Cloud detection using SEVIRI IR channels for the GERB processing  

A. Ipe, I. Gonzalez Sotelino, E. Baudrez, N. Clerbaux, I. Decoster, S. Devitte, S. Nevens, A. Velazquez Blazquez  

Royal Meteorological Institute of Belgium, Department of Observations, Section Remote Sensing from Space, Avenue Circulaire 3 8, B-1180 Brussels, Belgium  
gerboma.be – http://gerb oma.be

1 Motivations

The Geostationary Earth Radiation Budget (GERB) processing currently relies on a scene identification applied to the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) for the estimation of the top-of-the-atmosphere (TOA) solar fluxes on a near real-time basis from the GERB broadband radiometers. More specifically, this scheme leverages a cloud mask based on the visible SEVIRI channels to properly select the angular dependency models (ADMs) for the radiance-to-flux conversion. It results that such product is unavailable during nighttime. Therefore, we have decided to develop a cloud detection algorithm solely based on thermal SEVIRI measurements. Its associated thermal cloud mask will be available under all observing conditions including sunglint and will allow to perform studies on clouds and aerosols radiative forcing over any time period.

2 Strategy

Major cloud detection algorithms based on the multispectral threshold technique in the thermal wavelengths rely on ancillary data. For example, the Nowcasting Satellite Application Facility (NWCSAF) [1] as well as the EUMETSAT Cloud Mask (CLM) [3] need numerical weather prediction (NWP) model fields to correctly discriminate clearsky from cloudy scenes. These fields include surface temperature, air temperatures at various altitudes as well as the total water vapor content of the atmosphere. While these data are not mandatory for the NWCSAF software, its accuracy is significantly decreased when there are missing since climatologies are then used. Such multispectral threshold approach for the GERB processing is discarded for several reasons: 1. GERB products would not anymore allow a fully independent validation of GCsMs, 2. the latter would rely on an external data source which could be unavailable, thus resulting in delayed processing (near-real-time constraint), and 3. the quality of these ancillary data could not be guaranteed over long time period due to NWP software updates, and therefore bias could be introduced in the thermal cloud mask (release of consistent product datasets).

The method suffers from a major drawback 3 Algorithm

Motivations

As mentioned above, we are considering a 60 days time-series of BTs for each pixel and infrared channel at a given time of day. These 60 values can usually be classified into 3 groups with decreasing BT:  

- clearsky
- thin or warm (low) clouds
- thick cold clouds

This classification is achieved by means of a modified k-means unsupervised clustering algorithm [2] as described below:

1. Let \( i = 0 \) and initialize the 3 clusters’ centers \( C_k^0 \) ( \( k = 1,2,3 \)).
2. Classify all 60 BTs according to their “nearest” cluster’s center \( C_k^i \).
3. Recompute the 3 clusters’ centers \( C_k^{i+1} \).
4. Let \( i = i + 1 \) and repeat from step 2 until \( C_k^i \) and \( C_k^{i+1} \) does not change more than 0.01 K.

The modification lies in the distance \( d \) used for the assignment of each sample BT \( T(\mu) \) ( \( \mu = 1, \ldots, 60 \)) to a specific class \( C_k \). We estimated by taking the monthly percentile at 95% of these 10 years of instantaneous \( A(L) \) values. The initialization of the scheme of \( C_k^0 \) then follows:

1. For \( C_k^0 \) (clearsky): \( C_k^0 = BT_{max} - \Delta(L) \) and \( \Delta(L) = \Delta/3.25 \).
2. For \( C_k^0 \) (thin or warm clouds): \( C_k^0 = BT_{max} - \Delta \) and \( \Delta = \Delta/3.25 \).
3. For \( C_k^0 \) (thick cold clouds): \( C_k^0 = BT_{min} \) and \( \Delta(L) = \Delta/3.25 \).

However, if \( C_k^0 \) and \( C_k^1 \) are not at least separated by \( L \), then the clustering is only performed on the 2 upper clusters. If the initialization fails with 2 clusters, then the time-series is assumed to be entirely clearsky.

4 Preliminary results

In the following figures, we have plotted the resulting cloud mask when our method is applied to the SEVIRI 8.7 um channel. One may note that our modified k-means clustering is successfully detecting low contrasted clouds in terms of brightness temperatures over the Atlantic Ocean. Moreover, convective cloud fields near the Equator are also correctly identified by our scheme. However, further quantitative comparisons with another cloud mask are also required.

Cloud mask based on the 8.7 um band (left) where black is for clearsky, grey is for low contrasted clouds, white is for high contrasted clouds and associated BTs [K] (right) for March 11 2007 at 0:00 GMT

5 Preliminary comparisons

As mentioned in section 2, two cloud masks are routinely derived from SEVIRI imagery. While the product from NWCSAF can be considered as one of the most accurate cloud detection scheme, the EUMETSAT CLM was primarily designed for robustness. Nevertheless, both are using NWP fields in their numerous threshold tests as well as visible and NIR channels during daytime. Thus, in the following, we are considering the NWCSAF cloud mask as the truth and comparing both the CLM and GERB products to it. Since the CLM is not delivering results at viewing zenith angles above 75°, we are restricting GERB comparisons to pixels below this limit. Depending on the surface type and its associated emissivity, it is obvious that the classification is improved for SEVIRI cloud channels exhibiting the highest contrast between clearsky and cloudy objects. This is illustrated in the following table.

### Cloud mask

<table>
<thead>
<tr>
<th>Cloud mask</th>
<th>Geotype</th>
<th>ocean</th>
<th>vegetation</th>
<th>desert</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWCSAF CMa</td>
<td>ocean vegetation desert</td>
<td>10.8 µm</td>
<td>73.80</td>
<td>91.19</td>
</tr>
<tr>
<td>GERB IR</td>
<td>ocean vegetation desert</td>
<td>86.12</td>
<td>89.12</td>
<td>94.88</td>
</tr>
<tr>
<td>GERB IR+</td>
<td>ocean vegetation desert</td>
<td>89.97</td>
<td>90.67</td>
<td>96.02</td>
</tr>
</tbody>
</table>

### Weighted means according to the number of night pixels of the hourly pixels' agreement (in percent) between the NWCSAF CMa and the GERB IR cloud masks with respect to the NWCSAF CMa cloud mask for August 11–17 2007 for the 3 IR channels and geotypes. The uncertainties are given at ±0.1 %

<table>
<thead>
<tr>
<th>Cloud mask</th>
<th>Geotype</th>
<th>ocean</th>
<th>vegetation</th>
<th>desert</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWCSAF CMa</td>
<td>ocean vegetation desert</td>
<td>10.8 µm</td>
<td>73.80</td>
<td>91.19</td>
</tr>
<tr>
<td>GERB IR</td>
<td>ocean vegetation desert</td>
<td>84.51</td>
<td>86.20</td>
<td>94.00</td>
</tr>
<tr>
<td>GERB IR+</td>
<td>ocean vegetation desert</td>
<td>89.01</td>
<td>83.33</td>
<td>84.77</td>
</tr>
</tbody>
</table>

### Averages of the daily weighted means according to the number of night pixels of the hourly pixels' agreement (in percent) between the MPEF CLM, GERB IR cloud masks with respect to the NWCSAF CMa cloud mask for August 11–17 2007 for the 3 IR channels and geotypes. The uncertainties are given at ±0.1 %

<table>
<thead>
<tr>
<th>Cloud mask</th>
<th>Geotype</th>
<th>ocean</th>
<th>vegetation</th>
<th>desert</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEF CLM</td>
<td>ocean vegetation desert</td>
<td>10.8 µm</td>
<td>73.80</td>
<td>91.19</td>
</tr>
<tr>
<td>GERB IR</td>
<td>ocean vegetation desert</td>
<td>84.51</td>
<td>86.20</td>
<td>94.00</td>
</tr>
<tr>
<td>GERB IR+</td>
<td>ocean vegetation desert</td>
<td>89.01</td>
<td>83.33</td>
<td>84.77</td>
</tr>
</tbody>
</table>

Since no spatial textural filtering is performed within the GERB IR cloud detection algorithm compared to the 2 other schemes, we also have calculated the agreement with an “augmented” GERB IR method by performing the NWCSAF CMa spatial filtering as post-processing (denoted in table by “GERB IR+”). As expected, the “augmented” GERB IR method always performs better than without the spatial filter. It can be noted that the MPEF CLM systematically exhibits a lower agreement than the standard GERB IR cloud detection over all geotypes for nighttime conditions. Nevertheless, low water clouds (stratuscumulus) are usually misidentified due to low BT contrast (<1 K). The use of the 3.7 µm channel through BT differences with the 10.8 and 12.0 µm channels (NWCSAF-like) should be investigated.

### References


### Acknowledgment

This study was supported by EUMETSAT and the Belgian Science Policy Office under the contract EUMET-GERB.