following parameters at the observation position for this specific call:

- surface type and/or surface elevation when the information is not available from the main program: two atlas of surface type and surface elevation at a 1/6 *1/6 degrees resolution (extracted from the AAPP package) are provided with the MAIA software.
- climatological SST, visible reflectivity and total water content data
- forecast data (T2m, TWVC, T500hPa, T700hPa, T850hPa, surface elevation)
- secant of the solar zenith angle
- threshold values depending on the previous information

These information are saved for all the following calls until the “new_bg’ parameter is put to .false. Of course, the user can systematiquely put “new_bg” to .true. and the processes are done at each call.

4.3 GEOGRAPHICAL INFORMATION SELECTION

The input climatologies and forecast data are used to compute in-line adapted thresholds to the geographical position of the pixel. Some priorities are given to the geographical information depending on the surface type and the data available. One of the purpose is that the software can be run even without forecast data. This section summarizes the priorities.

4.3.1 Surface temperature

Over sea: the climatological SST (from the monthly SST climatologies) is always used.

Over land: when the forecast is available, the forecast air surface temperature is used. Otherwise, the T11µminput observation.

Remark: the surface temperatures over land archived inside the SST climatologies are never used.

4.3.2 Visible reflectivity

Over sea: the visible reflectance is computed with a theoretical model (in phulpin.f90) and no input external data is used.

Over land, the visible reflectance comes from the monthly visible reflectance climatologies provided with the software.

4.3.3 Total water vapor

When no input TWVC is available from the calling program from a previous process (ex: AMSUA) :

- the forecast TWVC is used.
- If forecast data are not available, the TWVC is extracted from the specific humidity atlas. This makes possible to run the software in all cases, even without forecast data. However, the total water vapor content is a very variable information and a monthly mean value in place of the real information, has to be taken with caution in wet tropical area. See for example the variation of the T3-T5 threshold with TWVC (figure 1).

5. THRESHOLDS DETERMINATION

Terminology: in the software, all the fixed thresholds or offset values in hard inside are named cst_*. The computed values from interpolation inside thresholds tables are named s_*

5.1 CONSTANT THRESHOLDS OR OFFSET VALUES

There are often 2 different values for each threshold, one for the sea (s for sea) and one for the land and coast. Unit is K*100 or degrees*100 when used as the difference of 2 temperatures.

THRESHOLDS FOR DYNAMIC INTERPOLATION IN TABLES:

cst_tempcold = 278.	cst_tempwarm = 288.
cst_sstcold = 278.	cst_sstwarm = 288.
cst_solnight = 90.	cst_solhigh = 60.
cst_veg_max_alb = 20.1	cst_des_min_alb = 30.1
cst_adiarate =0.006	

IR THRESHOLDS:

cst_ir =1000	! offset for IR threshold ; land /night
cst_irld =700	! offset for IR threshold ; land /day
cst_sst =400	! offset for SST threshold ; sea
cst_irs =300	! offset for IR threshold ; sea. When there is no climatologic SST

UNIFORMITY:

cst_sd4s =20.	! for local standard deviations, T4 ; sea
cst_sd43s =50.	! same for T4-T3 ; sea
cst_sda2 =50.	! same for A2 ; always sea
cst_sd4l =100.	! same for T4 ; land
cst_sd43l =100.	! same for T4-T3 ; land
mountain =1500	! to determine mountain

SNOW/ ICE DETECTION:

cst_ice4 =500	! to calculate IR threshold
cst_ice34 =1500	! to calculate 3-4
cst_2shadow =2000	! to determine cloud shadow with A2
cst_ice45 =200	! maximum t4-t5 to determine snow

CLIMATOLOGICAL THRESHOLDS:

cst_lstmin =26315	! minimum land surface temperature
cst_sstmin =27315	! minimum sea surface temperature

CHANNEL 3 IDENTIFICATION:

separ1637 =18000	! to know if 1.6µm or 3.7µm is available

VEGETATION/DESERT INDEXES:

cst_desertalbmin =2500	! minimum visible reflectance over desert
cst_rlivthr =-0.1	! constant used over desert
cst_rsivthr =-0.04	! constant used over land

THRESHOLDS IF SUNGLINT:

cst_glint2 =1500	! visible threshold
cst_glint34 =500	! threshold for t3-t4 difference
cst_glint2coef =1.7	! coefficient for visible threshold

BLACK BODY THRESHOLDS:

bb45thr =100	! black body threshold

5.2 THRESHOLD FILES

Sixteen thresholds files per satellite are given with the scheme and are used to determine the thresholds depending of the measurement and environmental conditions (total water vapor content, surface temperature, secant of the zenith angle and solar zenith angle). The files have been computed off-line for a specific satellite by using the ECMWF climatological data-set **[5]** and RTTOV6 [**6**]. They are of the form:

T108t120_ocean _-3:+3_noaaxx.dta
T038t108_veget_+0:+0_noaaxx.dta…

In the present version, files for Noaa14 to17 are provided. New files must be created for future satellite.

They are fully read with the routine iniseuil.f90 to fill 16 tables (logical unit 70):

tab45_sea_cold	tab45_sea_warm	tabts4_sea_cold	tabts4_sea_warm
tab43_sea_cold	tab43_sea_warm	tab34_sea_cold	tab34_sea_warm
tab45_ln	tab45_ld	tabts4_ln_veg	tabts4_ld_veg
tabts4_ln_des	tabts4_ld_des		
tab43_l_cold_veg	tab43_l_warm_veg	tab43_l_cold_des	tab43_l_warm_des
tab34_l_cold_veg	tab34_l_warm_veg	tab34_l_cold_des	tab34_l_warm_des
tab34_tab_opaq	tab45_tab_opaq		

The method of threshold files creation is described on chapter 10. An example of threshold table is given on next figure.

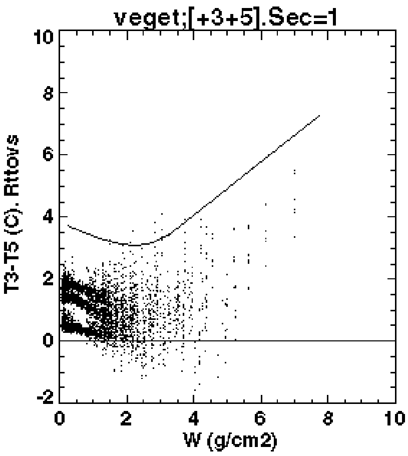


Figure 1 : Threshold 3-5 over land, satsec = 1

We can see that the threshold is really dependent of the TWVC. This one changes if the climate is tropical, temperate, polar or other. For example, if TWVC is equal to 7 g/cm^2, then the T3-T5 threshold value is 6.5 C and for TWVC equal to 3 g/cm^2 , the threshold is 3.5 C.

5.3 THRESHOLDS COMPUTATION

Threshold values for the pixel conditions are computed through interpolation in the previous tables by the valseuil_sea.f90, valseuil_land.f90 and valseuil_opaq.f90 routines. The interpolation is done inside a specific table with the background TWVC and the secant of the zenith angle. Interpolations are also done between tables function of the background surface temperature (sst_clim over sea, t2m forecast over land) and the solar zenith angle.

5.3.1 thresholds for infra-red channels

Over sea, s_45sea, s_sst, s_43sea, s_34sea are computed from tables:

tab45_sea_cold	tab45_sea_warm	tabts4_sea_cold	tabts4_sea_warm
tab43_sea_cold	tab43_sea_warm	tab34_sea_cold	tab34_sea_warm

Over land, s_45land, s_lst, s_43land, s_34land are computed from tables:

tab45_ln	tab45_ld	tabts4_ln_veg	tabts4_ld_veg
tabts4_ln_des	tabts4_ld_des	tab43_l_cold_veg	
tab43_l_warm_veg	tab43_l_cold_des	tab43_l_warm_des	tab34_l_cold_veg

tab34_l_warm_veg tab34_l_cold_des tab34_l_warm_des

For cloudy data s45_opaq,s34_opaq are computed from tables:

tab34_tab_opaq tab45_tab_opaq

5.3.2 thresholds for visible channels

Visible thresholds s_a1, s_a2, s_a3a are determined with the measurement conditions only (geometry angles). Different models are used over sea (albmer.f90) and land (albter.f90) .

The subroutine albmer.f90 simulates the visible reflectances over sea as function of the geometry angles. For the input channels 1 and 3a, the maximum reflectance (albsans) over the sea surface is computed with the subroutine phulpin.f90, following the Cox and Munck theory. Then, the simulatmos_ocean_35.f90 subroutine returns a0, a1 and a2 values for correcting the surface albedo of the atmospheric diffusion over sea.

The visible reflectance is then computed as follow:

s_a1 (s_a3a) = cosol * 100 * (brdf6s + offset) where : brdf6s = 100.*(a0+a1*surface/(1-surface*a2))
surface = (albsans/100.) / 100.
offset=5.
cosol= cos(θs) θs= zenithal angle

If the value is negative, the threshold is put to 1500* cosol

The albter.f90 subroutine simulates the visible reflectances over land in a very similar way than does albmer.f90 for sea. For the input channels 1 and 2, the surface reflectance over the land surface is computed with the subroutine roujean.f90 and depends on the climatological albedo (alb_clim) and the geometry angles (it applies an angle correction to the climatological albedo). Then, the simulatmos_terre_35.f90 subroutine returns a0, a1 and a2 values for the atmospheric diffusion over land.

The visible reflectance is then computed as follow:

s_a1 (s_a2) = cosol *100*brdf6s where: brdf6s = 100.*(a0+a1*surface/(1.-surface*a2))
surface = (alb_clim/100. * brdf6s +offset)/100.
offset = 8.

If the value is negative, the threshold is put to 1500* cosol

The sn16 threshold concerns the detection of the snow and ice. It is computed by the subroutine seuil16neige.f90 which determinesthe theoretical reflectance in channel 3 (for 1.6μm only) over snow for the input geometric angles. It is done in a very similar way than for visible channel over land. First, the surface theorical reflectance albsans over the snow surface is computed with the subroutine leroux.f90 and depends on the geometry angles (it applies an angle correction to a theoretical reflectance). Then, the simulatmos_terre_35.f90 subroutine returns a0, a1 and a2 values for the atmospheric diffusion over land.

The channel 1.6μm reflectance is then computed as follow:

seuil16neige = cosol*100*(brdf6s+offset) where brdf6s = 100.*(a0+a1*surface/(1.surface*a2))
surface = (albsans/100.)/100.
offset=5.

6. CLOUD DETECTION TESTS

See for more details the reference document **[1]**, for the scientific explaination of the tests and results of simulations.

Channels 3, 4 and 5 (BTs) are in K*100 and channels 1 and 2 (albedo) are in %*100.

6.1 T4 TEST

channel 11µm.

Its aim is to detect low temperature pixels corresponding to medium or high clouds. The main problem encountered is the threshold definition, which must be as high as possible to make the test efficient. It is very important during night-time or at dawn for detecting medium clouds.

Over ocean, a local SST value is computed using the split-window algorithm (tempsurfm.f90) with the pixel observations in input and is compared to the background climatological SST. Four degrees Kelvin (cst_sst) are then subtracted to account for water vapour absorption and imperfection of climatology.

Over land, radiative AVHRR temperatures are estimated using air surface temperature forecast from the most recent forecasts (no more than 48 hour of delay). Each T2m of the 4 nearest modes are first reduced from the grid altitude to the sea surface (alt = 0) T2m by using the geometric altitude of the forecast nodes. 3 steps are done :

- reduction of the T2m of the nodes at alt=0
- interpolation of T2m at the pixel position
- computation of T2m at the correct altitude (use of the pixel elevation)

The threshold s_ir is computed from an interpolation in threshold tables (s_lst) with the background forecast value in input. The quality of this test mainly depends on the forecast quality and the real vertical structure of the atmosphere (a dry adiabatic law is used to account for the height effect), but also on the land type (emissivity effect, presence of snow), the solar conditions (radiative cooling during night-time or maximum warming at around 14 h solar local time depending on the land type), and the delay between the time of AVHRR observations and the time of the forecast field.

If no forecast temperatures are available, the T4 threshold is a constant value (cst_ir during night, cst_irld for daytime periods) and the background surface temperature is the maximum value between the observed T4 and a cst_lstmin value. This maximum T4 brightness temperature is supposed to represent cloud-free surface temperature which desactives the test for most cases.

Over sea, when the climatologic SST < 1.8 °C, sea ice is expected with sea ice temperature as cold as land, and the land tests will be used with specific thresholds (s_sst, cst_sstmin).

6.2 T4-T5 TEST

T11µm - T12µm.

This test is applied to detect cirrus clouds and cloud edges, which are characterized by higher T11µm - T12µm brightness temperature differences than cloud-free surfaces. Channel 4 and 5 radiation absorptions are affected by the water vapor continuum and lines (of course different for the two channels), by CO_2 lines, surface temperature and emissivities. As a result, T4-T5 depends on the TWVC in the atmosphere and of the type of surface. This test will be useless if the estimated clear-sky T4-T5 is too high, which may be the case at daytime.

The threshold used is calculated from thresholds tables which depends on satellite zenith angle and total water vapor content. Interpolation is done between tables to the correct surface temperature and solar zenith angle. Over coast, the threshold used is the maximum value between computed values over sea and land.

6.3 T4-T3 TEST DURING NIGHT-TIME

T11µm - T3.7µm

This test is used to detect low water clouds. It is only applied during night-time, since it assumes that the 3.7µm infrared channel is not affected by the solar irradiance. Its efficiency is based on the spectral variation of the water clouds emissivity, which is lower at 3.7µm than at 11µm. The T4-T3 brightness temperature difference is large for small water-particle clouds, whereas continental or oceanic surfaces (except the sandy deserts in Africa) have similar brightness temperatures in the two channels.

The threshold is calculated from thresholds tables. Over sea, tab43_sea_cold and tab43_sea_warm are used ; over land for small background albedos (less than 20%), tab43_l_cold_veg and tab43_l_warm_veg are used, otherwise tab43_l_cold_des and tab43_l_warm_des are used. Over coast, the threshold used is the minimum value between computed values over sea and land.

6.4 T4-T3 TEST DURING DAYTIME

The T3.7µm -T11µm is used to detect shadows over low clouds. A basic assumption is that the T3.7µm is not affected by solar irradiance which should be the case in shadows. In this version, the thresholds are the same than for night-time.

6.5 T3-T4 TEST DURING NIGHT-TIME

T3.7µm - T11µm

This test is only applied at night-time to detect semi-transparent ice clouds or subpixel cold clouds. It is based on the fact that the contribution to the brightness temperature of relatively warm ground is higher at 3.7µm than at 11µm, due to the lower transmittance of ice cloud and to the high non linearity of the Plank function at 3.7µm. The brightness temperature difference T3-T4 is a function of the cloud height, thickness (for cirrus) and cloudiness (for subpixel clouds).

The thresholds are calculated from tables tab34_sea_cold and tab34_sea_land over sea ; over land from tab34_l_cold_veg and tab34_l_warm_veg for small background albedos , otherwise from tab34_l_cold_des and tab34_l_warm_des. Over coast, the threshold used is the maximum value between computed values over sea and land.

6.6 LOCAL UNIFORMITY TESTS

The tests are used to detect cloud edges, thin cirrus and small cumulus, by using their high spatial variations in the visible, near infra-red or infrared channels. For each AVHRR pixel, a local standard deviation for channels 1, 2, 4, 4-3 and local max values with neighbours for channels 2, 4 computed in a small box of 3x3 pixels centered on the pixel are input arguments of the MAIA routine. Subroutines sd_box.f90 and dtmax_box.f90 are provided with the sofware for computing these values. If this value is higher than a threshold corresponding to the ground heterogeneity, function of the surface type and of the channel, the central pixel is said to be cloudy.

This test is applied during daytime and night-time over land and oceans.

The thresholds are constant values (cst_sd*) depending of the surface type and channel.

6.7 A1 OR A2 TEST

Visible channel 0.6µm (A1) and near-infrared channel 0.9µm (A2)

This test, applied to the visible or near-infrared channel, is very useful in detecting low clouds, having a reflectance higher than the underlying surface.

Over sea, the reflectance measured over the oceans corresponds mainly to Rayleigh and aerosol scattering (if no sunglint is assumed), weaker in the near-infrared (0.9µm) than in the visible (0.6µm) and to the solar reflection. The near-infrared channel 2 reflectances are used, as they are less sensitive to aerosol and molecular scattering effects (Rayleigh) than the visible. High reflectances are mainly due to clouds. Sea surfaces have a low reflectance. All pixels with reflectances above a reflectance threshold are assumed to be cloud-contaminated. The threshold is a theorical value depending on the measurement conditions and is computed with the subroutine albmer.f90 which calls phulpin.f90 and simulatmos_ocean_35.f90.

Over land, channel 1 reflectances are used since the reflectance of land surfaces in channel 1 is much less than in channel 2 due to the vegetation spectral radiance behavior at these two wavelengths. This increases the contrast between land and cloud. The threshold computation determines the cloud-free land reflectance depending on the atmosphere (scattering and absorption) and on the land cover, but also on the viewing geometry. This calculation is made by the routine albter.f90 which calls roujean.f90 and simulatmos_terre_35.f90 subroutines. It is based on values from monthly reflectance climatologies (albedo), on which directional and atmospheric corrections are applied.

6.8 A2 AND T3-T4 TESTS IN SUNGLINT CONDITIONS

A2 and T3-T4

The near-infrared test over the oceans is usually very efficient, except inside sunglint areas. Thus, it is not possible to simply used the near infrared channel, since the ocean reflectance can reach very high values. Low cloud detection is then a problem.

The portion of the AVHRR passes that may be affected by this phenomenon can be determined by simple geometrical considerations on the sun and satellite respective positions (glint.f90).

It has been shown that simultaneous use of the near-infrared (A2/0.9µm) and medium infrared (T3/3.7µm) channels allows the detection of low clouds even in case of specular reflection. In fact, the solar reflection at 3.7µm, approximated by T3-T4 brightness temperature difference, is much higher over ocean in the case of specular reflection than over clouds for a given 0.9µm reflectance.

Thresholds used are constant values (cst_glint34, cst_glint2, cst_glint2coeff) normalized by the solar zenith angle.

6.9 SNOW AND ICE DETECTION UNDER CLOUD-FREE CONDITIONS DURING THE DAY

The snow and ice tests are applied for daytime period only if the solar elevation is greater than 20 degrees (10 degrees for 1.6µm), since it relies on the analysis of the solar reflection in the visible (0.6µm) and the medium-infrared (3.7µm, 1.6µm) wavelengths. It is mainly based on the fact that cloud-free snow reflects sunlight relatively weakly at 3.7µm and has high reflectance at 0.6µm, whereas water clouds have relatively high reflectance in both channels.

The 3.7µm channel measurements includes solar reflection and thermal emission. The solar reflection part can be roughly approximated by the T3-T4 brightness temperature difference. It can be also computed when assuming that the surface does not show transmittance at 3.7µm and has emissivity at 11µm equal to 1. In fact, distinction between clouds and snow is better performed using the brightness temperature difference.

Rayleigh scattering is more important at 0.6µm than at 0.9µm and is negligible at 3.7µm. Consequently, shadows of high clouds over low clouds are characterized by no solar reflection at 3.7µm, but relatively high visible reflectances and can therefore be confused with snow. However, they can be distinguished from snow by their low near-infrared reflectance and their near-infrared to visible reflectance ratio smaller than 0.75.

Moreover, some clouds (cirrus and stratus or stratocumulus) may also have a relatively small T3-T4 brightness temperature difference. Cirrus can be easily detected by their high T4-T5 difference, but remaining stratus or stratocumulus are very difficult to identify. The only remaining possibility is using the near infrared channel in the scheme. However, even the use of near-infrared channel, together with the channels at 0.6μm, 3.7μm and 11μm, does not allow a perfect separation of clouds from snow.

7. TESTS SERIES FOR CLOUD DETECTION

General comment: the test T3<separ1637 is used to discriminate the two 1.6μm and 3.7μm channels.

The values of all the thresholds are given in K*100 and %*100.

7.1 INPUT ARGUMENTS

- Over land:

  tsurf : t2m from forecast or if not available max value (cst_lstmin, T4)
  s_ir: s_lst or cst_irld if the forecast is not available
  s_45: s_45land
  s_43: s_43land
  s_34: s_34land

- Over sea:

  tsurf : sst_clim
  or over cold surfaces forecast t2m or if not available max value (cst_sstmin, T4obs)
  s_ir: s_sst or cst_irs if the forecast is not available
  s_45: s_45sea
  s_43: s_43sea
  s_34: s_34sea

- Over coast:

  tsurf : min(sst_clim, tsurf) with tsurf: t2m from forecast or if not available max value (cst_lstmin, T4obs)
  s_ir: min(s_lst, s_sst)
  s_45: max(s_45sea, s_45land)
  s_43: min(s_43_sea,s_43land)
  s_34: max(s_34sea,s_34land)

7.2 SNOW AND ICE DETECTION

These tests are only applied during the day. Snow detection is determined by a succession of tests with the test_snow.f90 routine. If snow is detected, the pixel is said to be 'clear'. This test is only applied over land. A similar test exists to detect ice over sea during day with the test_ice.f90 routine.

Snow detection	**Ice detection**
A3a and θs ≤ 80 deg. : If T4 < 27715 and T4 ≥ tsurf-s_ir - cst_ice4 and A1 > s_a1 and A2 > cst_2shadow*cosol and T3 < sn16 and T4-T5 < cst_ice45 ⇒ **Snow**	A3a and θs ≤ 80 deg. : If: sst_clim < 27715 and T4 < 27715 T4 ≥ tsurf-s_ir - cst_ice4 and T4-T5 < cst_ice45 and A1 > s_a1 and A2 > cst_2shadow*cosol ⇒ **Ice**
A3b and θs ≤ 70 deg: If T3-T4 < cst_ice34*cosol and T4 ≤ 27715 and T4 ≥ tsurf-s_ir - cst_ice4 and T4-T5 < cst_ice45 and A1 >s_a1 and A2 > cst_2shadow*cosol ⇒ **Snow**	A3b and θs ≤ 70 deg: If: sst_clim < 27715 and T3-T4 < cst_ice34 and T4 ≤ 27715 and T4 ≥ tsurf-s_ir - cst_ice4 and T4-T5 < cst_ice45 and A1 >s_a1 and A2 > cst_2shadow*cosol ⇒ **Ice**

Table 7-1: Succession of tests to detect snow only applied over land and sea ice.

With: cosol= cos(θs) θs= zenithal angle

7.3 TESTS APPLIED DURING DAYTIME WITHOUT SUNGLINT

These tests are done in testsd.f90 (sea & day), testld.f90 (land & day) and testcd.f90 (coast & day) routines.

Land	Sea	Coast	
If snow ⇒ **clear**	If ice ⇒ **clear**	If snow ⇒ **clear**	
Else	Else	Index ≥ Rivthr ⇒ Land	Index < Rivthr ⇒ Sea
Tsurf-T4 ≤ s_ir	sstclim- sstloc ≤ cst_sst or tsurf-T4 ≤ s_ir	Tsurf-T4 ≤ s_ir	sstclim- sstloc ≤ cst_sst or tsurf-T4 ≤ s_ir
A1 ≤ s_a1	A2 ≤ s_a2	A1 ≤ s_a1	A2 ≤ s_a2
	.not.A3a or A3≤ s_a3a		.not.A3a or A3≤ s_a3a
T4-T5 ≤ s_45	and T4-T5 ≤ s_45	and T4-T5 ≤ s_45land	and T4-T5 ≤ s_45sea
T4-T3 < s_43 or A3a or albedo_clim>cst_desertalbmin	T4-T3 < s_43 or A3a	T4-T3 < s_43land or A3a or albedo_clim>cst_desertalbmin	T4-T3 < s_43sea or A3a
Uniformity: (SD4 ≤ s_sd4 or SD43 ≤ s_sd43) and DT1≤fct(DT4/DT1)	Uniformity: (SD4≤cst_sd4 or SD43≤ cst_sd43) and (SD4≤cst_sd4 or SD2≤ cst_sd2)		
⇒ **clear**	⇒ **clear**	⇒ **clear**	⇒ **clear**

Index = (A2-A1)/(A2+A1)

Table 7-2: Daytime tests over land, sea and coast without sunglint

Rivthr: used to determine if we are over sea or land with the vegetable index. If the climatological albedo is greater than a cst_desertalbmin value, then rivthr = cst_rlivthr, else rivthr =cst_rsivthr.

sstloc= tempsurfm(T3, T4, T5, satsec, szen, sst_clim) see chapter 8.1

s_sd4 = cst_sd4l + 1.67* (90.- θs)

s_sd43 = cst_sd43l + 1.67* (90.- θs)

SD4, SD43: local variances channels 4, 4-3

DT1, DT4: local max values channels 1, 4

7.4 TESTS APPLIED DURING DAYTIME WITH SUNGLINT

Theses tests are only applied over sea and coast, in the testsg.f90 (sea, glint) and testcg.f90 (coast, glint). The determination of the sunglint condition is done in the glint.f90 routine, called by maia.f90 before calling any test*.f90 routines.

7.4.1 Specular reflection test

It determines if specular reflection is present by computing the specular reflection (Phulpin model) in the routine glint.f90 :

$$\mu_n = \frac{(\mu_{so} + \mu_{sa})}{\sqrt{2.*(1.+\sin(\theta_{so})*\sin(\theta_{sa})*\cos(\varphi_{sa} - \varphi_{so}) + (\mu_{so} * \mu_{sa}))}}$$

where :

$\mu_{so} = \cos(\theta_{so})$

$\theta_{so} = \theta_s * \pi / 180$ where θ_s is the solar zenith angle

$\mu_{sa} = \cos(\theta_{sa})$

$\theta_{sa} = \theta_l * \pi / 180$ where θ_l is the satellite zenith angle

$\varphi_{so} = \varphi_s * \pi / 180$ where φ_s is the solar azimuth angle

$\varphi_{sa} = \varphi_l * \pi / 180$ where φ_l is the satellite azimuth angle

If $\mu_n \geq 0.999 \Rightarrow$ **specular reflection**

Else if $\mu_n > 0.90$ then :

$$val = \frac{n\,\text{int}(200.*\exp(-1))}{4.*\mu_{sa} * \mu_{so} * abs(\mu_n^{\ 2} - \mu_n^{\ 4})}$$

If val > 1000 ⇒ **specular reflection**.

7.4.2 Tests series

Sea	coast	
If Ice ⇒ **clear**	If snow ⇒ **clear**	
Else	Index < Rivthr ⇒ Sea	Index ≥ Rivthr ⇒ Land
sstclim- sstloc ≤ cst_sst	sstclim- sstloc ≤ cst_sst	Tsurf-T4 ≤ s_ir
T4-T5 ≤ s_45	A2 ≤ cst_glint2*cosol or T3-T4≤cst_glint34*cosol or T3-T4 ≥ A2* cst_glint2coef or A3a	
Uniformity : SD4 ≤ cst_sd4 or SD43 ≤ cst_sd43		
A2 ≤ s_a2	A2 ≤ s_a2	A1 ≤ s_a1
A2 ≤ cst_glint2*cosol or T3-T4 ≤ cst_glint34*cosol or T3-T4 ≥A2*cst_glint2coef or A3a	T4-T5 ≤ s_45sea	T4-T5 ≤ s_45land
T4-T3 ≤ s_43 or A3a	T4-T3 < s_43sea or A3a	T4-T3 < s_43land or A3a or albedo_clim>cst_desertalbmin
⇒ **clear**	⇒ **clear**	⇒ **clear**

index = (A2-A1)/(A2+A1)

Table 7-3: Daytime tests over sea and coast with sunglint

7.5 TESTS APPLIED DURING NIGHT-TIME

They are done in the subroutines testsn.f90, testln.f90 and testcn.f90.

sea	land		coast
	A3a	.not. A3a	
T4 ≥ s_ir	Tsurf-T4 ≤ s_ir	Tsurf-T4 ≤ s_ir	Tsurf-T4 ≤ s_ir
T3-T4 ≤ s_34		T3-T4 ≤ s_34	A3a or T3-T4 ≤ s_34
T4-T3 ≤ s_43		T4-T3 < s_43 or albedo_clim>cst_desertalbmin	T4-T3 < s_43 or albedo_clim>cst_desertalbmin
Uniformity: SD4≤cst_sd33s	T4-T5 ≤ s_45	T4-T5 ≤ s_45	T4-T5 ≤ s_45
T4-T5 ≤ s_45	SD4 ≤ cst_sd4l or SD43≤ cst_sd43l	Montain or SD4≤ cst_sd4l or SD43≤ cst_sd43l or albedo_clim<cst_desertalbmin	
⇒ **clear**	⇒ **clear**	⇒ **clear**	⇒ **clear**

Table 7-4: Night-time tests over sea, land and coast.

7.6 TESTS APPLIED DURING TWILIGHT

They are done in the subroutines testst.f90, testlt.f90 and testct.f90.

Sea	Land	Coast
sstclim- sstloc ≤ cst_sst	Tsurf-T4 ≤ s_ir	Tsurf-T4 ≤ s_ir
T4-T5 ≤ s_45	T4-T5 ≤ s_45	T4-T5 ≤ s_45
Uniformity: SD4 ≤ s_sd4 or SD43 ≤ s_sd43	T4-T3 < s_43 or A3a or albedo_clim >cst_desertalbmin	T4-T3 < s_43 or A3a or albedo_clim > cst_glint2coef
T4-T3 < s_43	A1 ≤ s_a1	
A2 ≤ s_a2	SD4 ≤ s_sd4 or SD43 ≤ s_sd43	A2 ≤ s_a2 or A1 ≤ s_a1
.not. A3a or A3 ≤ s_a3a		
⇒ **clear**	⇒ **clear**	⇒ **clear**

Table 7-5: Tests applied during twilight.

8. AVHRR SURFACE AND CLOUD TOP TEMPERATURES DETERMINATION

8.1 SEA SURFACE TEMPERATURE

For each AVHRR situation declared clear by the test series and over sea, a surface temperature is computed with the following function from **[7]**, in the tempsurfm.f90 routines. The coefficients are read from a file dependant of the satellite.

for day (θs < 110 °):

tempsurfm = c_nl(1) *t4 + (c_nl(2)*tclim + c_nl(3)*steta)*(t4-t5) + c_nl(4) + b_nl

for night:

tempsurfm = (c_t37(1) + c_t37(2)*steta)*t3 + (c_t37(3) + c_t37(4)*steta)*(t4-t5) + c_t37(5)*steta + c_t37(6)+ b_t37

steta =(θs sec-1) with sec is the secant of the local zenith angle.

8.2 LAND SURFACE TEMPERATURE

For clear pixels over land the same process is applied as for sea conditions but with an other function and constant coefficients independent of the satellite, in the tempsurft.f90 routine :

tempsurft= t4+(1.31+0.27*(t4-t5))*(t4-t5)+1.16

This routine is not very efficient mainly over desert with possible radiative cooling during night or large warming in the afternoon.

8.3 CLOUD TOP TEMPERATURE

At the end of each testxx.f90 routine, for each AVHRR pixel rejected with one of the tests (then supposed to be cloudy), a step is applied to flag the cloud opaque or not (cornoir.f90 routine), by comparing T4-T5 to a threshold :

T4-T5 < bb45thr. With bb45thr=1K

If the situation is flagged cloudy and opaque (usually said black-body), a cloud top temperature is given which is put to the channel 4 input observation.The assumption is that there is no absorption by the atmosphere above the cloud, which of course is not true and this cloud top temperature should be considered with care in case of low cloud layers and corrected afterwards.

9. CLOUD TYPE DETERMINATION

When a situation is flagged cloudy, a further process is done to determine its cloud type. The input AVHRR channels vector goes through a classification tests sequence governed by its illumination (day, night, dawn).

Nine cloud categories are defined :

- five opaque cloud classes according to their altitude: very low, low, medium, high and very high
- three semi-transparent classes according to their thickness: thick, mean and thin
- one class of semi-transparent clouds above lower clouds
- one fractional clouds class

9.1 THRESHOLDS DETERMINATION

A set of specific thresholds adapted to the situation conditions is computed for the classification tests.

The infra-red thresholds (s_45opaq, s_34opaq) used to separate opaque clouds from semi-transparent clouds are computed through interpolation in the two tables tab34_tab_opaq and tab45_tab_opaq.

The visible thresholds (max06, min06), also used to discriminate between opaque and semi-transparent clouds, are computed using the measurement conditions and the surface temperature.

The temperature thresholds (maxt4lo,maxt4me, maxt4hi, maxt4vh) used to determine the height of the cloud, are computed using a regression with the forecast temperatures at standard levels (T500hPa, T700hPa, T850hPa).

9.2 CLASSIFICATION TESTS APPLIED DURING NIGHT

Very High Opaque	T4<maxt4vh	T4-T5 <t45_thick
High Opaque	maxt4vh <T4<maxt4hi	T4-T5 <t45_thick
Medium	maxt4hi<T4<maxt4me	T4-T5 <t45_thick (A3a) or T3-T4 <t34_thin (A3b)
Low	maxt4me<T4<maxt4lo	T4-T5 <s_45 (A3a) or T3-T4 <t34_thin (A3b)
Very Low	maxt4lo<T4<maxt4thin maxt4lo<T4<maxt4thin maxt4thin<T4	T4-T5 <t45_thick and T3-T4 >t34_thin T3-T4 <t34_thin and (T4-T5 <t45_thick or T3-T4 <t34_low) T4-T5 <s_45 or T3-T4 <t34_low
Fractional	maxt4lo<T4<maxt4thin maxt4lo<T4<maxt4thin maxt4thin<T4	T4-T5 >t45_thick and T3-T4 <t34_thin and T3-T4 >t34_low T4-T5 >t45_thick and T3-T4 >t34_thin T4-T5 >s_45 and T3-T4 >t34_low
Semi-transparent thick	T4<maxt4vh maxt4vh <T4<maxt4hi	T4-T5 >t45_thick T4-T5 >t45_thick
Semi-transparent mean	maxt4hi<T4<maxt4me	T4-T5 >t45_thick (A3a) or T3-T4 >t34_thin (A3b)
Semi-transparent thin	maxt4me<T4<maxt4lo	T4-T5 >s_45 (A3a) or T3-T4 > t34_thin (A3b)

With

t45_thick= s_45opaq if less than cst_45_opaq_max. Otherwise, it is put to cst_45_opaq_max*100

t34_thin = s_34opaq if less than cst_34_semi_max. Otherwise, it is put to cst_34_semi_max*100

t34_low = s_34opaq - cst_34_low_delta*100

maxt4_thin = maxt4lo +2*s_45

9.3 CLASSIFICATION TESTS APPLIED DURING DAY

Very High Opaque	T4<maxt4vh	T4-T5 <t45_thick	
High Opaque	maxt4vh <T4<maxt4hi	T4-T5 <t45_thick	
Medium	maxt4hi<T4<maxt4me	T4-T5 <t45_thick T4-T5 <t45_above	A1>max06 A1<max06
Low	maxt4me<T4<maxt4lo	T4-T5 <t45_above or Sdlog4 <Sdlog1	A1>max06
Very Low	maxt4lo<T4<maxt4thin maxt4thin<T4	T4-T5 <t45_edge T4-T5 <t45_edge	A1>max06 A1>min06
Fractional	maxt4me<T4<maxt4lo maxt4lo<T4<maxt4thin maxt4lo<T4<maxt4thin	T4-T5 <t45_above T4-T5 >t45_edge T4-T5 <t45_edge	A1<max06 A1>max06 A1<max06
	maxt4thin<T4	T4-T5 >t45_edge or A1< min06	
Semi-transparent thick	T4<maxt4vh maxt4vh <T4<maxt4hi	T4-T5 >t45_thick T4-T5 >t45_thick	
Semi-transparent mean	maxt4hi<T4<maxt4me	T4-T5 >t45_above	A1<max06
Semi-transparent thin	maxt4me<T4<maxt4lo maxt4me<T4<maxt4lo maxt4lo<T4<maxt4thin	T4-T5 >t45_above and Sdlog4 >Sdlog1 T4-T5 >t45_above T4-T5 >t45_edge	A1>max06 A1<max06 A1<max06
Semi-transparent above	maxt4hi<T4<maxt4me	T4-T5 >t45_above	A1>max06

With

t45_thick= s_45opaq if less than cst_45_opaq_max. Otherwise, it is put to cst_45_opaq_max

t45_edge = s_45

t45_above = t45_thick

maxt4_thin = maxt4lo +2*s_45

Sdlog4 = 100.*(15.*(log(1+sd33_t4/100.)))

Sdlog1 = 30.*(log(1+sd33_a1/100.))

9.4 CLASSIFICATION TESTS APPLIED DURING DAWN

Very High Opaque	T4<maxt4vh	T4-T5 <t45_thick
High Opaque	maxt4vh <T4<maxt4hi	T4-T5 <t45_thick
Medium	maxt4hi<T4<maxt4me	T4-T5 <t45_thick
Low	maxt4me<T4<maxt4lo	T4-T5 <s_45
Very Low	maxt4lo<T4<maxt4thin maxt4thin<T4	T4-T5 <t45_thick T4-T5 <s_45
Fractional	maxt4lo<T4<maxt4thin maxt4thin<T4	T4-T5> t45_thick T4-T5 >s_45
Semi-transparent thick	T4<maxt4vh maxt4vh <T4<maxt4hi	T4-T5 >t45_thick T4-T5 >t45_thick
Semi-transparent mean	maxt4hi<T4<maxt4me	T4-T5 >t45_thick
Semi-transparent thin	maxt4me<T4<maxt4lo	T4-T5 >s_45

With

t45_thick= s_45opaq if less than cst_45_opaq_max. Otherwise, it is put to cst_45_opaq_max*100

maxt4thin = maxt4lo + s_45

10. OFF-LINE THRESHOLD FILES CREATION

The threshold files were created off line to give the variation of the different threshold files (t4-t5, t4-t3, t3-t4, surface temperature-t4) with the secante of the zenith angle and the total water vapor content in the atmosphere, for different surface temperature, emissivity and solar elevation.

First, we used a sub-set of the ECMWF dataset (2995 profiles) which represents the global atmosphere. Each profile is documented with a profile code (sea, coast, land), its position and date, its total water vapor content.
The RTTOV6 fast forward model was used to compute synthetic brightness temperatures for AVHRR channels 3, 4 and 5, for 5 different secant angles (from 1 to 2 with a step of 0.25), 7 surface-air skin temperatures (-10,-5,-3,0,+3,+5,+10) and 60 emissivities from 0.8 to 1. (with 2 different steps of 0.005 and 0.0025). The brightness temperatures depend on the channel characteristics and consequently on the satellite number. The output file noaaxx_ecmwf3000.res is very large of about 500MB.

Secondly, sub-files of smaller size were extracted differently for the 4 conditions: sea/coast, vegetation, desert and cloud. Only profiles with the correct code are kept, and a selection in the range of emissivity is done. For vegetation, 3 sets of emissivities are considered and 2 over desert. Over sea, the emissivity depends on the secant. We get 4 independent files of much smaller size (between 3 and 10MB).

noaaxx_ocean_ecmwf.dta
noaaxx_veget_ecmwf.dta
noaaxx_desert_ecmwf.dta
noaaxx_nuage_ecmwf.dta

The extraction takes about 15 minutes for one satellite.

The 4 preceeding files are read by the crea_table_* routines, which create the tables of the channel differences for all the possible air-sol differences. The means and standard deviations for the secant angles, and 7 twvc (from 0.25 to 7.75) are computed and curves of maximum values of the channels differences are estimated by :

Channel difference= mean + 2*std +noise

The noise comes from statistics between RTTOV synthetic brightness temperatures and observations.

The resulting curves are then interpolated and extrapolated on 7 secant angles (from 1. to 2.5) and 16 twvc (from 0.25 to 7.75).

The results correspond to more than 100 files per satellite for t45, t43, t35, ts4 (sea, vegetation, desert) , several air-skin surface temperaure departures and three for cloud conditions. The resulting files are of the form: Ex: t108120_veget_-3 :+5_noaaxx.dta

In pratice, only sixteen of them are used and provided with the MAIA software :

t108t120_ocean_+3:+3_noaa14.dta	t108t120_ocean_+0:+0_noaa14.dta
tsurt108_ocean_+0:+0_noaa14.dta	t108t038_ocean_+3:+3_noaa14.dta
t038t108_ocean_+0:+0_noaa14.dta	t038t108_ocean_-3:-3_noaa14.dta
t108t038_veget_+3:+3_noaa14.dta	t108t120_veget_-10:-10_noaa14.dta
tsurt108_veget_+5:+5_noaa14.dta	tsurt108_veget_-5:-5_noaa14.dta
t108t038_veget_+5:+5_noaa14.dta	t108t038_desert_+5:+5_noaa14.dta
t038t108_veget_+3:+3_noaa14.dta	t038t108_veget_+0:+0_noaa14.dta
t038t108_nuage_+0:+0_noaa14.dta	t108t120_nuage_+0:+0_noaa14.dta

11. VALIDATION

11.1 TEST FILE DESCRIPTION

An estimate of the accuracy and limits of the cloud mask developed for NWC SAF applied to GOES data can be found in **[1]**. It used an interactive test file containing the geometrical and radiative characteristics of many small targets (20000) for GOES. A similar test file with the same information for AVHRR (7007 targets of 5x5 pixels over 3 years) was also created. Only the AVHRR center pixel is processed and the neighbour pixels are used to compute the local variances. Nearest NWP fields were collocated to each situation.This file was used to validate the previous version of the MAIA cloud mask **[8]**.

To summarize, the targets were manually flagged cloud free or contaminated (with an estimation of the type of clouds). 38 target categories have been manually identified by CMS nephanalysts. Each target is collocated with the nearest NWP forecast to compute the air surface temperature and total water vapor content.

The situations come from 3 different satellites: 342 targets for Noaa12, 2221 for Noaa14 and 4464 targets for Noaa15. Most of the Noaa15 data contains daily situations with a 1.6µm for channel 3. Figure 2 shows the spatial distribution of the targets forNoaa15. The different colors identify the 38 cloud categories. All situations have been extracted from data acquired at the CMS center and are situated over Europe. 20 % are for latitudes above 55N. The distribution of the targets with the date (month and hour of the day), with the surface temperature and the total water vapor content is given in figure 3.

Table 11.1 gives the distribution of the targets with the measurement conditions:

day over sea	twilight, sea	sunglint. sea	day, land	twilight,land	night , sea	night , land
42%	0.12%	9.6%	42.8%	0.17%	3%	2,3%

Table 11.1: distribution of the targets with the measurement conditions.

We can found 2 main weaknesses to the test file: it is not representative of night situations and of climates outside Europe.

This chapter presents the validation on this AVHRR test file of MAIA v3 cloud mask and classification.

The validation is done to access the accuracy of the main outputs of the routine: cloud flag, cloud type, black-body flag, skin surface temperature from the AVHRR split window and the cloud top temperature.

11.2 CLOUD FLAG ACCURACY

Table 11.2 illustrates the efficiency of the cloud mask depending on the illumination and surface conditions. As expected, the software is well detecting clouds for all conditions (about 99.3% over sea, 93.9% over land, all cloud types). Concerning the clear targets, 4.1% of them over sea and 6.1% over land are mis-classified as cloudy: these results are much better than in the previous MAIA version **[8]** and are very similar than in the GOES validation. That is due to the fact that in this version most of the thresholds are computed through interpolation with the viewing angle and total water vapor inside off-line computed tables or even interpolation between tables with the solar zenithal angle.

	Cloudy targets correctly detected	Cloud free targets correctly classified
sea, day 3.7µm	99.3% (294)	95.9% (219)
sea, day, 1,6µm	99.3% (1618)	95.5% (490)
sea, night	100% (65)	97.6% (125)
sea: twilight	91.4% (58)	92.9% (14)
land, day, 3.7µm	87.3% (221)	95% (179)
land, day 1.6µm	96.6% (742)	93.2% (556)
land night	97.1% (35)	95.2% (124)
land: twilight	68.2% (26)	100% (17)

Table 11.2: Overall cloud mask efficiency for the different illumination, surface conditions and separatly for channel 3a and 3b. Number of situations is under bracknet.

Table 3 illustrates the efficiency of the cloud mask and of the cloud classification for all illumination conditions. For an easiest read of the table, the numbers correspond to the nearest integer of the statistical scores in percent.

The main deficiences seen by the mask validation are the following:

- Clear targets over snow or ice with a solar zenithal angle larger than 70 degree have been discarded from the statistics, as no tests are done to detect snow or ice and these situations are systematically mis-classified as cloudy.
- A large number of clear sea with shadow appairs as fractional clouds due to the test on uniformity. Over sea, this threshold is very low and it is also possible that when computing the local variance using the neighbours, part of the cloud is taken into account.
- Low clouds shadowed by higher clouds may be not very well detected.
- Too thin cirrus over land or brocken clouds may be characterized as the surface they cover

Target type	*nb*	*clear* %	*snow* %	*ice* %	*cloud* %	*very low* %	*low*	*medium* %	*high* %	*very high* %	*str-thick* %	*str-mean* %	*str-thin* %	*str-above* %	*frac* %
open sea	**885**	**96**	0	0	4	2	0	0	0	0	0	0	0	0	1
sea with shadow	**23**	**43**	0	0	57	17	13	0	0	0	0	0	0	0	26
sea with sand or aerosol	**304**	23	0	0	77	72	5	0	0	0	0	0	0	1	0
sea with sunglint	**249**	**98**	0	0	2	2	2	0	0	0	0	0	0	0	0
land	**910**	**94**	0	0	6	4	1	0	0	0	0	0	0	0	1
land with shadow	**38**	**95**	0	0	5	0	3	0	0	0	0	0	0	0	3
land with sand or aerosol	**98**	44	0	0	56	47	2	3	0	0	0	1	3	0	0
ice (solzen < 70)	**13**	0	0	**85**	15	0	8	0	0	0	0	0	8	0	0
snow (solzen < 70)	**417**	1	**89**	0	10	4	4	2	0	0	0	0	0	0	0
stratus	**546**	7	0	0	**92**	59	30	3	0	0	0	0	0	0	0
statocumulus	**1009**	1	0	0	**99**	38	50	10	0	0	0	0	0	0	1
shadow over low clouds	**33**	15	0	0	**85**	36	30	15	3	0	0	0	0	0	0
small Cu ocer sea	**107**	4	0	0	**96**	78	10	0	0	0	0	2	0	2	5
small Cu over land	**57**	12	2	0	**86**	65	14	0	0	0	0	0	0	4	4
Cu congestus over sea	**48**	0	0	0	**100**	23	19	44	6	0	2	2	0	4	0
Cu congestus over land	**11**	0	0	0	**100**	0	36	36	18	0	0	0	0	9	0
cumulonimbus	**89**	0	0	0	**100**	0	0	7	69	24	0	1	0	0	0
extensive cumulonimbus	**48**	0	0	0	**100**	0	0	0	65	35	0	0	0	0	0
thin Ci over sea	**103**	3	0	0	**97**	4	11	4	0	0	0	7	22	32	17
thin Ci over land	**98**	6	1	0	**93**	5	13	2	0	0	1	10	18	37	6
thin Ci over snow	**35**	11	17	0	**71**	0	20	3	3	0	9	6	6	26	0
thin Ci over st/sc	**171**	0	0	0	**100**	2	8	33	4	0	2	2	1	46	1
thin ci over cu	**17**	0	0	0	**100**	6	6	41	24	0	0	0	0	24	0
thin ci over ac/as	**99**	0	0	0	**100**	0	1	43	26	0	10	1	0	18	0
altocumulus/altostratus	**135**	0	0	0	**100**	0	4	89	1	0	1	0	0	4	0
altocumulus	**119**	0	0	0	**100**	0	3	86	0	0	2	0	0	10	0
cirrostratus	**165**	0	1	0	**99**	0	1	10	30	1	32	1	0	24	0
cirrostratus over ac/as	**252**	0	0	0	**100**	0	0	3	78	9	9	0	0	2	0
cloudy (unknown)	**47**	4	0	0	**96**	0	15	47	19	0	2	0	0	0	0

Table 11.3: Overall cloud mask and cloud classification accuracy

The first colomn gives the list of cloud and earth types available in the AVHRR test file. The second colomn corresponds to the number of targets for each category and the following colomns are in %. Colomns 'clear', 'snow','ice' and 'cloudy' concerns the cloud mask flag output. The next colomns concerns the MAIA cloud description given when a cloud is found in the target.

Str stands for semi-transparent

11.3 CLOUD TYPE AND BLACK-BODY FLAG ACCURACY

Table 11.4 presents the efficiency of the cloud classification. First line concerns the accuracy of the cloud detection for the 10 cloud classes: a situation is considered well cloud classified when nephanalysts visualized clouds, sand or aerosols in the scene. Also, to access the accuracy of the classification, we have gathered the observed cloud descriptions in 3 meta-classes :

- stratus, stratocumulus, low clouds and cumulus in "low" opaque clouds,
- cumulonimbus, altocumulus and cirrostratus in "high"opaque clouds
- situations covered by cirrus in semi-transparents

The second line of table 4 presents the correlation between the MAIA classification and the meta classes. Most of the clouds are correctly classified with an agrrement of more than 90%, excepted to thin cirrus are sometimes not seen by the classification algorithm: some situations with thin cirrus over medium clouds are classified 'medium', some others (ex: over land) are classified fractional.

When a target is cloudy, a test is done to determine if the cloud is black-body or not. The last line is the table black-body flag occurrence. As expected, there is a good agreement between then black-body flag and the cloud type.

	very low	*low*	*medium*	*high*	*very high*	*str-thick*	*str-mean*	*str-thin*	*str-above*	*frac*
Nb	**1453**	**906**	**593**	**411**	**70**	**103**	**31**	**49**	**268**	**79**
correctly cloud classified (%)	95	97	98	100	100	100	100	98	100	57
correct "meta type" (%)	98	92	78	91	100	98	83	100	88	39
Black-body flagged (%)	90	95	93	97	96	11	13	9	8	7

Table 11.4:Statistics on the cloud type accuracy (first line), the agreement with meta-types (second line) and the black-body flag occurrence (last line) function of the MAIA cloud classification

11.4 SURFACE TEMPERATURE ACCURACY

An indirect way to validate the cloud mask and the retrieved surface temperature is to compute the brightness temperatures of the AVHRR channels 4 and 5 window channels for only the situations classified 'clear' and to compare them to the observations. If some cloudy situations remain in the computation, the departure between synthetic and observed Tbs is large.

The synthetic brightness temperatures are computed using the RTTOV6 forward radiative transfer model, the collocated NWP profile and the surface information. The surface temperature is defined by two ways to see the impact of this parameter. First we put the background surface temperature (left parts of tables 5 and 6) used by the cloud mask software – the climatological SST over sea or the forecast air temperature at the surface over land – and then we used the retrieved skin surface temperature from AVHRR split-window (right parts of tables). The surface emissivity is a function of the viewing angle over sea and a constant of 0.98 over land.

The synthetic brightness temperatures are compared to the observations. The departure statistics ofcourse include several extra incertities (from collocation, forecast profile, forward model, surface emissivity) but mainly depend on the determnination on the cloud mask flag and of the surface temperature. The results show a large improvement of the departure statistics when using the split-window surface temperature compared to the background surface temperature. The standard deviation is about 0.5K for channel 4 over sea and slightly larger over land due to the fact that the split window is not satellite dependant and that the surface emissivity is a constant. Nevetheless, the results are much better than when using the T2m forecast as surface temperature, showing the importance of a correct definition of this parameter.

	SST from climatology			*SST from AVHRR split-window*		
	Noaa12	***Noaa14***	***Noaa15***	***Noaa12***	***Noaa14***	***Noaa15***
N	59	271	449	59	271	449
A4 (bias/std)	0.87 / 1.7	-1.50 / 1.59	-0.95 / 1.26	0.55 / 0.94	0.11 / 0.57	-0.01 / 0.44
A5 (bias/std)	-0.28 / 1.82	-0.88 / 1.53	-0.73 / 1.25	0.92 / 1.28	0.45 / 0.78	0.11 / 0.61

Table 11.5: Sea Surface temperature accuracy. Statistics of the departure between synthetic brightness temperature and onservations

	Tsurf = T2m forecast			*Tsurf from AVHRR split-window*		
	Noaa12	***Noaa14***	***Noaa15***	***Noaa12***	***Noaa14***	***Noaa15***
N	40	263	554	40	263	554
A4 (bias/std)	4.86 / 2.55	-1.89 / 6.72	2.34 / 3.46	0.74 / 0.99	-0.03 / 1.15	-0.52 / 0.81
A5 (bias/std)	4.47 / 2.34	-1.54 / 6.15	2.17 / 3.08	1.07 / 1.37	0.032 / 1.63	0.51 / 1.22

Table 11.6: Land Surface temperature accuracy. Statistics of the departure between synthetic brightness temperature and observations

11.5 CLOUD TOP TEMPERATURE ACCURACY

A similar comparison is done for the cloudy opaque situations for which the black-body flag is on. For these cases, the software determines a cloud top temperature (in fact the observed AVHRR channel 4). The cloud top pressure is computed from the cloud top temperature and the collocated forecast profile. A correction is done for low-level clouds: when the cloud pressure is below 750hPa, the cloud top temperature is warmed by a 1K value and the cloud pressure is recomputed. This takes into account the atmosphere above low-level clouds.

The synthetic brightness temperature is computed by putting a blackbody cloud (cloud emissivity =1) at that pressure. The departure statistics are very good.

	Noaa12	***Noaa14***	***Noaa15***
N	49	318	1887
A4 (bias/std)	-0.46 / 0.48	-0.37 / 0.29	-0.51 / 0.38
A5 (bias/std)	-0.81 / 0.79	-0.21 / 0.63	-0.47 / 0.60

Table 11.7: cloud top temperature accuracy. Statistics of the departure between synthetic brightness temperature and observations

12. EXAMPLE ON A NOAA17 GLOBAL LAC REVOLUTION

This study was proposed in the framework of the verification of the EPS/CGS level 1 and 2 processing softwares **[9]**. It was necessary to generate realistic IASI spectra simulations for a complete orbit, consistent with the AVHRR measurements. For that purpose, a complete orbit (08/08/2002 ; 18h12 – 20h24) of Noaa17 AVHRR observations at full resolution (LAC data) was provided by Eumetsat and we run the MAIA cloud mask on these data at full AVHRR resolution, 1 pixel/10, 1 line/10. The nearest ECMWF NWP field (08/08/2002 18h00) was used as forecast file. Figures 4 shows the resulting MAIA classification.

Synthetic brightness temperatures were computed using the RTTOV6 forward radiative transfer model, the collocated NWP profile and the MAIA surface information, for all the clear and 'opaque cloud' situations, by the same way as described in chapter 11. The synthetic brightness temperatures are compared to the observations. This can't be considered as an absolute validation because the skin surface temperature and the cloud top temperature are determined from the observations but it is an easy way to focus on area of possible problems.

Figure 5 shows the map of the differences. With the choiced color table, departures less than two standard deviations (expected from the previous validation with targets) appears white over sea and light blue, white and yellow over land. It is mostly the case everywhere.

Figures 6 show a zoom over the tropical Pacific, where the largest departures are found over sea. Cirrus edges (figure 6.B) were not well detected and gives large positive differences for some few pixels (in red on figure 6.A). This part of the orbit also corresponds to large variation of the total water wapor content (figure 6.D) and the large negative differences (in blue on figure 6.A), corresponding to clear situations, correlated with the TWVC features, are probably caused by the NWP description in the forward radiative transfer computation and not by the cloud detection itself: for these pixels, the retrieved SST is correct (figure 6.C).

Resulting statistics, given in table 12.1, are of same order than over Europe for the clear land and the opaque cloudy situations. They are slightly higher for the clear sea data: std=0.75 for channel 4 compared to about 0.5 over Europe. If removing the data over the tropical Pacific (shown on figure 6), the resulting clear sea statistics are improved up to a bias= -0.01, std=0.62 for channel 4.

	Clear sea	***Clear land***	***Opaque cloudy***
N	179840	121901	417995
A4 (bias/std)	-0.12 / 0.75	-0.34 / 0.99	-0.13 / 0.44
A5 (bias/std)	-0.04 / 0.94	-0.35 / 1.31	-0.02 / 0.63

Table 12.1: cloud mask accuracy on a global revolution: Statistics of the departure between synthetic brightness temperatures and observations for the clear and the opaque cloudy situations

13. REFERENCES

[1] LeGleau H., M. Derrien, 2000: Prototype scientific description for Meteo-France/CMS. SAF/NWC/MFCMS/MTR/PSD, isuue 1

(you can get a copy at the web page : www.meteorologie.eu.fr/safnwc)

[2] Gutman G., D. Tarpley,A. Ignatov, S. Olson, The enhanced NOAA global dataset from the advanced very high resolution radiometer. Bulletin of the American Meteorological Society. 1995.

[3] Oort A.: Global Atmospheric Circulation Statistics. 1958 –1973.

[4] Klaes D, R. Schraidt 1999. The european ATOVS and AVHRR processing package (AAPP). ITSC10

[5] Chevallier, 2001 TIGR-like sampled databases of atmospheric profiles from the ECMWF 50 level forecast model. NWP-SAF report 1.

[6] Saunders R., M. Matricardi, P. Brunel 1999, An improved fast radiative transfer model for assimilation of satellite radiance observations. Quartt. J. Roy. Meteor. Soc, 125:1407-1426.

[7] Brisson, LeBorgne,Marsouin, 1998: Development of algorithms for SST retrieval at O&ISAF Low and Mid latitudes. Report to EUMETSAT

[8] Lavanant L 2000: Retrieval of cloud and surface parameter with AVHRR for IASI processing. Validation part. MASSIF2 report.

[9] Noveltis-CMS 2002 : Validation of the IASI test data simulation chain. NOV-3147-NT-1342

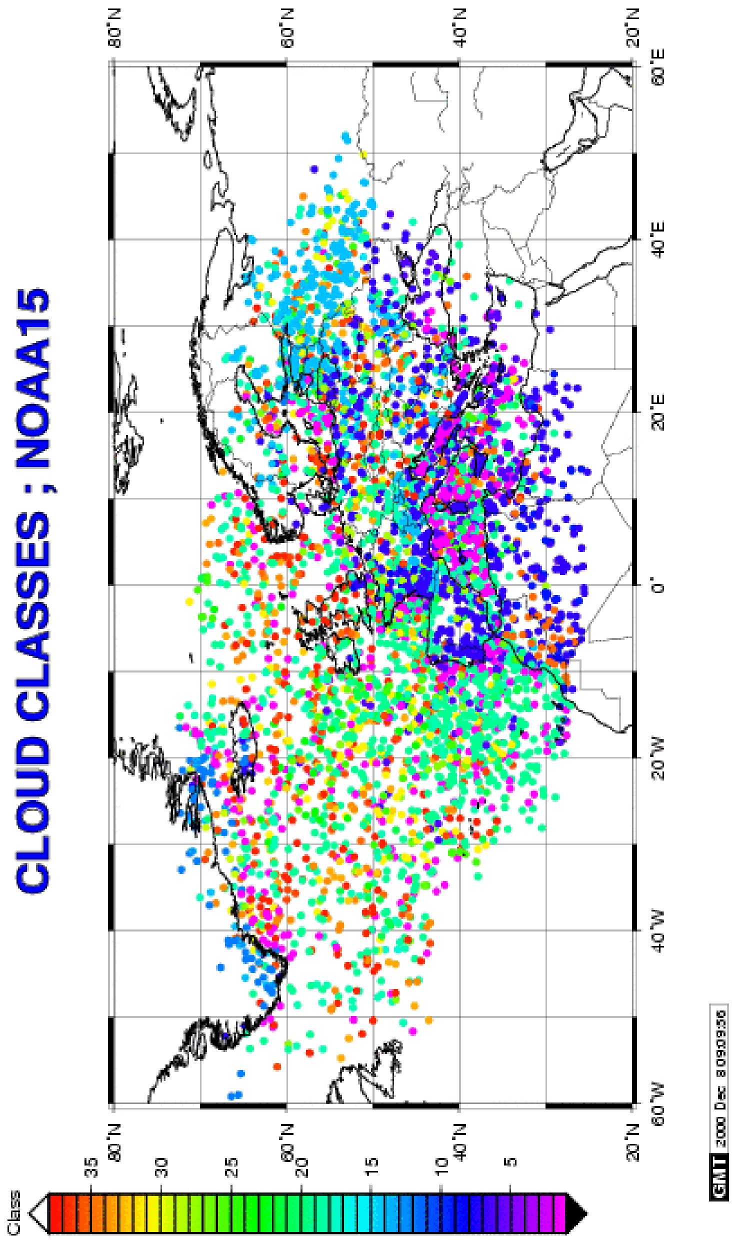
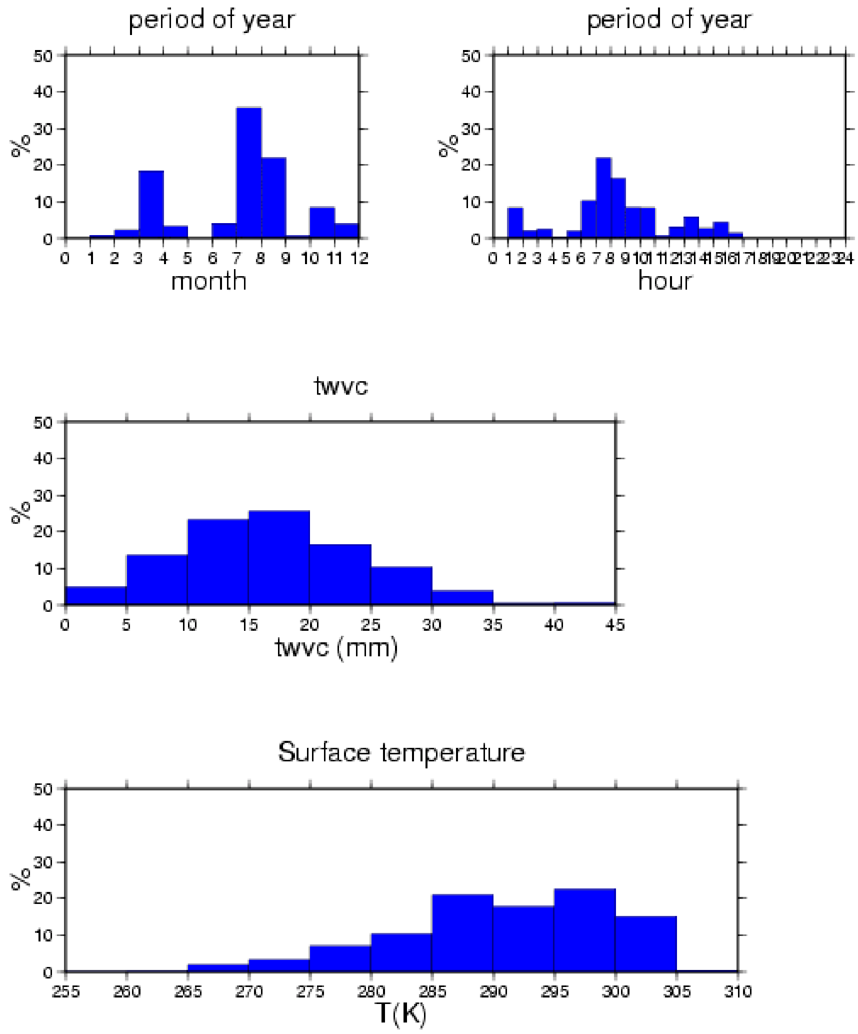


Figure 2: Spatial distribution of Noaa15 targets . Colors correspond to the different target types.



Figures 3: Distribution of the targets with the month and hour (upper figures), the total water vapor content (middle) and the surface temperature (below)

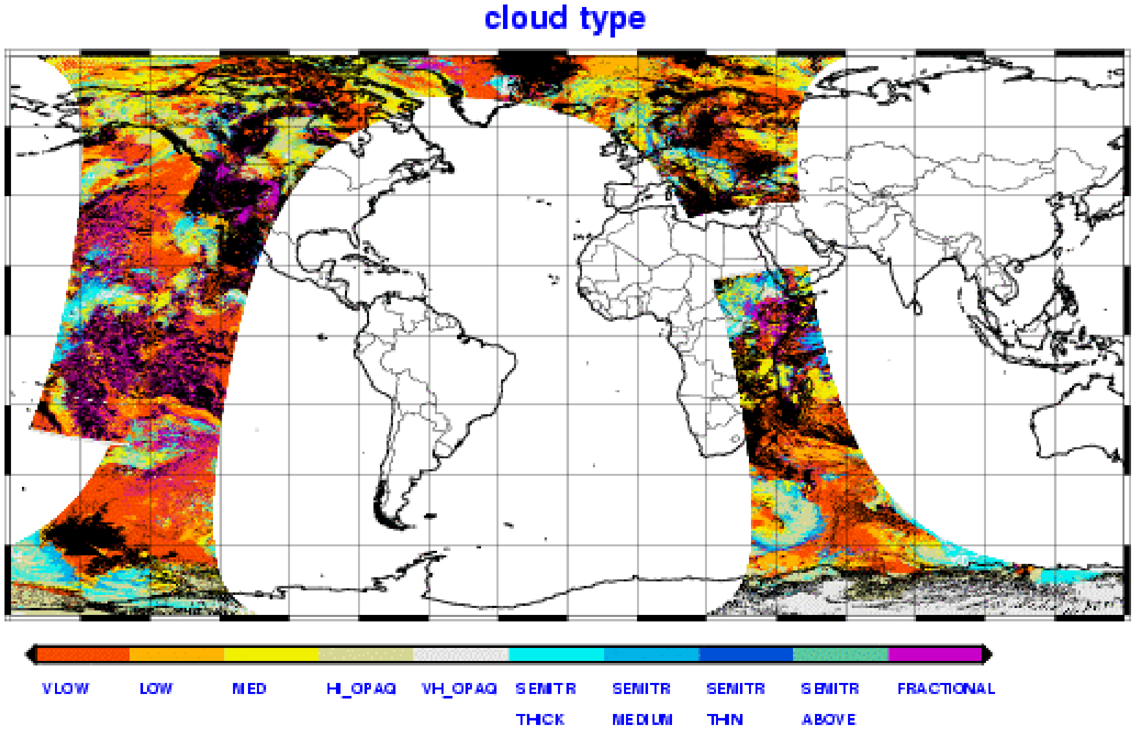


Figure 4: Noaa17 orbit. Cloud classification.

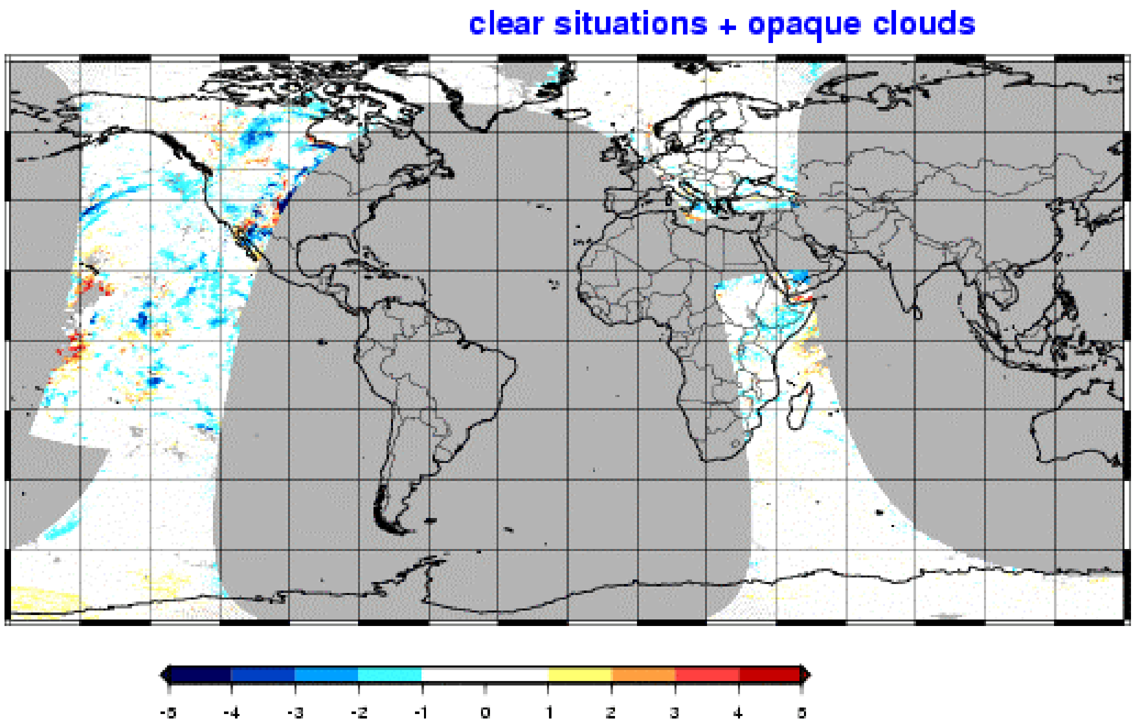
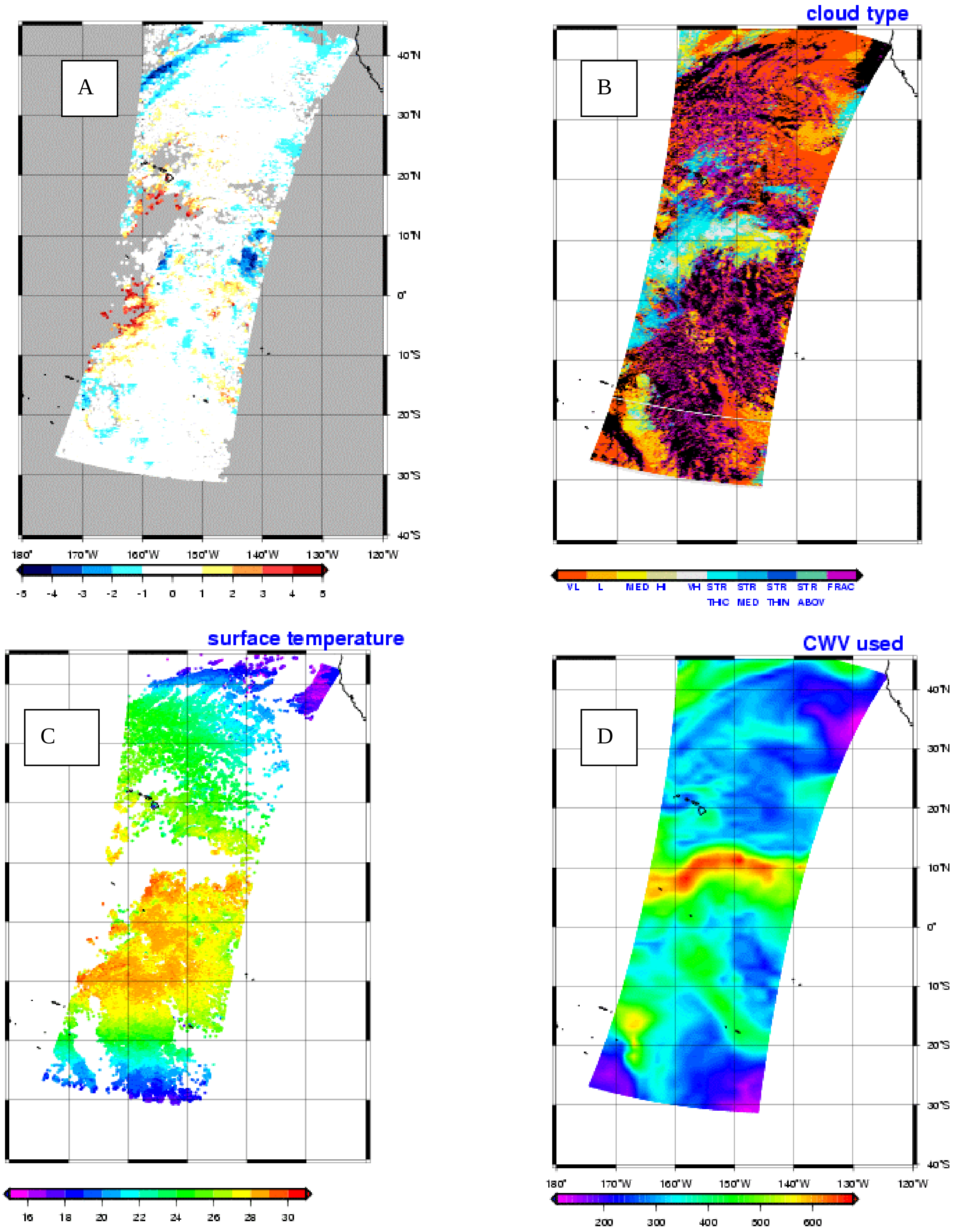


Figure 5: Noaa17 orbit. Departure between AVHRR channel 4 computed brightness temperatures and observations.



Figures 6: zoom on tropical Pacific ocean where the largest departures were found. The large positive differences (6.A) correspond to mis-detection of cirrus edges (6.B). The situation corresponds to large variation of the total water wapor content (6.D) and negative differences are correlated with the TWVC gradients and are probably caused by the NWP description.

Annexe 1 : FORECAST ASCII FORMAT

The files contain one header describing the grid for all the fields which are inside, and for all fields 3 sub-header lines and the values in 20i4 or 16i5 format

Header of 7 lines

-line 1 grid_type ch*12 type of data (analysis or forecast)
-line 2 grid_name ch*12 reference name of the grid
-line 3 grid_refdate i4,i2,i2,2x,i2 reference date and time of the fields: year, month, day, hour
-line 4 nb_hours_forecast i3 number of hours for the forecast
the date time of validity of the fields will be
grid_refdate + nb_hours_forecast
For analysis nb_hours_forecast is 0
-line 5 lat1, lon1 2f10.3 latitude and longitude of the first grid point
latitudes north are positive
longitudes east are positive
-line 6 step_lat, step_lon 2f10.3 latitude longitude increment between 2 grid nodes
step_lat should be negative (North to South)
step_lon should be positive (West to East)
-line 7 nbl, nbp 2i10 number of lines and pixels of the grid
lines are in the north-south direction
pixels are in the west-east direction

For each field:

-line 1 character*12 parameter name, one of the following:
T = temperature
HU = humidity
P = pressure
ALTITUDE = altitude over sea level
-line 2 character*12 level type one of the following:
ISOBARE
SURFACE
MER = sea level
HAUTEUR = altitude above surface
-line 3 integer i8 level value with respect to the level type
ex: 850 with level type ISOBARE means 850hPa
ex: 10 with level type HAUTEUR means 10m above surface
-line 4 to n integer values of the field in an array of (pixels,lines)
where pixels are on a parallel and lines on a meridian
latitude of array(i,j) =lat1 + (step_lat * (j-1))
longitude of array(i,j) =lon1 + (step_lon * (i-1))
format 20i4 unless format 16i5 for Z and P

storage units are: temperatures are K * 10 pressures are hPa*10
humidity in percentage * 10
land-sea in percentage altitude in meters