# Study on Impact of undetected clouds on Image Navigation Quality Assessment

Study Report, Final Issue

For the attention of: Mr Johannes MULLER - EUMETSAT

| | Function | Name | Signature | Date |
|---|---|---|---|---|
| Prepared by | Project manager/ Project Team | Florian POUSTOMIS<br>Océane LASSERRE | | 20/04/18 |
| Approved by | Business Unit Manager | Jean François VINUESA | | 20/04/18 |

153 rue du Lac – 31670 Labège – France
Tél. : +33 (0)562 88 11 11 – Fax : +33 (0)562 88 11 12 – E-mail : noveltis@noveltis.fr

# Distribution list

| INTERNAL | EXTERNAL | |
|---|---|---|
| **Name** | **Name** | **Company / Organisation** |
| Documentation NOVELTIS<br>Richard BRU<br>Cindy HORTALA<br>Jean François VINUESA<br>Océane LASSERRE | Johannes MULLER<br>Dieter JUST<br>Jagjeet NAIN<br>Sandrine FOSTER | EUMETSAT<br>EUMETSAT<br>EUMETSAT<br>EUMETSAT |

# Document status

| Study on Impact of undetected clouds on Image Navigation Quality Assessment<br>Study report, draft #2 | | | |
|---|---|---|---|
| **Issue** | **Revision** | **Date** | **Reason for the revision** |
| 1 | 0 | 02/02/2018 | Draft issue #1 |
| 2 | 0 | 16/03/2018 | Draft issue #2 |
| 3 | 0 | 20/04/2018 | Final version |

| **Modification status** | | | | |
|---|---|---|---|---|
| Issue | Rev | Status * | Modified section | Reason for the modification |
| 2 | 0 | M | §1 | Improvement of the draft issue 1 |
| 2 | 0 | M | §2 | Improvement of the draft issue 1 |
| 2 | 0 | M | §3 | Improvement of the draft issue 1 |
| 2 | 0 | I | §4 | Analysis of the impact of clouds on the chip acceptance process and the GQA quality |
| 2 | 0 | I | §5 | Conclusion of the study |
| 3 | 0 | M | §2.1 | Add precisions on the global methodology following the final meeting |
| 3 | 0 | M | §2.2.2 | Add precisions on the generation of the reference images (RID #1, 2, 3) |
| 3 | 0 | M | §2.3.1 | Add a reference concerning the Lanczos interpolation (RID #4) |
| 3 | 0 | M | §2.3.2 | Correct the scaling coefficient (RID #5) |
| 3 | 0 | M | §2.4 | Add precisions on the edge detection methodology (RID #7 and 8) |
| 3 | 0 | M | §2.5 | Add precisions on the matching method (RID #9, 10) |
| 3 | 0 | M | §0 | Add precisions and clarifications concerning the matching acceptance criterion and how it has been defined (RID #13, 14, 15, 16, 17, 18, 39) |
| 3 | 0 | M | §2.7 | Add clarifications on the GQA formulation and on the different coefficients in the formulation (RID #19, 20, 21) |
| 3 | 0 | M | §2.8 | Add precisions on the cloud cover threshold (RID #22 and 23) |

| | | | | |
|---|---|---|---|---|
| 3 | 0 | M | §3.1 | Add precisions on the GQA definition (RID #24) |
| 3 | 0 | M | §3.2.1, 3.2.2, 3.2.3 | Add precisions to clarify the methodology for the GQA quality definition (RID #25, 26, 27, 28) |
| 3 | 0 | M | §3.2.4 | Add precision on the formulation of the GQA quality (RID # 30) |
| 3 | 0 | M | §3.2.5 | Add precision on the validation of the GQA quality (RID #30 and 32) |
| 3 | 0 | M | §3.3 | Add precision on the first GQA results (RID #33 and 34) |
| 3 | 0 | M | §4.1.2 | Replace a figure (RID #35) |
| 3 | 0 | M | §4.1.3 | Add precision on the analysis of the cloud coverage (RID # 41) |
| 3 | 0 | D | §4.3.2.1 | Delete a table related to the cloud cover on the whole chips not used in the analysis (RID #42) |
| 3 | 0 | M | §4.3.2.1 | Replace the figures to be in accordance with the previous ones (cloud mask in white) (RID #44) |
| 3 | 0 | I | §4.3.2.3 | Add a section concerning the sensitive study regarding the cloud cover threshold (RID #22, 23, 40) |
| 3 | 0 | M | Appendix A | Move this section in appendix and add precision on the limits of the method (RID #11 and 12) |
| 3 | 0 | M | Appendix B | Add precision for the standard deviation of the offsets of the accepted chips (RID #29) |

* *I = Inserted* *D = deleted* *M = Modified*

# Acronyms

| | |
|---|---|
| GSHHG | Global Self-consistent, Hierarchical, High-resolution Geography Database |
| GQA | Geometric Quality Assessment |
| MSE | Mean Square Error |
| RMS | Root Mean Square error |
| STD | Standard Deviation |
| SVM | Support Vector Machine |

# Applicable and references documents

| | |
|---|---|
| [AD1] | Contract EUM/CO/4600001983/JML PO No 4500016033 - Study on Impact of undetected clouds on Image Navigation Quality Assessment |

| | |
|---|---|
| [RD1] | Wessel, P., and W. H. F. Smith, A Global Self-consistent, Hierarchical, High-resolution Shoreline Database, J. Geophys. Res., 101, #B4, pp. 8741-8743, 1996. |
| [RD2] | HRIT/LRIT Mission Specific Implementation<br>https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_TEN_05057_SPE_MSG_LRIT_HRI&RevisionSelectionMethod=LatestReleased&Rendition=Web |
| [RD3] | HRIT/LRIT Global Specification<br>*http://www.cgms-info.org/documents/cgms-lrit-hrit-global-specification-%28v2-8-of-30-oct-2013%29.pdf* |
| [RD4] | *Trigonometric interpolation of empirical and analytical functions*<br>C. Lanczos, Studies in Applied Mathematics, vol. 17, no. 1-4, pp. 123-199, 1938 |
| [RD5] | *A Computational Approach to Edge Detection*<br>JOHN CANNY, 1986<br>https://pdfs.semanticscholar.org/55e6/6333402df1a75664260501522800cf3d26b9.pdf |
| [RD6] | Sobel operator: *https://en.wikipedia.org/wiki/Sobel_operator* |
| [RD7] | OpenCV Canny Edge Detector<br>https://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/canny_detector/canny_detector.html |
| [RD8] | D. Huttenlocher, G. Klankerman, W. Reucklidge. *Comparing images using the Hausdorff distance*<br>https://people.eecs.berkeley.edu/~malik/cs294/Huttenlocher93.pdf |
| [RD9] | E. Baudrier, F. Marain-Nicolier, G. Millon, S. Ruan. *Une méthode de comparaison d'images quantifiant les dissimilarités locales – Application à la Classification d'impressions anciennes* |
| [RD10] | S. Narayanan, P.K. Thirivikrman. *Image Similarity using Fourier Transform.* |
| [RD11] | OpenCV Template Matching Function<br>https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html |
| [RD12] | *Geometric Quality Assessment of Orthorectified VHR Space Image Data*<br>Simon Kay, Peter Spruyt, and Kyriacos Alexandrou, 2003<br>https://pdfs.semanticscholar.org/7deb/a8e5e62e521c5a14ce527ee44478f187229c.pdf |
| [RD13] | *Geometric Quality Assessment of CARTOSAT-1 Data Products*<br>S. Muralikrishnan, A. Senthil Kumar, A.S. Manjunath and K.M.M. Rao<br>http://www.isprs.org/proceedings/XXXVI/part4/C-SAP-2.pdf |

# Table of contents

**1. INTRODUCTION ........ 8**

1.1. CONTEXT OF THE STUDY ........ 8

*1.1.1. Context of the study ........ 8*

*1.1.2. Document's content ........ 8*

**2. IMPLEMENTATION OF THE METHOD ........ 9**

2.1. OVERVIEW ........ 9

2.2. REFERENCE GENERATION ........ 10

*2.2.1. Reference Map ........ 10*

*2.2.2. Methodology ........ 11*

2.3. IMAGE CHIP ENHANCEMENT ........ 12

*2.3.1. Oversampling ........ 12*

*2.3.2. Conversion from 10 bits to 8 bits ........ 12*

*2.3.3. Contrast enhancement ........ 13*

2.4. EDGE EXTRACTION ........ 13

*2.4.1. Methodology ........ 13*

*2.4.2. Implementation ........ 14*

2.5. MATCHING ........ 15

*2.5.1. Methodology ........ 15*

*2.5.2. Implementation ........ 15*

2.6. MATCHING ACCEPTANCE ........ 17

*2.6.1. Analysis of the matching grid ........ 17*

*2.6.2. Matching criterion ........ 18*

2.7. GQA RESULTS ........ 19

2.8. CLOUD MASK ........ 19

**3. GQA QUALITY ........ 20**

3.1. DEFINITION ........ 20

3.2. MATHEMATICAL FORMULATION ........ 20

*3.2.1. Methodology ........ 20*

*3.2.2. Mean distance to the separation hyper plane ........ 22*

*3.2.3. Number of accepted chips ........ 22*

*3.2.4. Figure 11: Standard deviation of the GQA results according to the number of accepted chipsFormulation 23*

*3.2.5. Validation ........ 23*

3.3. FIRST RESULTS ........ 28

**4. IMPACT OF UNDETECTED CLOUDS ON GQA ........ 30**

4.1. ANALYSIS OF THE CHIP ACCEPTANCE BY THE TWO SCENARIOS ........ 30

*4.1.1. Presentation ........ 30*

*4.1.2. Detailed analysis of each classification cases ........ 30*

*4.1.3. Global synthesis ........ 34*

4.2. IMPROVEMENT OF THE SCENARIOS ........ 36

*4.2.1. Coastline mask ........ 36*

*4.2.2. Cloud threshold in the coastline mask ........ 37*

4.3. ANALYSIS OF THE RESULTS AFTER IMPROVEMENT OF THE SCENARIOS ........ 37

*4.3.1. Validation of the improvement of the scenarios ........ 37*

*4.3.2. Impact of clouds with the new implementation ........ 40*

**5. CONCLUSION ........ 46**

# 1. Introduction

## 1.1. Context of the study

### 1.1.1. Context of the study

Precise geolocation of Earth Observation satellite measurements in Earth coordinates is critical for users exploiting L1 and L2 satellite products. In theory, determining on-ground pixel location is possible from the information on the exact orbit position, attitude, and position of optical elements such as scanning mirrors. In practice, however, various errors, such as deviations from prescribed orbit, attitude variations, and uncertainty on the precise position of optical instrument components make this task prone to errors that are not acceptable for end-users. For this reason, Image Navigation and Registration (INR) models rely on the known positions of fixed landmarks to continually correct or minimize model errors. The position of landmarks (latitude and longitude), such as sea-costs or lake-costs are known with high accuracy and are used to calculate the difference between the true value and the values estimated by the model for pixels corresponding to these landmarks.

Currently, the quality of Image Navigation and Registration (INR) models and of the Geometric Quality Assessment (GQA) in the instrument data processing is highly depending on the quality of the landmark recognition and matching step. In the calculation process, the performance of the matching step is greatly influenced by cloud contamination in the images. The effects of this contamination reverberate on the whole process.

Currently, it is not possible to be sure to know the positions of each cloudy pixel of an image and to work only on cloud free pixels.

So in the scope of the study “Impact of undetected clouds on Image Navigation Quality Assessment”, [AD1], EUMETSAT wants to understand and to quantify the influence of this phenomenon in order to limit its effects. The present study is thus dedicated to:

- The development of a methodology (process, selection criteria, evaluation method…) in order to better measure the impact of clouds in the matching process;
- The statistical study of the characteristics of “bad” pixels/Image Chip and the definition of tools to detect them;
- The improvement, based on the study results, of the matching process, in order to better limit the impact of cloud contamination of the image; and, if possible, to get rid of the cloud masking step.

The study has been performed on SEVIRI images acquired in 2010 and provided by EUMETSAT.

### 1.1.2. Document’s content

This document presents the NOVELTIS study report on the project “Study on Impact of undetected clouds on Image Navigation Quality Assessment”.

The section 2 presents the proposed methodology to assess the GQA results of the SEVIRI images.

In section 3 the GQA quality is defined and its statistical relevance is demonstrated.

In section 4 an analysis is performed to assess the impact of undetected clouds on the GQA results and GQA quality.

Finally, the section 5 is the conclusion of the study.

# 2. Implementation of the method

This section is dedicated to the implementation of the two scenarios in order to assess the impact of undetected clouds on the GQA results. First and foremost, an overview of the global methodology is presented. Then each module of the scenarios is detailed.

## 2.1. Overview

In order to measure the impact of clouds on the GQA of SEVIRI images, the different image chips, extracted from the SEVIRI image, are compared to reference images. Two scenarios have been defined, which are presented in Figure 1.

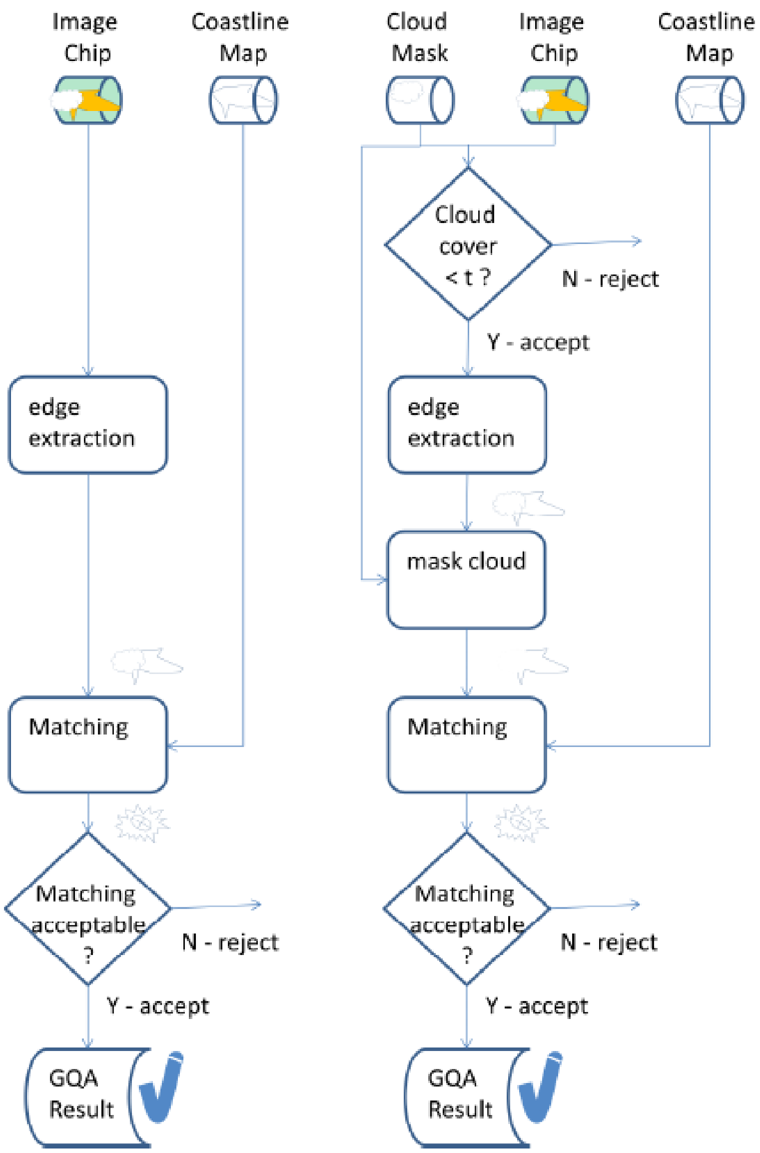



**Figure 1 - Method for GQA evaluation with and without cloud mask**

In the first scenario (called scenario 1 in the document), cloud information is omitted whereas, in the second one (called scenario 2 in the document), cloud information is used. In both cases, edges are extracted from image chips and matched with corresponding reference images representing specific landmarks on the Earth surface. If the matching is acceptable the image chips are kept in order to compute the GQA of the SEVIRI image. In the second scenario, cloud information is used before matching: image chips with important cloud cover are rejected before the matching and for other image chips, a cloud mask is applied to the edges of the image chips before matching in order to remove artefacts due to clouds. **A comparison of the results of these two scenarios will highlight the impact of the clouds in the GQA computation process**.

Actually the GQA is estimated for each SEVIRI image band independently. The number of the studied SEVIRI bands and the relation with the corresponding band ID is provided in Table 1. Then the image chips extracted from the SEVIRI images are processed band by band.

**Table 1: Lookup table for the correspondence of the studied SEVIRI band numbers**

| Band number | 0 | 1 | 2 | 3 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Band ID | VIS 0.6 | VIS 0.8 | IR 1.6 | IR 3.9 | IR 8.7 | IR 9.7 | IR 10.8 | IR 12.0 | IR 13.4 |
| Band type | VNIR<br>Core Imager | VNIR<br>Core Imager | VNIR<br>Core Imager | IR / Window<br>Core Imager | IR / Window<br>Core Imager | IR / Ozone<br>Pseudo-Sounding | IR / Window<br>Core Imager | IR / Window<br>Core Imager | IR / Carbon Dioxide<br>Pseudo-Sounding |

The first task of the study is to implement all the steps of the two scenarios. Common steps are implemented in the same way in both scenarios. The implementation details are documented in this section.

## 2.2. Reference generation

First and foremost, the reference images corresponding to the 200 image chips of the database have to be generated.

### 2.2.1. Reference Map

The reference coastline map chosen to perform the matching is **GSHHG - A Global Self-consistent, Hierarchical, High-resolution Geography Database**. This dataset has already been used by NOVELTIS in previous studies and is fully adapted to the needs of the study. This section presents the dataset.

The GSHHG is a high-resolution geography data set, amalgamated from two databases in the public domain: World Vector Shorelines (WVS) and CIA World Data Bank II (WDBII). The former is the basis for shorelines while the latter is that for lakes, although there are instances where differences in coastline representations necessitated the adding of WDBII islands to GSHHG. The WDBII source also provides all political borders and rivers. GSHHG data have undergone extensive processing and should be free of internal inconsistencies such as erratic points and crossing segments. The shorelines are constructed entirely from hierarchically arranged closed polygons.

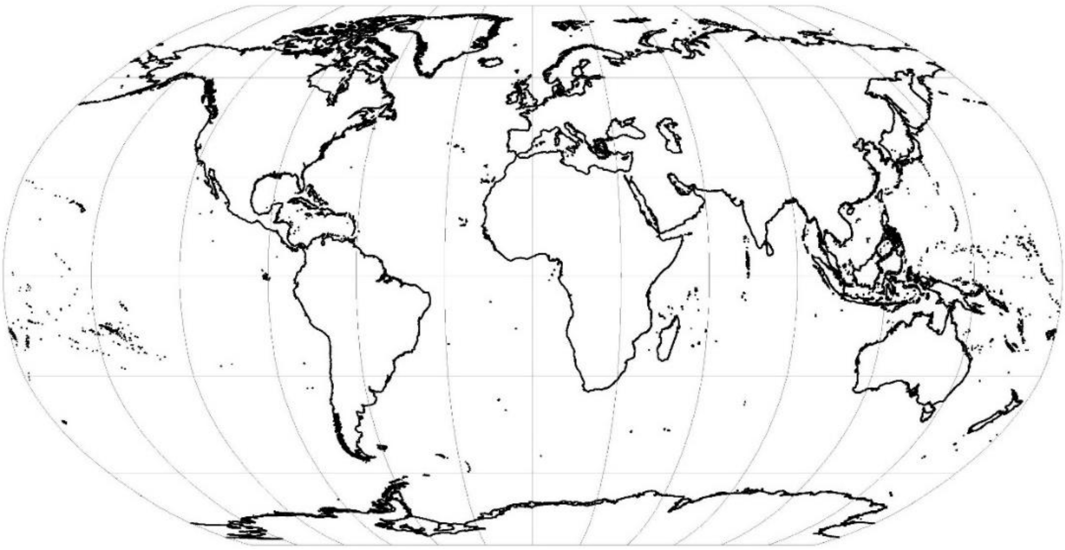


**Figure 2: Coastlines and islands represented by vectorial data**

The GSHHG geography data come in five different resolutions: crude(c), low(l), intermediate(i), high(h), and full(f). Shorelines are further organized into 4 hierarchical levels: boundary between land and ocean (L1), boundary between lake and land (L2), boundary between island-in-lake and lake (L3), and boundary between pond-in-island and island (L4).

GSHHG is distributed in shapefile format including the complete GSHHS polygons and WDBII lineaments in the five resolutions under the GNU Lesser General Public license. The GSHHG data version 2.3.5 (April 12, 2016) used for the study is available for direct download from: https://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html. GSHHG data production methodology is described in [RD1]. Under this study the full resolution dataset has been used to get the best precision (L1f).

### 2.2.2. Methodology

The generation of the reference images is made by using the GSHHG dataset but also the landmarkChipFile provided by EUMETSAT and containing the centre position and the size of each landmark reference image, and the formulas provided by EUMETSAT too and allowing to convert (row, column) coordinates into the SEVIRI grid to (latitude, longitude) according to the GEOS projection, [RD2], [RD3].

First and foremost, for each reference image, the centre and size are read in the landmarkChipFile. The counting of centre coordinates read in the file is (0,0) based whereas for the conversion formulas from (row, column) to (latitude, longitude), the counting of the SEVIRI grid cells is (1,1) based. So 1 is added to the row centre and column centre coordinates. However, as there is a mistake in the indexing of the image chips provided by EUMETSAT for this study, 1 should be subtracted to the row centre and column centre coordinates of the reference images in order to fit with the corresponding image chips. Thereby, if the reference images are relevant and correct for this study, regarding the provided image chips, they are not for other studies if the image chips are extracted without offset.

Then the size of each reference image is multiplied by the oversampling factor (4 in this study).

Finally, knowing the centre coordinates in the SEVIRI grid, the number of row and column of the reference image and the required resolution, the (row, column) coordinates of each pixel of the reference image in the SEVIRI grid are converted into (latitude, longitude) thanks to Navigation Functions described in [RD3], p. 24. The location of this (latitude, longitude) point is checked in the GSHHG coastline map dataset and the nature of the point (water or ground) is reported to the corresponding (row, column) pixel of the reference image (0 or 1). The obtained reference images are binary images (0 for water and 1 for ground) as is shown in Figure 3. Later, the edges of these reference images will be extracted (see section 2.4) to perform the matching with the edges of the different image chips.

**Figure 3: Corsica (landmark 56) and Cyprus (landmark 60) reference images**

## 2.3. Image chip enhancement

Preliminary steps are applied to the extracted image chips from the SEVIRI images in order to facilitate the edge extraction and the matching to the reference images.

### 2.3.1. Oversampling

The first preliminary step is an **oversampling of factor 4**. Indeed, the offsets we have to find are often smaller than one low resolution SEVIRI pixel (*i.e.* 3km) so the image chips have to be oversampled. Figure 4 presents the results of the Lanczos oversampling, described in [RD4], and we can note that the coasts appear more precisely after the oversampling.

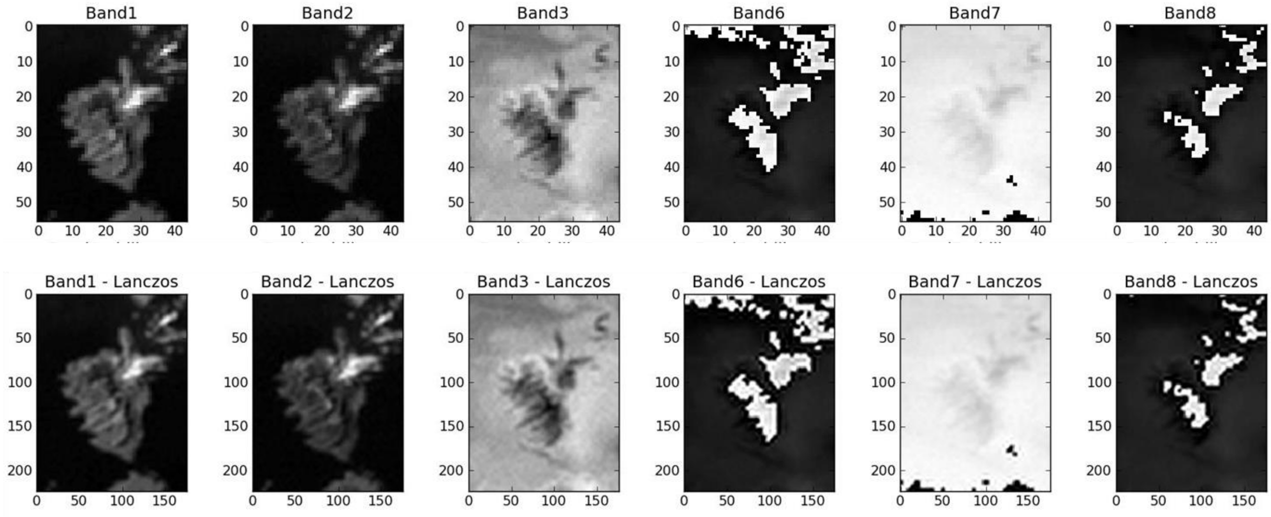



**Figure 4: Several bands of MSG2-SEVI-MSG15-0100-NA-20100817171241 chip 56 (Corsica) without oversampling (first row) and after oversampling by a factor 4 (second row)**

### 2.3.2. Conversion from 10 bits to 8 bits

The second preliminary step consists in converting the input image chips from 10 bits to 8 bits (between 0 and 255) because the edge detector used in the following works only with 8 bits images.

This conversion could have a small impact in the study result because two pixels with different close values could have the same value after conversion. However, as the used edge extraction method works only with 8 bits images, this impact cannot be quantified.

### 2.3.3. Contrast enhancement

The third preliminary step consists in **contrast enhancement**. Indeed, because of bright clouds or specific illumination conditions, the bands of the image chips can be saturated. Then thresholding 2% of the pixel with the lowest intensity and 2% of the pixel with the highest intensity and rescaling the image band chips between 0 and 255 allows to enhance the band contrast of the image chips and to detect more easily the edges on the images as is shown in Figure 5.

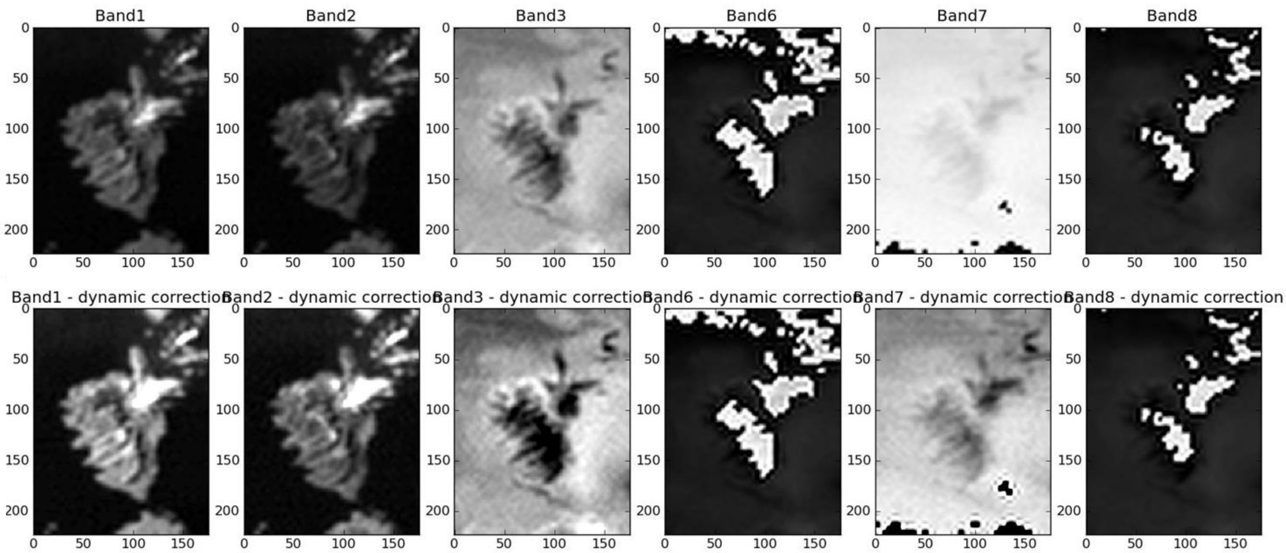



**Figure 5: Several bands of MSG2-SEVI-MSG15-0100-NA-20100817171241 chip 56 (Corsica) without contrast enhancement (first row) and after contrast enhancement (second row)**

## 2.4. Edge extraction

### 2.4.1. Methodology

The edge extraction of the images has been carried out with a **Canny Edge Detector** [RD5] implemented in the openCV library. The edges are extracted exactly the same way for image chips and binary reference images.

The canny edge detection process is divided into the following steps:

1. First and foremost, the gradient of the image is computed. A low gradient means that the area is uniform. A high gradient means that there is a great intensity gap between two areas, it is an edge. The direction of the gradient shows how the edge is oriented. A Sobel filter is used to find the intensity of the gradient in the image, [RD6].

2. The following step of the method enables to find the evident edges, where the gradient is higher than a defined high threshold.

3. The last step is used to improve the edges and ensure the continuity. All the gradient pixels between the high threshold and a second lower threshold are evaluated. If a pixel, the gradient of which is between these two thresholds, is linked to a pixel flagged as an edge previously, then it is flagged as an edge too. If a pixel, the gradient of which is between these two thresholds, is not linked to a pixel flagged as an edge previously, then it is not considered as an edge.

The methodology is illustrated by the one dimensional example in Figure 6. The gradient of an image chip is computed. All the pixels “A”, the gradient of which is higher than the defined “maxVal” threshold, are flagged as edge pixels. Then all the pixels the gradient of which is between the defined “minVal” and “maxVal” thresholds, are considered. The pixel sequence “C”, linked with pixels defined as edges previously, is considered as edges too. However, the pixel sequence “B”, not linked to pixels flagged as edges previously is not defined as edges. At the end of the process, a binary edge image is obtained.

After the edge extraction an edge correction step is applied to delete all the edges found far from the reference image edges, due to clouds or artefacts on the image chip bands. A margin equal to the maximum search distance defined for the matching (see section 2.5) is used.

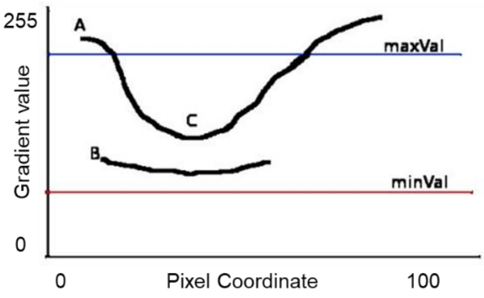



**Figure 6: Canny edge detector principle**

### 2.4.2. Implementation

In this study, the OpenCV's Canny edge detector [RD7] has been used.

The thresholds have been defined empirically and depend on the standard deviation of each enhanced image chip band. Indeed it is not efficient to set constant thresholds for all the bands and all the SEVIRI images acquired with different illumination conditions as the difference of contrast between the water and the ground can change between these different cases. Then adaptive thresholds have been implemented according to the image chip bands the edges of which have to be extracted. **The low threshold is equal to the standard deviation of the image chip band multiplied by 1.5 whereas the high threshold is equal to the standard deviation of the image chip band multiplied by 2**.

The same process is applied to the binary reference images in order to detect the edges of the landmarks. Of course, with binary images, the third step of the Canny methodology is useless.

Results of the edge extraction are shown in Figure 7. The reference image for this chip is shown in Figure 3 (left). Comparing these edges to the reference image justifies the benefit of the oversampling (differences between the edges of the first and second rows for the band 1 for example), the contrast enhancement (differences between the edges of the second and third rows for the band 2 for example) and edge correction (differences between the edges of the second and third rows).

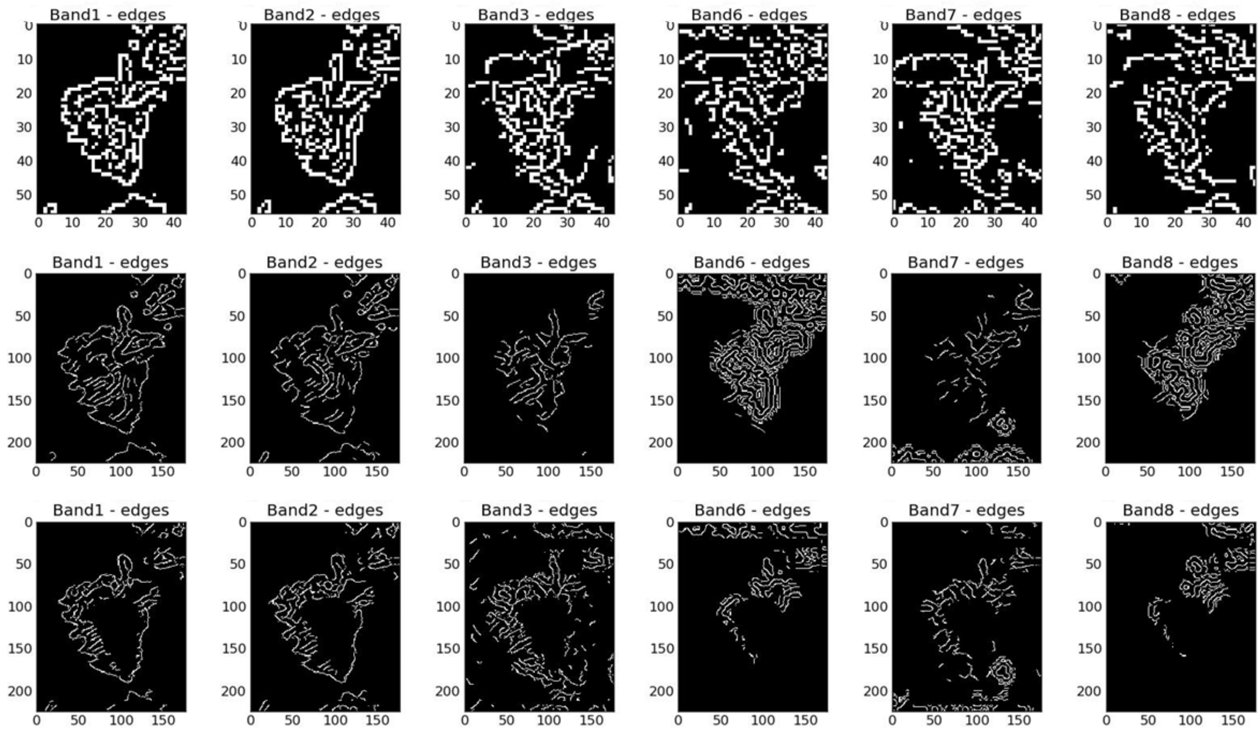


**Figure 7: Edge extraction for several bands of MSG2-SEVI-MSG15-0100-NA-20100817171241 chip 56 (Corsica), without image chip oversampling, image chip enhancement, and edge correction (first row), with image chip oversampling and without image chip enhancement, and edge correction (second row), and with image chip oversampling, contrast enhancement and edge correction (third row)**

## 2.5. Matching

As the different SEVIRI bands are processed independently, a matching is performed between the extracted edges of a given band of an image chip, and the extracted edges of the corresponding reference image. For these two images, the same edge extraction methodology is used, see section 2.4 and provides binary output images.

Several matching methods have been tested in this study, such as that of Hausdorff distance [RD8] and [RD9] and the MSE between Fourier transforms [RD10] but the selected one is based on cross correlation and described below.

### 2.5.1. Methodology

The used metric to compute the matching score between the two binary images – the extracted edges of an image chip band and the extracted edges of the corresponding reference image – is the **cross correlation**.

For a given reference image, the best match is searched with a moving matching window of same size as the reference image into a search window centred in the image chip band.

The matching result is a matching grid giving for each pixel corresponding to the centre of the moving matching window (*i.e.* for each offset between the matching windows) the matching score (*i.e.* the cross correlation) between the two matching windows. So the matching grid is square and its size is equal to twice the maximum search distance in pixel plus one pixel. **The most relevant offset between the image chip band and the reference image is indicated by the location in the matching grid of the pixel with the highest matching value**.

### 2.5.2. Implementation

The openCV function *matchTemplate* [RD11] is used to compute the cross correlation between two matching patches for each possible position of the mobile matching window into the search window.

The maximum search distance has been defined as 3 low resolution SEVIRI pixels (*i.e.* 3km pixels) in each direction, or 12 pixels in each direction after oversampling so the high of the search window is equal to the high of the reference image plus 24 oversampled pixels, and the wide of the search window is equal to the wide of the reference image plus 24 oversampled pixels. Thus, the size of the resulting matching grid is (25, 25).

The output correlation values of the *matchTemplate* function in the matching grid are between 0 (bad match) and 1 (perfect match) and are multiplied by 100 to be presented as a percentage.

To illustrate, the edges of the reference image for the landmark 56 (Corsica) have been shifted of one low resolution SEVIRI pixel to the East and two low resolution SEVIRI pixels to the South as is shown in transparency in left Figure 8. Then the matching algorithm has been applied to the edges of the reference image and the edges shifted of the reference image and the matching grid on the right is obtained. Regarding the maximum correlation value the initial offset can be retrieved.

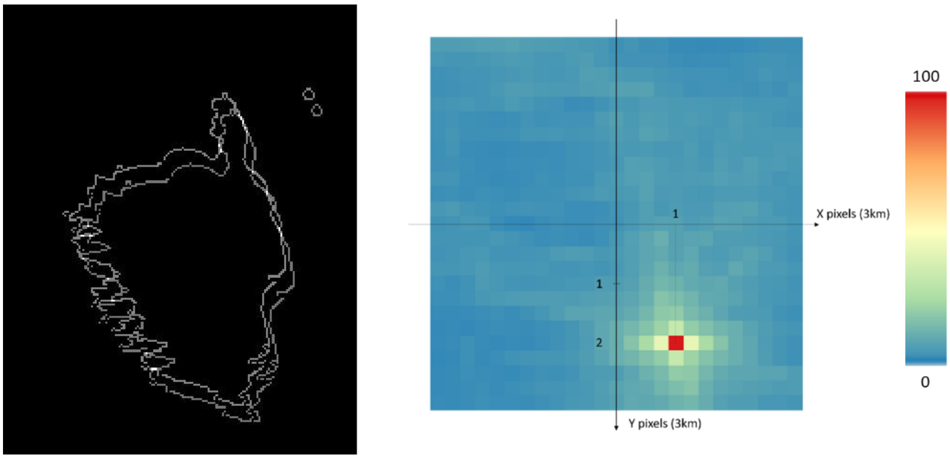



**Figure 8: Matching process, moving matching patches superimposed by transparency (left) and resulting matching grid (right)**

An example of matching result of a band of a SEVIRI image chip with its corresponding reference image is show in Figure 9.

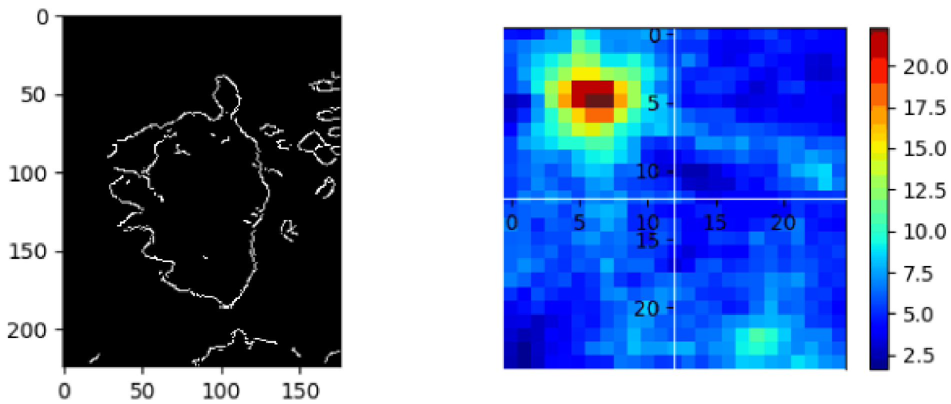



**Figure 9: Edges of the MSG2-SEVI-MSG15-0100-NA-20100521211241 chip 56 band 3 (left) and result of the matching with the edges of the corresponding reference image (landmark 56, Corsica)**

Of course, the shape of the correlation peak in a matching grid can be impacted by the shape of the corresponding landmark, and the auto-correlation of the edges of the reference image. However, the landmarks have often edges in several directions, as the Corsica landmark illustrated above, so the auto-correlation surface of the corresponding reference images is small, and the impact on the shape of the correlation peak resulting of the matching an between image chips band is small too.

By purpose of concision, it could be written in the following of the document “the matching between the image chip band and the reference image”, however, the matching is always performed between the binary extracted edges of these two images.

An alternative method have been investigated but abandoned. It is presented in Appendix A.

## 2.6. Matching acceptance

Once the matching grid is generated for a given image chip band, a **matching criterion is applied in order to reject or accept the image chip band to compute the GQA of the SEVIRI image band** (see section 2.7).

### 2.6.1. Analysis of the matching grid

As is mentioned above, the value of a pixel in the matching grid is the matching score between a reference image and an image chip band if the image chip band is shifted at the location of the studied pixel in the matching grid. So the location in the matching grid of the pixel with the highest matching score gives the most relevant offset of the image chip band regarding the reference image.

Sometimes the matching is really good and the resulting offset between the image chip band and the reference image is well defined. So for a given SEVIRI image band, these image chips can be used to compute the GQA of the SEVIRI image band. However, sometimes the matching is not good because of the presence of clouds or artefacts, or because of bad illumination conditions, and the offset between the image chip band and the reference image is not well defined. So for a given SEVIRI image band, these image chips should not be used to compute the GQA of the SEVIRI image band. Thus, in order to decide whether the matching between a reference image and an image chip band is good enough to compute the GQA of a SEVIRI image band, a matching criterion should be defined based on the matching grids.

Several coefficients have been extracted to **characterize the relevance of the matching result** and to accept or not the image chip for further steps:

- **Maximum correlation value:** higher the maximum correlation value, higher the number of pixels of the edges located at the good place regarding the reference image. The maximum correlation value is defined as follow:

  $$Max_{corr} = \max_{i,j} corr(i,j)$$

  With:

  - $i$: *line index of the matching grid*
  - $j$: *column index of the matching grid*
  - $corr(i,j)$: *cross-correlation value in the matching grid at the location* $(i,j)$*;*

- **Median of the matching grid:** lower the median, more concentrate the correlation peak and more precise is the resulting offset. The median of the matching grid is the value of one pixel of the matching grid such as half of the pixels of the matching grid has a value higher than the defined median and the other half has a value smaller than the defined median.

  The median of a matching grid depends on the shape of a landmark because it is linked to the auto-correlation of the edges of a reference image. Indeed, as mentioned above, if a landmark has edges in all the directions, the auto-correlation surface of the edges of the corresponding reference image is small. So the median of the matching result, between an image chip band and this reference image, has more chance to be small than the median of a matching result between an image chip and a reference image with a bigger auto-correlation surface (representing a landmark with edges in a unique direction for example).

  However, as the median is used with other coefficients to characterize a matching grid, the impact of the shape of the landmark is reduced in the matching criterion.

- **Weighted average of the distances between the global correlation peak and other high local correlation peaks:** this coefficient enables to know if local correlation peaks higher than 33% of the global correlation peak are located near the global peak (that highlights some imprecisions in the edge extraction step but that is not determinant for geometric quality of the image chip band) or far from the global peak (that highlights some artefacts that can reflect a bad match). Thus for the local correlation peaks higher than 33% of the global correlation peak, the average distance to the global correlation pixel weighted by the correlation value of the local peaks is computed as follow:

  $$\overline{Dist} = \frac{\sum_P c_P \left\| i_{P_{max}} - i_P \right\|}{n}$$

  *With:*

  - $P_{max}$: *global correlation peak;*
  - $P$: *local correlation peak above the threshold;*
  - $i_P$: *location of the pixel* $P$*;*
  - $\|i_{P_{max}} - i_P\|$: *the Euclidian distance in pixels between the pixel locations* $i_{P_{max}}$ *and* $i_P$*;*
  - $c_P$: *correlation of the pixel* $P$*;*
  - $n$: *number of local correlation peak above the threshold;*

- **Band number**: as the range of the previous coefficients is different according to the SEVIRI image band, the band number should be taken into consideration when computing the matching criterion in order to accept image chip bands giving a good matching with the corresponding reference images for all the SEVIRI bands. The band numbers used for this study and the correspondence with the SEVRIR band ID is provided in Table 1.

### 2.6.2. Matching criterion

In order to decide if an image chip band should be accepted or not for the GQA computation of a SEVIRI image band, regarding the matching with a reference image, thresholds have to be implemented based on the matching grid coefficients presented above. A **multi-dimensional SVM approach** seems to be a good idea to choose a matching acceptance criterion according to the matching results since 4 coefficients characterizing these results have been defined.

However, a ground truth is necessary to train the SVM model. As this ground truth does not exist it has been built during the study by visually analysing the matching of each band (excepted for the water vapour bands and the HRV band) of 10 image chips selected randomly of 72 SEVIRI images distributed over the 12 months of the year 2010 and on the 24 hours of a day in order to process images acquired at all seasons and at all day/night times. So for the 6,480 studied image chip bands, the acceptance or rejection has been decided regarding the edge extraction and the resulting matching grid by using Scenario 2 (with cloud information). Among the 6,480 image chip bands, 3,265 have been rejected by Scenario 2 because of a high amount of clouds (see section 2.8) and the remaining 3,215 have been visually analysed. This step was time consuming (around one day of work) but is very important for the success of the study. As the choice to accept or reject a matching result is subjective, the building of the ground truth has been performed by two people, each one analysing the half of the total number of matches, thus the subjectivity of the ground truth generation has been reduced.

So using the generated ground truth, a SVM model has been trained by using 80% of the analysed matches, which allows to elaborate a binary classifier. The SVM model is a combination of support vectors of dimension 4 (corresponding to the 4 matching grid coefficients presented above) for each class, the acceptance class and the rejection class. The acceptance class is composed of 190 support vectors and the rejection class is composed of 189 support vectors. Then the SVM classifier has been tested by using the remaining 20%. The result of the classification test is:

- **95.03% of good classifications**;
- **2.64% of omissions** (image chip band rejected whereas it should be accepted according to the visual study);
- **2.33% of false alarms** (image chip band accepted whereas it should be rejected according to the visual study).

The classification results are very good and the SVM model has been saved.

In the 4 dimension SVM space each match is represented by a point of coordinates (correlation maximum, median, distance between the local correlation peaks and the global correlation peak, band number). Then a separation hyper plane segregates the points representing the accepted matches and the points representing the rejected matches. For each point, **the distance to the hyper plane reflects the reliability of the classification**. The greater the distance, the more reliable the matching acceptance/rejection. For the accepted matches, the distance is positive, whereas for the rejected matches, the distance is negative. This distance is normalized for each studied SEVIRI band and provided in percent, by dividing the Euclidian distance of the point representing a match to the separation hyper plane, by the Euclidian distance to the separation hyper plane of the farthest point to this hyper plane. For a given SEVIRI band number, the three other coordinates in the decision space (maximum, median, and weighted average distance of the local maxima to the global maximum) of the farthest point to the separation hyper plane are computed according a synthetic perfect matching grid (one pixel at the maximum of correlation (100%) and all the others at 0%).

The best match is represented by a matching grid with one pixel at the maximum correlation value (100%), and all the other pixels at 0%. As the band number has an impact on the distance of the point representing a match to the separation hyper plane

Other features, such as the temporal (image to image) behaviour could be studied in order to define another matching criterion.

## 2.7. GQA results

The GQA value is the output of the process of geometric quality assessment of an image. In the scope of this study, it describes the agreement of the location of a SEVIRI image band according to the known location of several landmarks on the Earth respecting the GEOS projection. Normally, the GQA is expressed as a vector in km (South /North, East/West). As it is a vector field, the GQA can vary with the position on the image. In order to simplify the GQA metric it may be reduced to a mean vector or a scalar by computing the RMS over the GQA across the image, [RD12] and [RD13]. Then **the GQA characterizes the distance between the expected location of the image and its real position**.

The GQA is not directly accessible but can be estimated as it is performed in this study by analysing the matching of each image chip band with corresponding reference sub-image showing a significant feature: a landmark containing a feature. Regarding the matching results, only the image chip bands giving a good matching result are used in order to estimate the GQA of a SEVIRI image band.

Thus, in this study, for a given SEVIRI image band, **the GQA is defined as the RMS of the offsets in South/North and East/West directions highlighted by each accepted match between the same band of the image chips and the corresponding reference images**. It is computed as:

$$GQA = \sqrt{\frac{\sum_{chips}({delta_X}^2 + {delta_Y}^2)}{nb_{chips}}}$$

Where $delta_X$ and $delta_Y$ are the column and line offsets obtained according to the matching of a given image chip band, and $nb_{chips}$ is the number of image chips, which the studied SEVIRI band number has been accepted by the different acceptance criteria of a scenario, and used to compute the GQA of the studied SEVIRI image band. The GQA is provided in low resolution SEVIRI pixels.

It is recalled here, that as the GQA is estimated for each SEVIRI band independently, the number of chips mentioned here, and in the following of the document, refers to the number of accepted chips which the studied SEVIRI band number has been accepted by the matching acceptance criterion (and the cloud acceptance criterion).

## 2.8. Cloud mask

In scenario 2 of the study, cloud information is used at two levels:

- First, the percentage of clouds in the image is computed. If this percentage is above a defined threshold (50%), the image chip is rejected before the matching step;
- If the percentage of cloudy pixels is under the defined threshold (50%), the matching of each image chip band is performed. But, before the matching step, a cloud mask is applied to remove edges caused by the presence of clouds.

The cloud cover threshold is arbitrary. However a sensitive study has been performed, see section 4.3.2.3.

# 3. GQA quality

This section is dedicated to the definition of the GQA quality. First and foremost, the GQA quality is clearly explained. Then a mathematical formulation is provided and the statistical relevance of this formula is demonstrated. Finally, first results of GQA results from the two scenarios are analysed thanks to the GQA quality.

## 3.1. Definition

**The GQA quality coefficient is a mean to assess the reliability and the relevance of the estimated GQA result**. It could be seen as an error bar to the GQA estimate, but in the scope of this study, for a given SEVIRI image band, it is sufficient to determine which of the two estimated GQA (the one, using scenario 1, and the other, using scenario 2) is closer to the truth.

In order to assess the quality of the estimated GQA for a SEVIRI band, **a two scale analysis on the accepted image chips individually and globally** can be performed:

- Analysis of the decision to accept an image chip to compute the GQA: are the accepted image chips really good considering the studied SEVIRI band? This can be assessed by computing the average distance to the separation hyper plane of all the accepted chips for the studied SEVIRI band;
- Analysis of the global value of the GQA: is the GQA result representative of the location precision in the whole SEVIRI image band? Several metrics can help answer this question:
  - Number of accepted image chips for a considered SEVIRI band in order to estimate the GQA of this SEVIRI band;
  - Geographical distribution of the accepted image chips for a considered SEVIRI image band;
  - Standard deviation of the offsets in South/North and East/West directions highlighted by the matching (and presented in section 2.7) of the accepted image chips for a considered band, defined as follows:

$$std_{offX} = \sqrt{\frac{\sum_{chips}\left(delta_X - mean_{delta_X}\right)^2}{nb_{chips}}}$$

$$std_{offY} = \sqrt{\frac{\sum_{chips}\left(delta_Y - mean_{delta_Y}\right)^2}{nb_{chips}}}$$

  Where $delta_X$ and $delta_Y$ are the column and line offsets obtained according to the matching of a given image chip band, $nb_{chips}$ is the number of image chips, which the studied SEVIRI band number has been accepted by the different acceptance criteria of a scenario and used to compute the GQA of the studied SEVIRI image band, and $mean_{delta_X}$ and $mean_{delta_Y}$ are the mean offset in East/West and South/North directions respectively regarding the accepted image chips. The standard deviation of the offsets is provided in low resolution SEVIRI pixels.

## 3.2. Mathematical formulation

### 3.2.1. Methodology

In order to define a mathematical formulation of the GQA quality, each coefficient identified above as having a possible impact on the GQA quality have been studied independently. Indeed, as there is no existing tool or formulation to evaluate the quality of the estimated GQA, the objective is to assess how the GQA is impacted by the variation of these different coefficients in order to find a formulation for the GQA quality.

All the studied coefficients impacting possibly the GQA quality and defined in section 3.2.1, have been normalized and expressed in percent:

- The normalization of the distance of one chip to the separation hyper plane has been explained in section 2.6.2, and as each distance is in %, the average distance of several chips to the hyper plane is in % too:

$$normalized\_average\_distance_{hyper\ plane} = \frac{\sum_{chips} normalized\_distance_{hyper\ plane}}{nb_{chips}}$$
$$= \frac{\sum_{chips} 100 \frac{distance_{hyper\ plane}}{\max_{distance}}}{nb_{chips}}$$

Where $nb_{chips}$ is the number of accepted chips for the GQA estimation of a given SEVIRI image band, $distance_{hyper\ plane}$ is the distance to the separation hyper plane of a point representing the match of a chip in the matching acceptance decision space, and $\max_{distance}$ is the maximum distance to the separation hyper plane of a point representing a simulated perfect matching;

- The number of accepted chips is normalized by the total number of chips in a SEVIRI image:

$$normalized\_nb_{chips} = 100 \frac{nb_{chips}}{total_{chips}}$$

  Where $nb_{chips}$ is the number of accepted chips for the GQA computation of a given SEVIRI image band, and $total_{chips}$ is the global number of chips in a SEVIRI image (200 for the study);
- The normalization of the geographical distribution of the accepted chips on the SEVIRI image is detailed in Appendix B;
- The standard deviation of the offsets in each direction has been normalized by the maximum value we can have:

$$normalized\_std_{offX} = 100 \frac{std_{off_X}}{max_{std_{offX}}}$$
$$normalized\_std_{offY} = 100 \frac{std_{off_Y}}{max_{std_{offY}}}$$

  Where $std_{off_X}$ ($std_{offY}$ respectively) is the standard deviation of the offsets in East/West (South/North respectively) direction, and $max_{std_{offX}}$ ($max_{std_{offY}}$ respectively) is the maximum standard deviation of the offsets that it is possible to compute regarding the maximum search offset in the matching step *i.e.* 3 low resolution SEVIRI pixels.

So, for a given SEVIRI image band, thousands of combinations of accepted chips have been built by selecting a random number of chips (between 2 and the number of accepted chips for the studied SEVIRI band minus 1) among all the accepted chips by scenario 2 (with cloud information) for the studied SEVIRI image band. For each generated combination of accepted chips, the average distance of the accepted chips to the separation hyper plane, the number of accepted chips, the distribution of the accepted chips on the SEVIRI image, the standard deviation of the offsets in South/North and East/West directions of the accepted chips, and the GQA estimation computed with the selected chips in the combination only (and not with all the accepted chips for this SEVIRI image band) have been computed.

To assess the impact of the number of accepted chips on the GQA, all the combinations with the same number of accepted chips have been store in a bin, and for each bin, the mean GQA and the GQA standard deviation of the combinations in the bin have been computed as follow:

$$mean_{GQA} = \frac{\sum_{combinations} GQA}{nb_{combinations}}$$

$$std_{GQA} = \sqrt{\frac{\sum_{combinations}(GQA - mean_{GQA})^2}{nb_{combinations}}}$$

Where $nb_{combinations}$ is a number of the different combinations in a given bin, and $GQA$ is the estimated GQA using the accepted chip in a given combination. The mean GQA and the GQA standard deviation of the combinations in a bin are expressed in SEVIRI low resolution pixels.

Finally the GQA standard deviation can be plotted according to the number of chips.

The methodology to assess the impact of the other coefficients defined above on the GQA is a little bit more complex. Indeed regarding for example the geographical distribution of the accepted chips, there is no reason that several combinations of accepted chips have the same value of distribution of the accepted chips. Thus for these other coefficients, combinations with similar (and not identical) values of the studied coefficient are clustered: equally sized bins are created and populated with the different combinations. Then for each bin, the mean GQA and GQA standard deviation are computed as previously, and the mean value of the studied coefficient in the cluster is computed too. Finally the GQA standard deviation can be plotted according to the studied coefficient, or more precisely, according to the computed mean value of the studied coefficient. In practice, for a given coefficient, the size of a bin is equal to 2%.

The methodology allows to know for a mean value of a studied coefficient, the corresponding GQA standard deviation and so, the GQA precision. For a given coefficient, the lower the GQA standard deviation, the more precise the GQA estimation. This technique has been applied for each coefficient identified above, and for several SEVIRI image bands acquired at different seasons and day/night times.

### 3.2.2. Mean distance to the separation hyper plane

Using the methodology presented above, Figure 10 shows the GQA standard deviation according to the average distance of the accepted chips to the separation hyper plane for several SEVIRI image bands. For each SEVIRI image band, a trend curve has been plotted with its corresponding equation.

As expected, the higher the average distance of the accepted chips to the separation hyper plane, the more reliable the chip classification, the lower the GQA standard deviation and the more precise the GQA estimation. Thus, the average distance of the accepted chips to the separation hyper plane is a god mean to characterize the GQA quality. As is explained above, this coefficient reflects the reliability and the relevance of the acceptance of the chips that will be used to estimate the GQA.

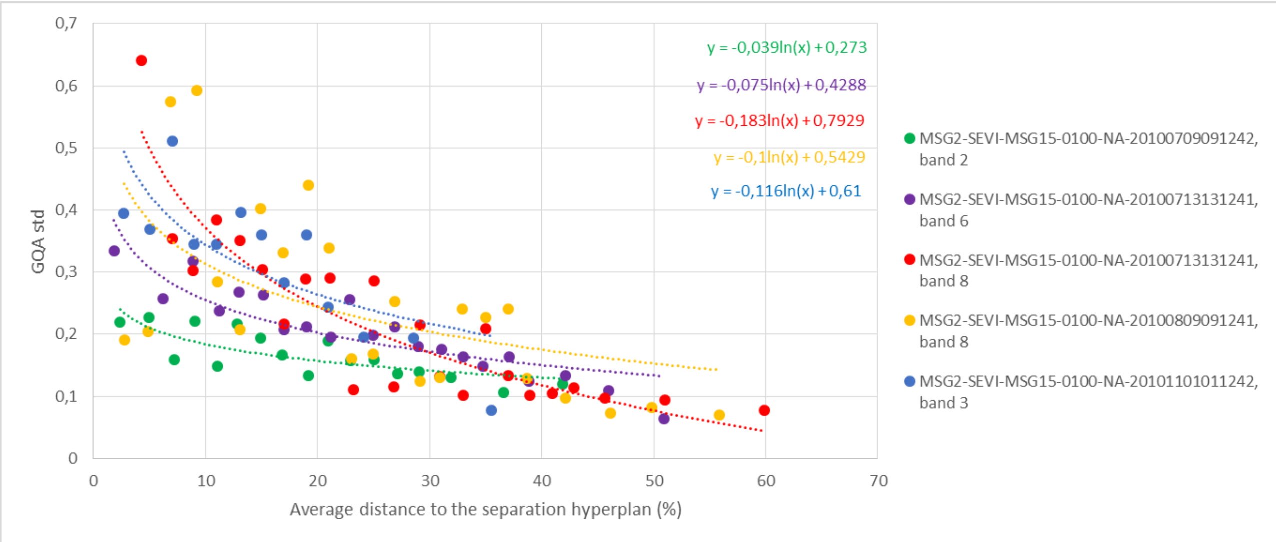
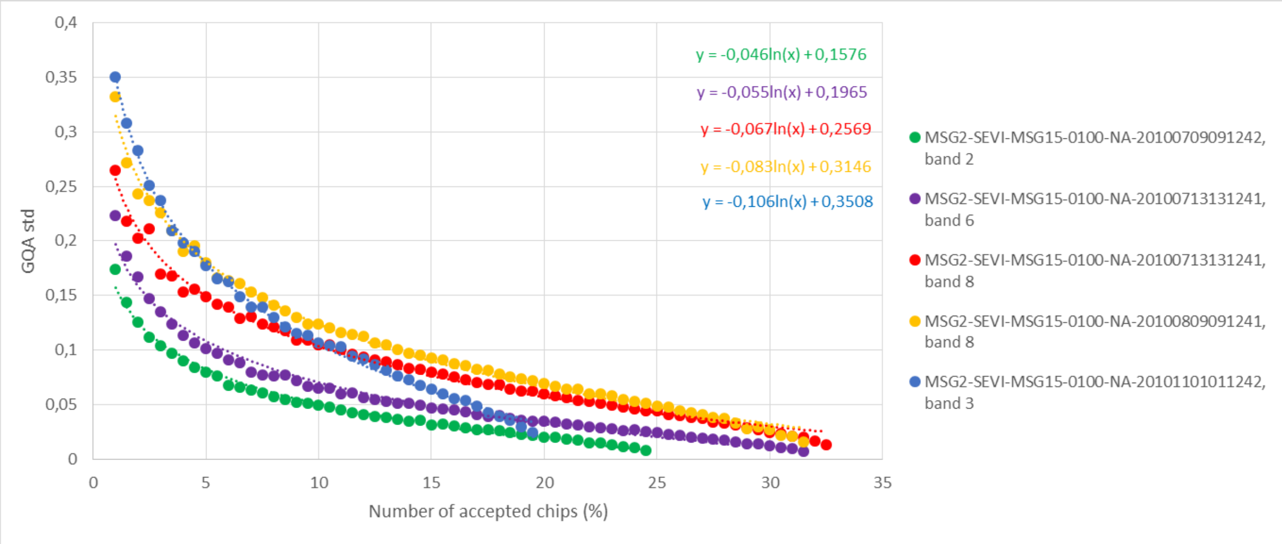



**Figure 10: Standard deviation of the GQA results according to the mean distance of the accepted chips to the separation hyper plane**

### 3.2.3. Number of accepted chips

Figure 11 presents the GQA standard deviation according to the number of accepted chips for several SEVIRI image bands. For each SEVIRI image band, a trend curve has been plotted with its corresponding equation.

The greater the number of accepted chips, the lower the GQA standard deviation and the more precise the GQA estimation. Thus, the number of accepted chips is a god mean to characterize the GQA quality. As is explained above, this coefficient reflects if the GQA estimation is reliable on the whole SEVIRI image band or not.




### 3.2.4. Figure 11: Standard deviation of the GQA results according to the number of accepted chipsFormulation

Regarding the analyses on the different identified coefficients, the following formula is used to assess the GQA quality:

$$GQA_{quality} = \left(-0.06 \times \ln\left(number_{chips}\right) + 0.25\right) + \left(-0.1 \times \ln\left(\overline{distance}\right) + 0.5\right)$$

Where $number_{chips}$ is the number of accepted chips to estimate the GQA for the studied SEVIRI band and $\overline{distance}$ is the average distance of the accepted chips to the separation hyper plane.

The different constants have been chosen as mean values of the constants of the trend curves in Figure 10 and Figure 11. These constants have not a primordial impact on the result, the shape of the function is more important.

Regarding this formulation, the GQA quality is always positive, and the better the GQA quality, the lower the GQA quality value. The minimum GQA quality value, for the best GQA quality, is equal to 0.0132, when the number of accepted chips and the average distance of the accepted chips to the separation hyper plane are equal to 100% (*i.e.* their maximal values).

By design, **the proposed GQA quality formulation assesses the GQA standard deviation and so the GQA precision of a SEVIRI image band, but the formulation provides no information on the GQA accuracy**. However, we have no information on the GQA accuracy of the SEVIRI image bands so it is not possible to elaborate and validate a mathematical formulation to assess the GQA accuracy, and hence, have a perfect definition of the GQA quality.

For the other coefficients identified in section 3.1 (geographical distribution of the accepted chips and standard deviation of the offsets of the accepted chips in South/North and East/West directions) as having a possible impact on the GQA quality, no relation with the GQA standard deviation is found, as detailed in Appendix B.

### 3.2.5. Validation

In order to demonstrate that the proposed formulation for the GQA quality enables to estimate the GQA standard deviation, and so the quality of the GQA, the GQA quality formulation has been validated by using the same methodology as the one presented in section 3.2.1.

For a given SEVIRI image band, thousands of combinations of accepted chips have been built by selecting a random number of chips (between 2 and the number of accepted chips for the studied SEVIRI band minus 1) among all the accepted chips by scenario 2 (with cloud information) for the studied SEVIRI image band. For each generated combination of accepted chips, the GQA quality according the proposed formulation and the GQA estimation computed with the selected chips in the combination only have been computed. Then combinations with similar values of GQA quality are clustered. Then for each cluster, the mean GQA quality, the mean GQA and the GQA standard deviation of the combinations in the cluster are computed. In this case, the clusters have been built by clustering in each pool 1.000 combinations of accepted chips with the closest GQA quality value between each other. Finally the mean GQA and GQA standard deviation can be plotted according to the GQA quality.

Figure 12, Figure 13, Figure 14, Figure 15 and Figure 16 show the variations of the mean GQA estimated by scenario 2, according to the GQA quality for several SEVIRI image bands, and error bars highlight the standard deviation of the estimated GQA.

It can be noted that the lower the GQA quality is, the lower the GQA standard deviation is and the more the GQA estimation converges to a constant value. So the GQA quality formulation allows to assess the GQA precision, and so the quality of the GQA.

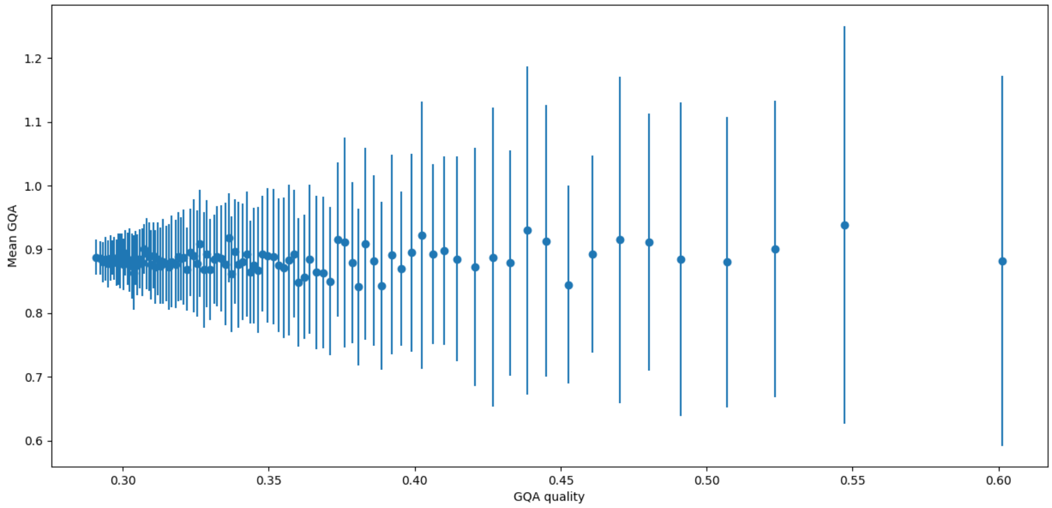



**Figure 12: Mean GQA estimated by scenario 2 and GQA std according to GQA quality for the band 0 of image MSG2-SEVI-MSG15-0100-NA-20100409091241**

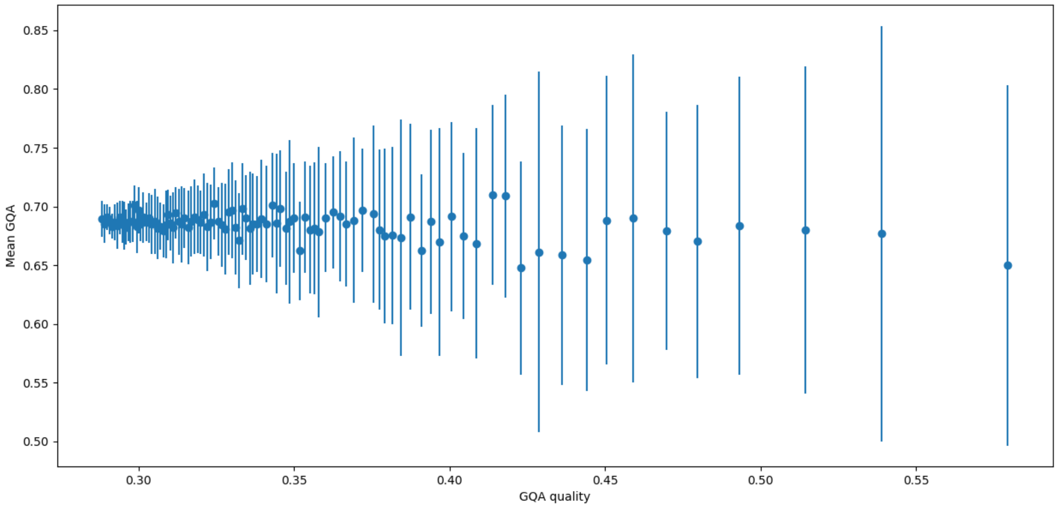



**Figure 13: Mean GQA estimated by scenario 2 and GQA std according to GQA quality for the band 2 of image MSG2-SEVI-MSG15-0100-NA-20100605051242**

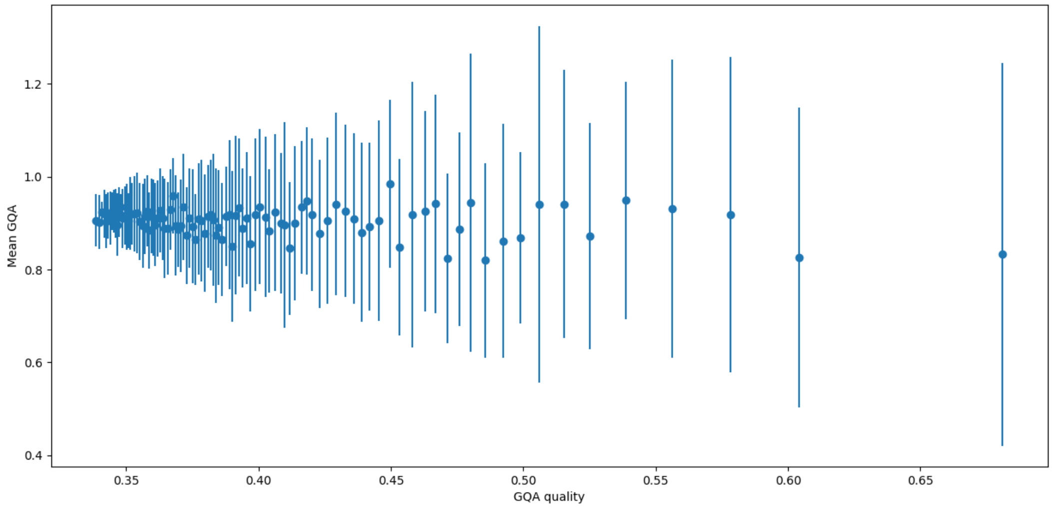



**Figure 14: Mean GQA estimated by scenario 2 and GQA std according to GQA quality for the band 6 of image MSG2-SEVI-MSG15-0100-NA-20101221211241**

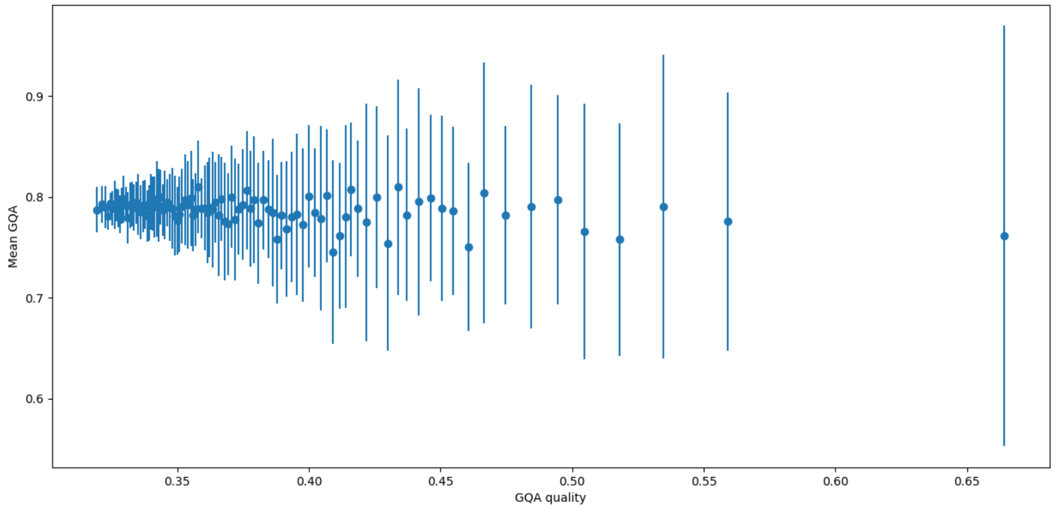



**Figure 15: Mean GQA estimated by scenario 2 and GQA std according to GQA quality for the band 7 of image MSG2-SEVI-MSG15-0100-NA-20100717171242**

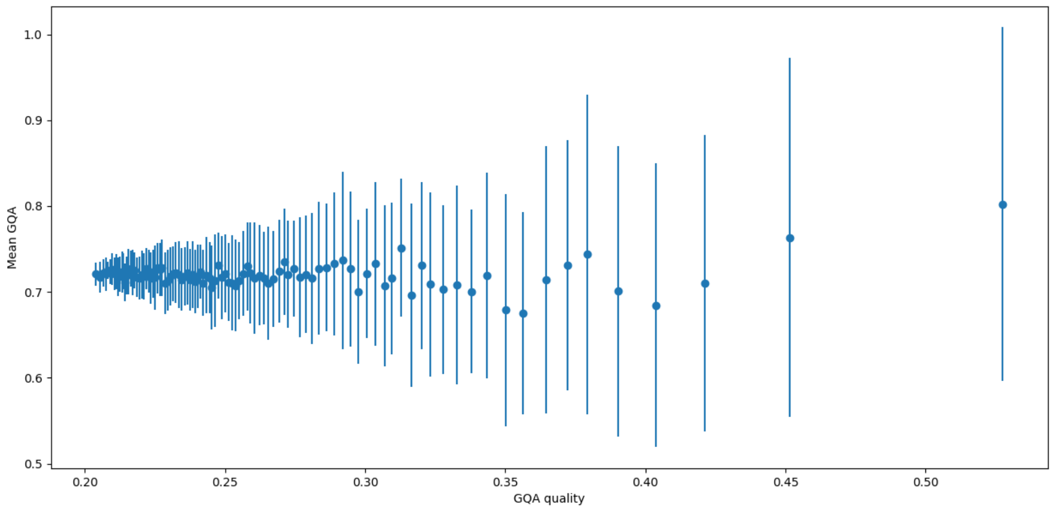



**Figure 16: Mean GQA estimated by scenario 2 and GQA std according to GQA quality for the band 8 of image MSG2-SEVI-MSG15-0100-NA-20100109091242**

Moreover the GQA quality formulation should be validated to assess the GQA estimated with scenario 1 also. Thereby Figure 17, Figure 18, Figure 19, Figure 20 and Figure 21 show the variations of the mean GQA estimated by scenario 1 (without cloud information), according to the GQA quality for the same SEVIRI image bands used to validate the GQA quality formulation for scenario 1, and error bars highlight the standard deviation of the estimated GQA.

As in previous cases with scenario 2, the lower the GQA quality is, the lower the GQA standard deviation computed by using scenario 1 is and the more the GQA estimation tends toward a constant value. **So the proposed GQA quality formulation allows to assess the GQA precision, and so the quality of the GQA estimated by both scenarios**.

It has not been analysed if the GQA quality of different SEVIRI image bands can be compared. To do so, the previous study should be repeated by taking together several random combinations of accepted chips among the accepted chips of several SEVIRI image bands and analysing their GQA and GQA quality values. However the previous demonstration shows that the GQA quality values of a same SEVIRI image band obtained from different ways can be compared. **That highlights the scenario providing the more precise GQA estimation for each SEVIRI image band.** And, according to the defined GQA quality formulation, the **lower the GQA quality value, the higher the precision of the GQA estimated.**

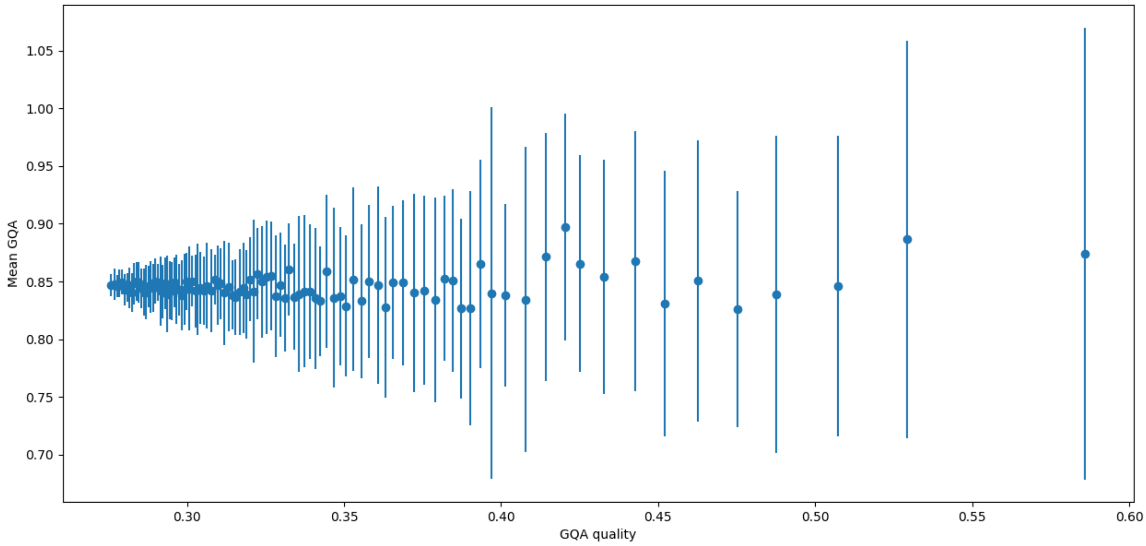



**Figure 17: Mean GQA estimated by scenario 1 and GQA std according to GQA quality for the band 0 of image MSG2-SEVI-MSG15-0100-NA-20100409091241**

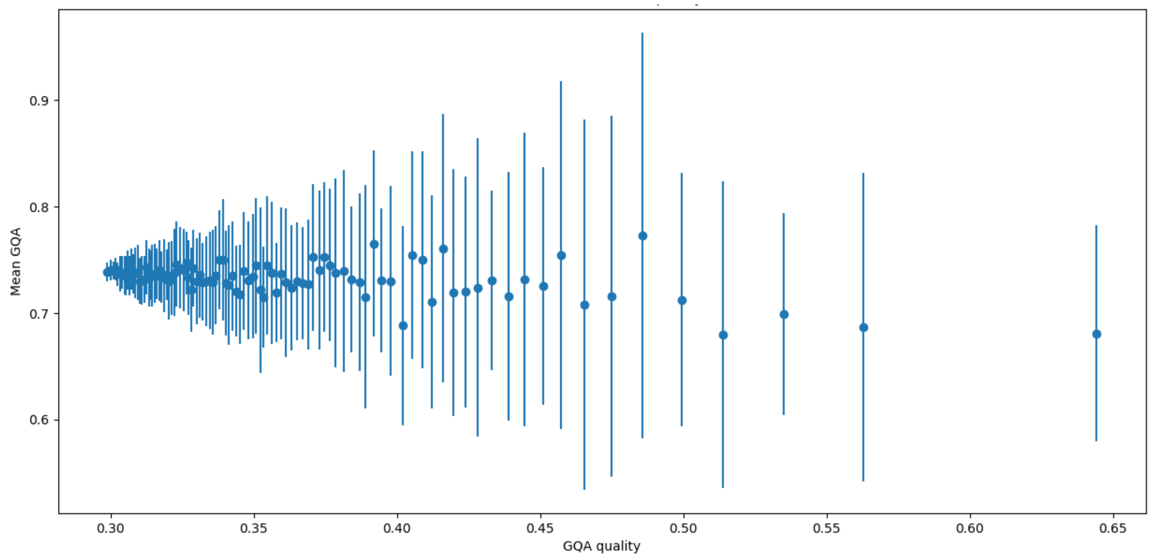



**Figure 18: Mean GQA estimated by scenario 1 and GQA std according to GQA quality for the band 2 of image MSG2-SEVI-MSG15-0100-NA-20100605051242**

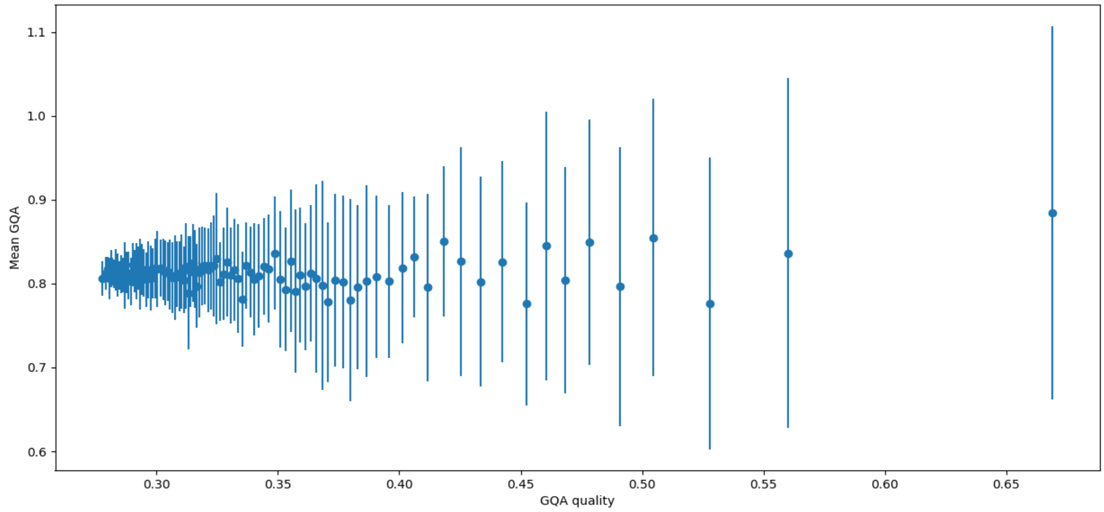



**Figure 19: Mean GQA estimated by scenario 1 and GQA std according to GQA quality for the band 6 of image MSG2-SEVI-MSG15-0100-NA-20101221211241**

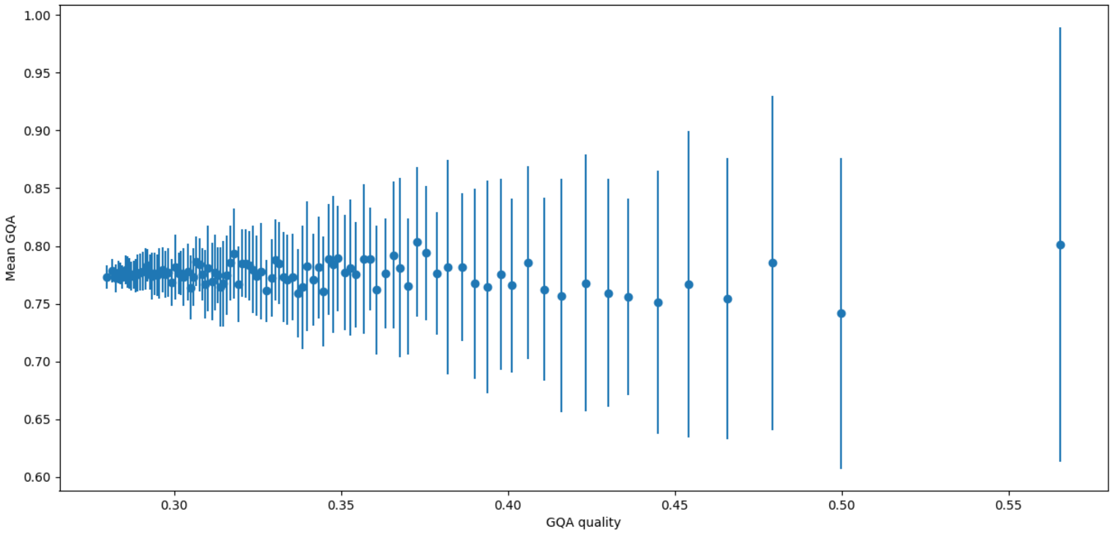



**Figure 20: Mean GQA estimated by scenario 1 and GQA std according to GQA quality for the band 7 of image MSG2-SEVI-MSG15-0100-NA-20100717171242**

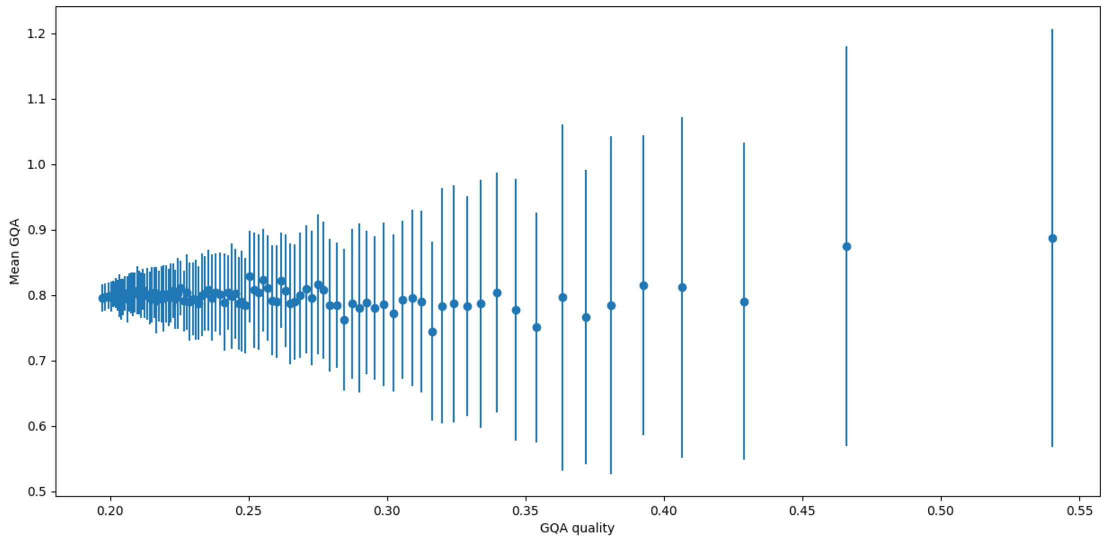



**Figure 21: Mean GQA estimated by scenario 1 and GQA std according to GQA quality for the band 8 of image MSG2-SEVI-MSG15-0100-NA-20100109091242**

## 3.3. First results

Both scenarios have been run on a selection of 72 input SEVIRI images distributed over the 12 months of the year 2010 and on the 24 hours of a day in order to process images acquired at all seasons and at all day/night time.

The GQA estimations (in low resolution SEVIRI pixels) and GQA quality of the different SEVIRI image bands computed by the two scenarios are plotted in Figure 22 and Figure 23 respectively. Each Id on the horizontal axis corresponds to a SEVIRI image band (organized by acquisition date firstly, and by SEVIRI band number secondly) and for a given Id, there is two GQA and GQA quality values, one, in orange, provided by scenario 1 and the other, in blue, provided by scenario 2. The comparison of these two values (for the GQA estimation, and the GQA quality respectively) for a given Id is important, more than the comparison of the GQA estimation or GQA quality provided by the same scenario for different SEVIRI images and different bands. Indeed, no particular behaviour occurs for a given SEVIRI band.

The GQA estimations seem to be in accordance with the expected known bias offset of the SEVIRI images acquired before 2017 (0.5 low resolution pixel to the North and to the West implying an average bias offset of 0.7 low resolution pixel).

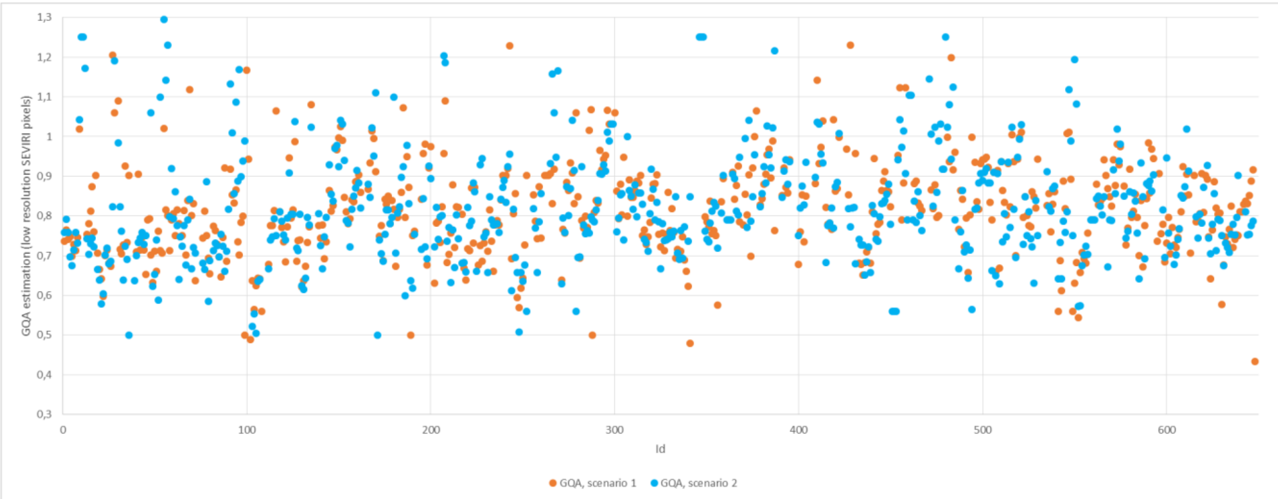



**Figure 22: GQA results obtained with the two scenarios for different SEVIRI images and different bands**

Figure 23 shows that the GQA quality is between the same range of values for the majority of the SEVIRI images and bands. This confirms that there is not a special behaviour according to the SEVIRI bands.

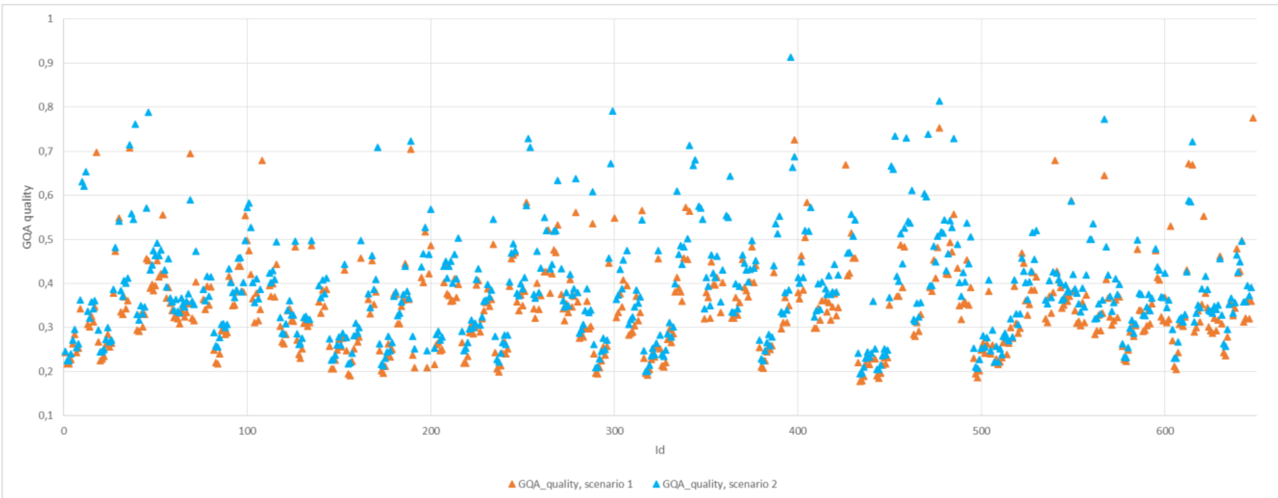



**Figure 23: GQA quality obtained with the two scenarios for different SEVIRI images and different bands**

Among the 648 (72 SEVIRI images, 9 bands per image) GQA and GQA quality values computed by each one the two scenarios, 104 are not studied because no chip have been accepted by one of the two scenario, or by the two scenarios. Among the remaining SEVIRI image bands, 292 GQA estimations computed by scenario 1 are higher than the corresponding GQA estimated by scenario 2 and 252 GQA estimations computed by scenario 2 are higher than the corresponding GQA estimated by scenario 1. So one scenario does not estimate always higher or smaller GQA values than the other scenario. Moreover the mean value between the GQA results computed by the two scenarios is equal to -0.009 low resolution SEVIRI pixels, which is very small regarding the range of GQA estimation in Figure 22, meaning that the GQA estimated by both scenarios are close each other.

In addition, as is described in Table 2, 524 GQA quality values computed by scenario 1 are lower than the corresponding GQA quality values computed by scenario 2 and 20 GQA quality values computed by scenario 2 are lower than the corresponding GQA quality values computed by scenario 1.

In average in case the GQA quality computed by scenario 2 is lower (better) than the GQA quality computed by scenario 1 (second row in Table 2):

- The GQA quality values given by the two scenarios are higher (worse) than in case the GQA quality computed by scenario 1 provides the most reliable GQA estimation (the best GQA quality);
- The GQA estimation provided by scenario 1 is closer to the best estimation than the GQA estimation given by scenario 2 in case scenario 1 provides the most reliable GQA estimation, regarding the average differences between the GQA quality values.

**Table 2: Analysis of GQA quality of the first implementation**

| | Number of cases | Mean GQA quality 1 | Mean GQA quality 2 | Mean difference between the GQA quality values |
|---|---|---|---|---|
| **Scenario 1 is the best** | 524 | 0.3315 | 0.3629 | 0.031 |
| **Scenario 2 is the best** | 20 | 0.5031 | 0.4836 | 0.020 |

Even if the GQA results estimated by the two scenarios are close each other, for 524 SEVIRI image bands (the majority), the GQA estimation computed by scenario 1 is more reliable than the GQA estimation computed by scenario 2. Thus, **scenario 1 (without cloud information) seems to give more reliable GQA estimation than scenario 2 (with cloud information)**. This result was not expected and will be analysed in the next section by studying the two components impacting the GQA quality: the number of accepted chips for the GQA estimation, and the average distance of these accepted chips to the separation hyper plane.

# 4. Impact of undetected clouds on GQA

This section is dedicated to the analysis of the impact of clouds on the GQA quality. First and foremost the results of scenario 1 and scenario 2 are compared and analysed. Then modifications are applied to the two scenarios in order to improve the results. Finally the results obtained with the new scenarios are presented.

## 4.1. Analysis of the chip acceptance by the two scenarios

### 4.1.1. Presentation

In order to explain the differences of GQA quality obtained with the two scenarios, and the impact of the clouds on the GQA quality, the two components impacting the GQA quality (the number of accepted chips for the GQA estimation, and the average distance of these accepted chips to the separation hyper plane reflecting the quality of the matching) are investigated. Then the classification of the chips by the two scenarios is analysed.

We recall here that for concision we refer to “chip accepted” or “chip rejected” (instead of “image chip band”), but the scenarios are applied to each SEVIRI image band separately as explained previously, so each band of an image chip is processed independently.

For this analysis the same SEVIRI image as in section 3.3 is processed. Table 3 explains the acronyms used for the different classification cases. The first two letters refer to the chip acceptance by scenario 1 (Matching Accepted or Matching Rejected). The letters after the underscore refer to the chip acceptance by scenario 2 ( Cloud ok Matching Accepted, Cloud ok Matching Rejected or Cloud Rejected).

**Table 3: Acronyms used for the different classification cases**

| | | **Scenario 2 (with cloud information)** | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| **Scenario 1 (without cloud information)** | Chip accepted | MA_CMA | MA_CMR | MA_CR |
| | Chip rejected | MR_CMA | MR_CMR | MR_CR |

To assess if omitting cloud information degrades or not the GQA quality and how, the chips in each classification case are analysed below.

### 4.1.2. Detailed analysis of each classification cases

An example of classification case MA_CMA is illustrated in Figure 24. The second row presents the results obtained with scenario 1 whereas the third row presents the results obtained with scenario 2. The mask are plotted in white on the cloud mask. Here, clouds have no direct impact on the acceptance of the chips.

However the mean distance to the separation hyper plane of the chips accepted by scenario 1 is equal to 21.2% whereas the mean distance the separation hyper plane of the chips accepted by scenario 2 is equal to 17.6% highlighting a small negative impact of using cloud information for this classification case because the chips are accepted with more reliability by scenario 1 than by scenario 2.

The mean cloud cover for all the chips classified as MA_CMA is equal to 12.6%.

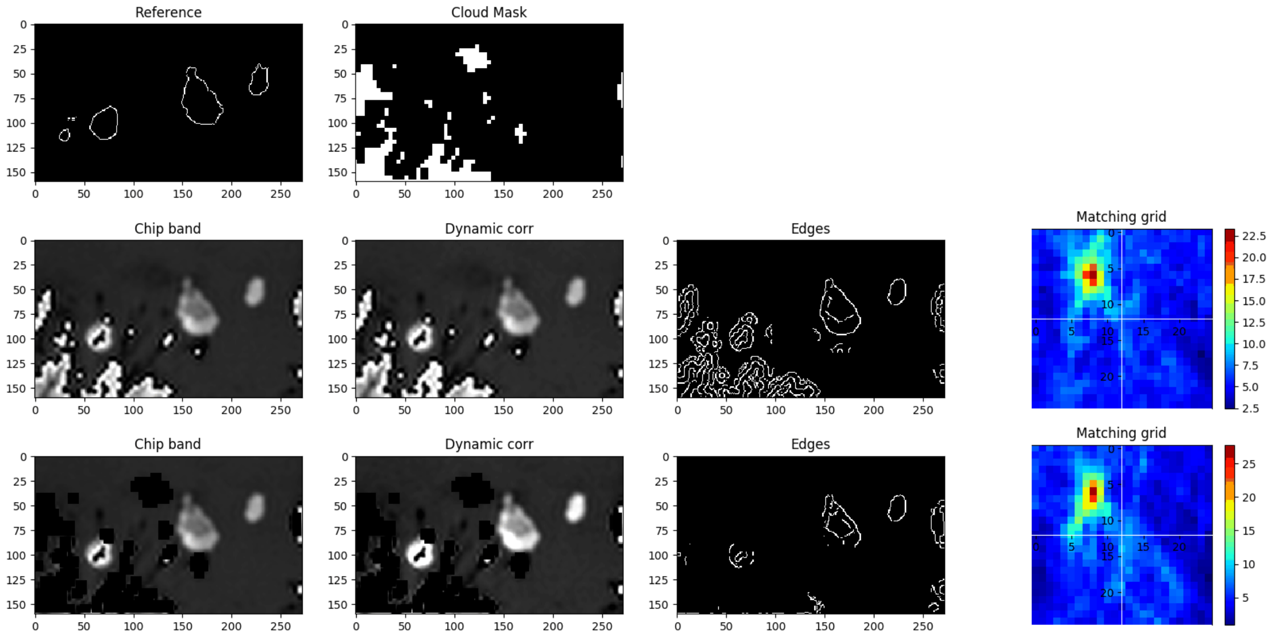



**Figure 24: Matching results for the band 6 of chip 178 of image MSG2-SEVI-MSG15-0100-NA-20100113131242**

An example of classification case MR_CMR is illustrated in Figure 25. Here the matching does not give a good result for the two scenarios because of bad illumination conditions and/or clouds and clouds have the same impact on both scenarios (no direct impact on the rejection of the chips in case of bad illumination conditions or a direct impact in case of presence of clouds near to the coastlines).

However the mean distance to the separation hyper plane of the chips rejected by scenario 1 is equal to -16.9.3% and the mean distance the separation hyper plane of the chips rejected after matching by scenario 2 is equal to -15.8% highlighting a small negative impact of using cloud information for this classification case because the chips are rejected after matching with more reliability by scenario 1 than by scenario 2.

The mean cloud cover for all the chips classified as MR_CMR is equal to 22.0%.

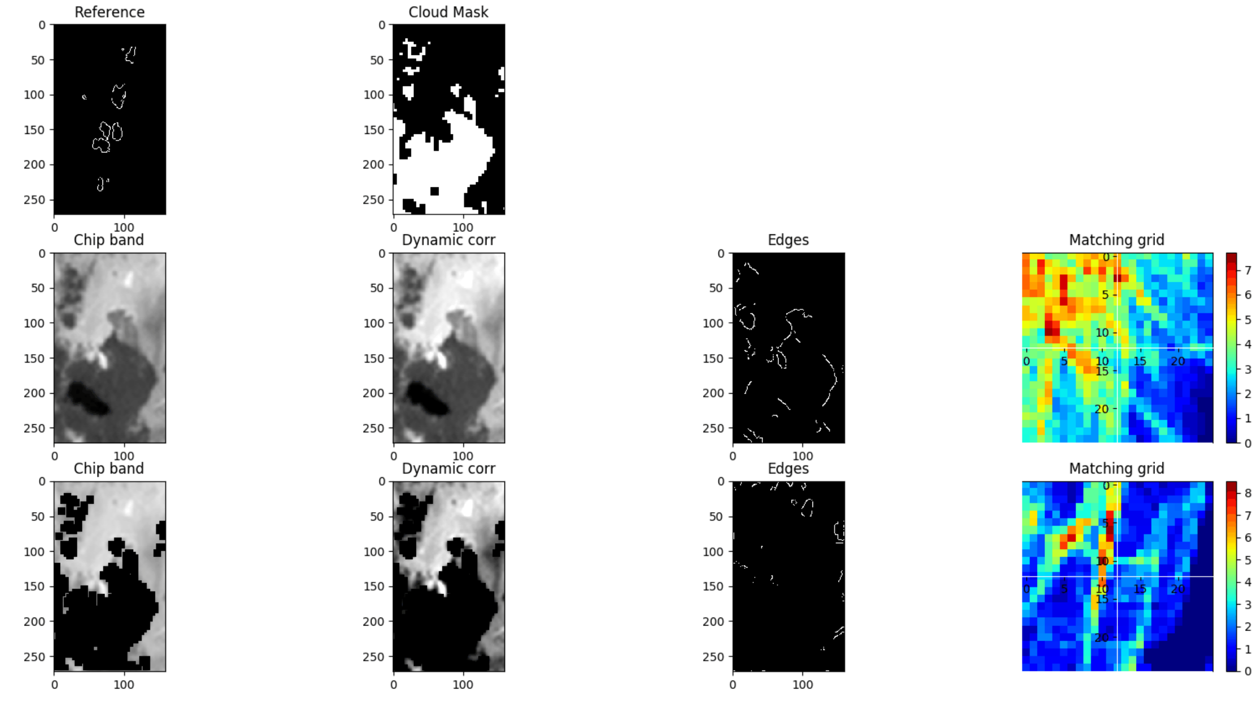



**Figure 25: Matching results for the band 3 of chip 1 of image MSG2-SEVI-MSG15-0100-NA-20100101011242**

Figure 26 illustrates an example of classification case MR_CR. In this example, the chip is rejected by scenario 2 because of a high amount of clouds whereas with scenario 1 the chip is rejected because the matching result is not good enough. However the edges seem correct and the matching result should not be far to be accepted. Moreover, the matching result obtained with scenario 2 is not so bad. Of course, this case is not representative of the majority of the chips in this classification case but shows that the two scenarios can be improved in order to obtain a better chip acceptance criterion.

The mean distance to the separation hyper plane of the chips rejected by scenario 1 is equal to -20.3% and the mean distance the separation hyper plane of the chips rejected because of clouds by scenario 2 is equal to -16.8% highlighting a small negative impact of using cloud information for this classification case because the chips are rejected after matching with more reliability by scenario 1 than by scenario 2.

The mean cloud cover for all the chips classified as MR_CR is equal to 83.8%.

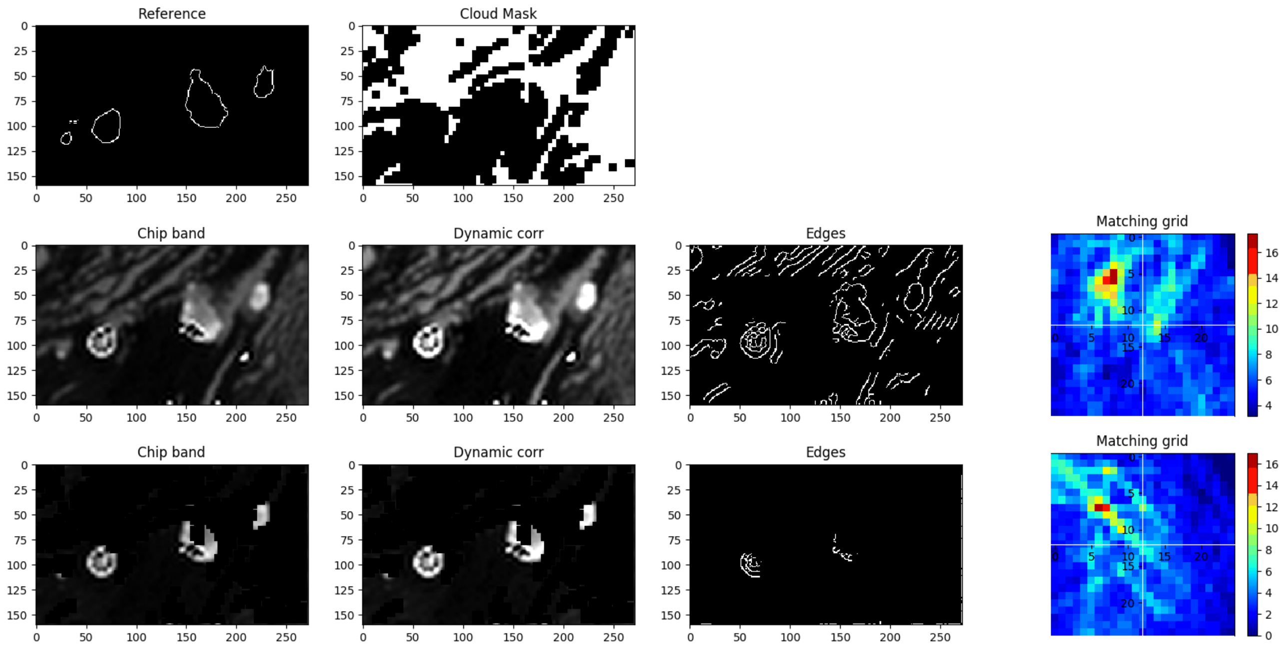



**Figure 26: Matching result for the band 3 of chip 178 of image MSG2-SEVI-MSG15-0100-NA-20100513131242**

The explanation of the case MA_CR is illustrated by an example in Figure 27. Without using the cloud mask the coastlines are clearly visible despite of the presence of clouds and the matching works well for scenario 1 (second row of the figure). However, as the percentage of clouds is higher than the defined threshold these chips are rejected by the cloud acceptance criterion. So in this case, using cloud information has a negative impact on the chip acceptance.

The mean distance to the separation hyper plane of the chips accepted by scenario 1 is equal to 12.2% and the mean distance the separation hyper plane of the chips rejected because of clouds by scenario 2 is equal to -2.9% highlighting that in average, even if these chips were accepted by the cloud acceptance criterion of scenario 2, they would have been rejected after matching because the mean distance to the hyper plane is negative. It is a good argument to omit cloud information.

The mean cloud cover for all the chips classified as MA_CR is equal to 67.4%.

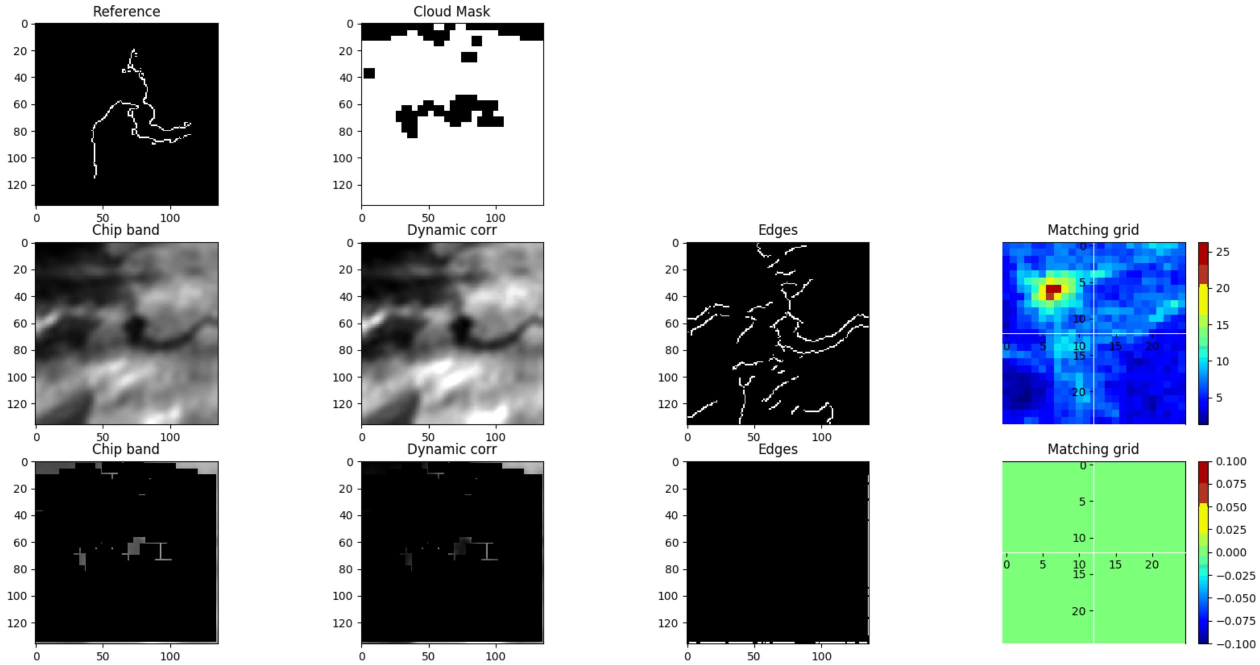



**Figure 27: Matching results for the band 1 of chip 24 of image MSG2-SEVI-MSG15-0100-NA-20100109091242**

Even if the previous chip were accepted by the cloud acceptance criterion in scenario 2, it is obvious that useful edges for the matching are masked by the cloud mask and this chip would have been rejected by the matching criterion. It is what happens for the chips classified as MA_CMR, as is illustrated by an example in Figure 28, and this highlights a negative impact on the matching process of using cloud information.

The mean distance to the separation hyper plane of the chips accepted by scenario 1 is equal to 7.2% and the mean distance the separation hyper plane of the chips rejected after matching by scenario 2 is equal to -5.3%.

The mean cloud cover for all the chips classified as MA_CMR is equal to 22.3%.

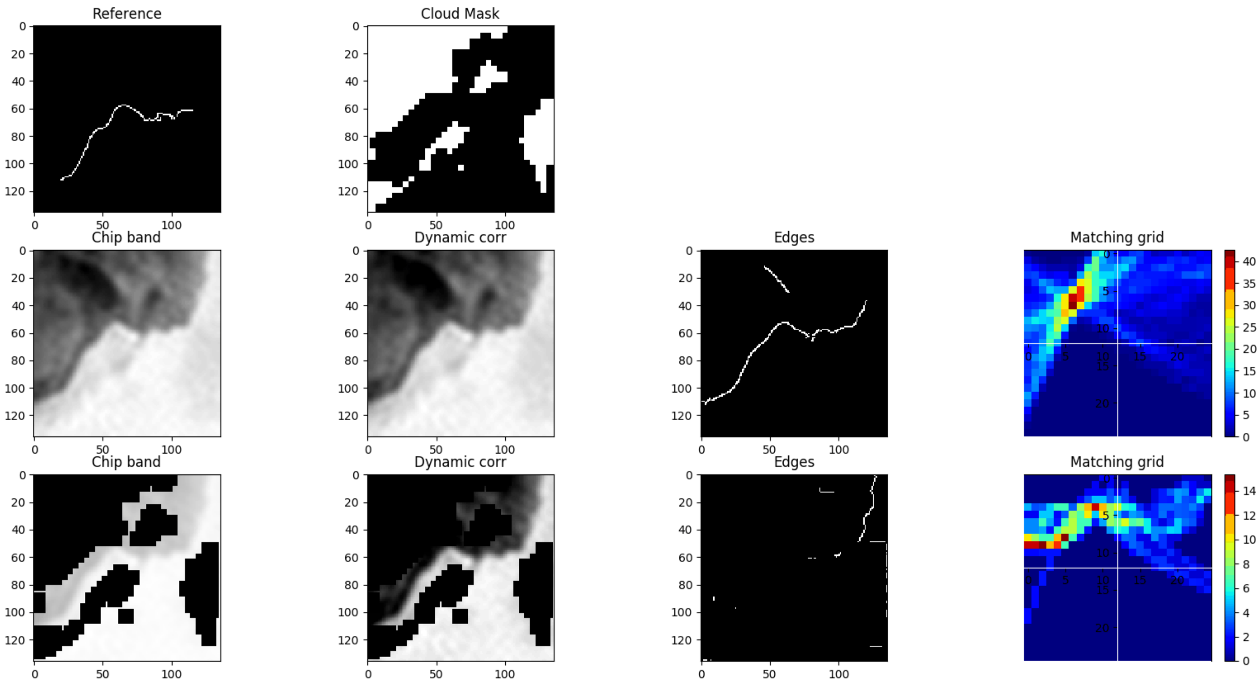



**Figure 28: Matching result for the band 3 of chip 21 of image MSG2-SEVI-MSG15-0100-NA-20100101011242**

On the other hand, using cloud information could have a positive impact by masking some clouds or edges contaminating the matching process in scenario 1: it is what happens for the chips classified as MR_CMA, as is illustrated by the example in Figure 29. In this case, masking the clouds allows to have a better contrast after contrast enhancement of the chip, and to detect the coastline with scenario 2 whereas the coastline is not detected in scenario 1.

The mean distance to the separation hyper plane of the chips rejected by scenario 1 is equal to -9.7% and the mean distance the separation hyper plane of the chips accepted by scenario 2 is equal to 4.6%.

The mean cloud cover for all the chips classified as MR_CMA is equal to 23.9%.

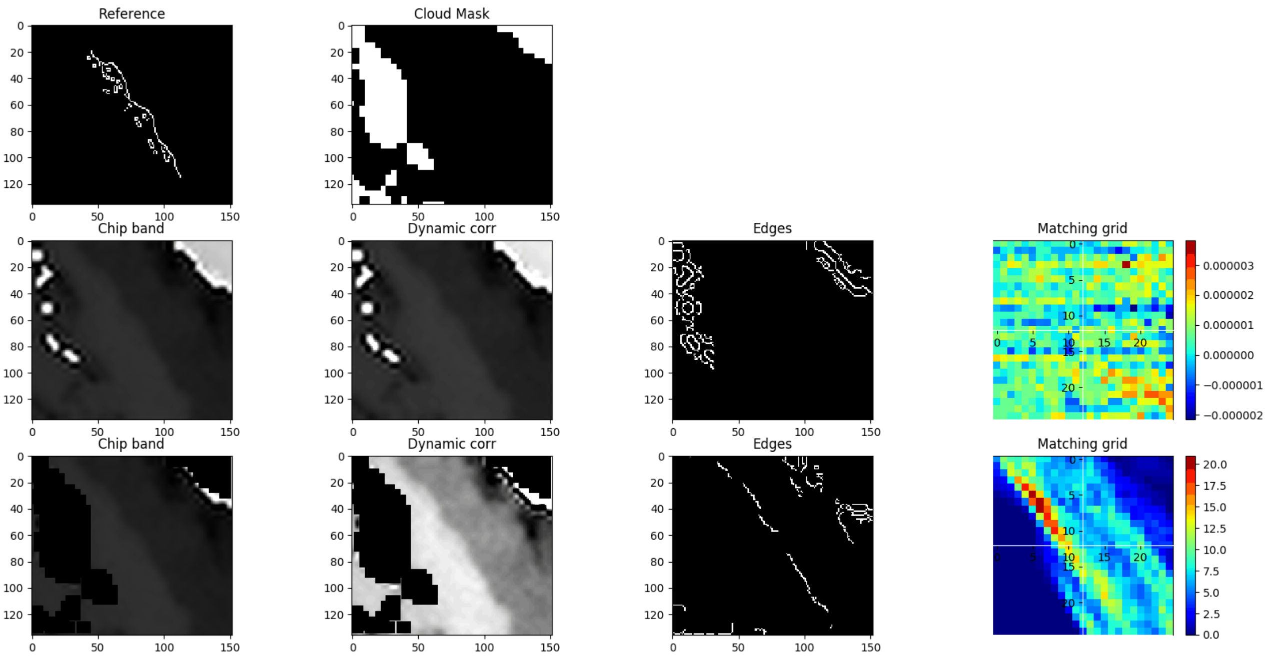



**Figure 29: Matching result for the band 8 of chip 5 of image MSG2-SEVI-MSG15-0100-NA-20100105051241**

### 4.1.3. Global synthesis

The global synthesis of this analysis regarding the global shape of the chips only is presented in Table 4.

**Table 4: Description of the different classification cases regarding the shape of the chips**

| | | Scenario 2 (with cloud information) | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| **Scenario 1 (without cloud information)** | Chip accepted | Chip accepted by the two scenarios (low amount of clouds and good illumination conditions): **no impact of using cloud information** in this case. | Chip accepted without masking the clouds and rejected by masking the clouds (low amount of clouds and good illumination conditions): **negative impact of using cloud information** on the result of scenario 2, which masks useful edges for the matching. | Chip accepted without masking the clouds but rejected by scenario 2 because of high amount of clouds: **negative impact of using cloud information** on the result of scenario 2 because edges are detected under the clouds by scenario 1. |
| | Chip rejected | Chip accepted by masking the clouds and rejected without masking the clouds: **positive impact of using cloud information** on the result of scenario 2 which masks edges contaminating the matching process in scenario 1 despite of the low amount of clouds. | Chip rejected by the two scenarios because of bad illumination condition (but low amount of clouds): **no impact of using cloud information** in this case. | Chip rejected because of high amount of clouds by scenario 2 and rejected after matching by scenario 1: **no impact of using cloud information** in this case (with scenario 1 the matching is rejected because it is contaminated by the clouds). |

Depending on the classification cases, using cloud information could have a positive, negative or no impact on the chip acceptance/rejection by each one of the two scenarios, then the analysis should be carried on.

Table 5 presents the number of chips classified in the different cases (the same image chip bands as processed previously in section 3.3 have been used).

First and foremost the global number of accepted chips by scenario 1 is higher than the number of accepted chips by scenario 2, suggesting that using cloud information has a negative impact in terms of number of accepted chips.

Moreover, over the half of the chips is rejected by scenario 2 because of the high amount of clouds (higher than 50% of the pixels of the chips, a sensitive study concerning this threshold is presented in section 4.3.2.3). But, the majority of these chips rejected by scenario 2 because of a high amount of clouds are also rejected by scenario 1 after matching (66 097 among 68 085). So, maybe cloud information could be omitted without degrading the GQA quality.

Thus, **using cloud information has a negative impact on the number of accepted/rejected chips**. But the quality of the accepted/rejected chips should be analysed too in order to conclude on the possibility to omit cloud information in the GQA process.

**Table 5: Number of chips in the different classification cases, first implementation**

| | | **Scenario with cloud information** | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds | | |
| **Scenario without cloud information** | Chip accepted | 13 432 | 2 927 | 1 988 | 18 347 | **Total** |
| | Chip rejected | 1 520 | 43 636 | 66 097 | 111 253 | |
| | | 14 952 | 46 563 | 68 085 | | |
| | | **Total** | | | | |

Table 6 details the average distance of the chips to the separation hyper plan for the different classification cases (the first percentage, in blue, is the average distance of the chips to the separation hyper plane classified by scenario 2 and the second one, in yellow, is the average distance of the chips to the separation hyper plane classified by scenario 1). The average distance to the separation hyper plane reflect the reliability of the classification. Higher the absolute value of the distance, more reliable the classification. It is recalled than for the accepted chips, the distance to the separation hyper plane is positive, whereas it is negative for the chips rejected by the matching acceptance criterion.

The mean distance to the separation hyper plane of the accepted chips by scenario 1 is globally higher than the mean distance to the separation hyper plane of the accepted chips by scenario 2 (in the classification cases MA_CMA and MA_CMR vs. MR_CMA). Then, the absolute value of the mean distance to the separation hyper plane of the rejected chips by scenario 1 is globally higher than the absolute value of the mean distance to the separation hyper plane of the rejected chips by scenario 2 (in the classification cases MR_CMR, MR_CR and MR_CMA vs. MA_CMR and MA_CR). Thus, **the chips are accepted or rejected with more reliability by scenario 1** because the distance between the chips and the separation hyper place is larger. .

It can be noted that even if the chips rejected by the cloud acceptance criterion of scenario 2 have been accepted, because of the cloud mask masks the coastlines these chips would have been rejected by the matching acceptance criterion in average (as the mean distance to the separation hyper plane of these chips is negative).

**Table 6: Mean distance of the chips to the separation hyper plane for the different classification cases**

| | | Scenario 2 (with cloud information) | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| Scenario 1 (without cloud information) | Chip accepted | 17.6%<br>21.2% | -5.3%<br>7.2% | -2.9%<br>12.2% |
| | Chip rejected | 4.6%<br>-9.7% | -15.8%<br>-16.9% | -16.8%<br>-20.3% |

Now it has been demonstrated that the reliability of the chip classification by scenario 1 is not worse (indeed better) than the chip classification by scenario 2, the impact of the percentage of clouds on the chip acceptance process by the two scenarios is studied. Table 7 presents the mean cloud coverage for the different classification cases.

When the cloud cover is low, the chips are accepted by the two scenarios (classification case MA_CMA) because the matching process works correctly.

As the cloud cover is very high, the chips are rejected by the two scenarios (by the matching acceptance criterion for the scenario1 and by the cloud acceptance criterion for scenario 2, classification case MR_CR).

However, as the cloud cover is high the chips could be accepted by scenario 1 (with a good reliability, as seen previously) whereas they are rejected by the cloud acceptance criterion of scenario 2 (classification case MA_CR).

**Table 7: Mean cloud cover for the different classification cases**

| | | Scenario 2 (with cloud information) | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| Scenario 1 (without cloud information) | Chip accepted | 12.6% | 22.3% | 67.4% |
| | Chip rejected | 23.9% | 22.0% | 83.8% |

**To summarize, not using cloud information allows to accept more chips for the GQA estimation. Moreover not using cloud information increases the chip acceptance reliability and enables to accept with good reliability some chips with a high amount of clouds that are rejected by the cloud acceptance criterion of scenario with cloud information. This explains why the GQA quality is better in average for GQA estimated by scenario 1 than by scenario 2. Considering the previous analysis, cloud information can be omitted without degrading the chip acceptance results and thereby the GQA quality. However the two scenarios can be improved.**

So in order to improve scenario 1 the feasibility to transfer some chips from case MR_CMA to case MA_CMA, from case MR_CMR to case MA_CMR or from case MR_CR to case MA_CR has to be assessed.

More globally, scenario 2 could be improved too, and the feasibility to transfer some chips from cases MA_CMR or MA_CR to MA_CMA or from cases MR_CMR or MR_CR to MR_CMA should be assessed too.

## 4.2. Improvement of the scenarios

### 4.2.1. Coastline mask

The last example presented above in Figure 26 shows that some modifications in scenario 1 could enable to transfer some chips from case MR_CR to case MA_CR (but also from case MR_CMA to case MA_CMA or, from case MR_CMR to case MA_CMR) and to improve the GQA results (by improving the chip acceptance).

Indeed the matching process presented in Figure 26 shows that the coastlines are well detected but that the matching is contaminated by some edges detected around clouds far from the coastlines. To get rid of this issue, a coastline mask is applied to the image chip band at the beginning of each scenario. This coastline mask is generated through the application of a dilatation on the edges of a given reference image. The radius of the structuring element used for the dilatation is equal to the maximum search distance in the matching process (*i.e.* 3 low resolution SEVIRI pixels).

In addition to getting rid of artefacts far from the edges, the coastline mask has another advantage: the contrast enhancement is performed over pixels that are not masked by the coastline mask only. So if a bright cloud is far from a coastline it will be masked and, after contrast enhancement, the coastlines will be detected more easily. Then the coastline mask will have the same effect as the cloud mask in the example shown in Figure 29.

### 4.2.2. Cloud threshold in the coastline mask

In order to give a better importance to the cloud information in scenario 2, chips will be rejected by the cloud criterion if the cloud amount computed over pixels that are not masked by the coastline mask only is higher than 50% of the pixels that are not masked by the coastline mask. Thus clouds far from the coastlines will not be taken into account for the cloud acceptance. This is not disturbing as the clouds masked by the coastline mask (so far from the coastlines) will not influence the matching process.

## 4.3. Analysis of the results after improvement of the scenarios

### 4.3.1. Validation of the improvement of the scenarios

Table 8 presents how the chips classified in the different cases with the first implementation are classified into the different cases by the new implementation (the same image chip bands as processed previously have been used).

**Table 8: Repartition of the chips in the classification cases of the two runs**

| | | First implementation | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | MA_CMA | MA_CMR | MA_CR | MR_CMA | MR_CMR | MR_CR | Total |
| New implementation | MA_CMA | 9 723 | 505 | 46 | 408 | 703 | 24 | 11 409 |
| | MA_CMR | 871 | 451 | 9 | 134 | 1 073 | 5 | 2 543 |
| | MA_CR | 2 437 | 1 453 | 1 788 | 109 | 757 | 750 | 7 294 |
| | MR_CMA | 204 | 92 | 5 | 290 | 798 | 15 | 1 404 |
| | MR_CMR | 139 | 235 | 0 | 199 | 24 117 | 175 | 24 865 |
| | MR_CR | 58 | 191 | 140 | 380 | 16 188 | 65 128 | 82 085 |
| | Total | 13 432 | 2 927 | 1 988 | 1 520 | 43 636 | 66 097 | |

First and foremost, the most impacted classification case is MA_CR, the number of chips of which has been multiplied by 3. This classification case is really important in the scope of the end of the study.

The majority of chips classified as MA_CMR by the first run are classified as MA_CR by the run after improvement of the scenarios highlighting that the matching in scenario 2 of the first implementation was not correct for these chips because of the high amount of clouds near the coastlines (cf. Figure 30). This phenomenon is confirmed by the high number of chips classified as MR_CR by the implementation after improvement of the scenarios whereas these chips were classified as MR_CMR by the first implementation. Reciprocally the great majority of chips classified as MA_CR (MR_CR respectively) by the first implementation are still classified as MA_CR (MR_CR respectively) by the implementation after improvement of the scenarios signifying that if there is a high amount of clouds in the chip, often there is a high amount of clouds near the coastlines too. **This validates the new cloud acceptance criterion in scenario 2 rejecting chips only if the cloud amount is high near the coastlines**.

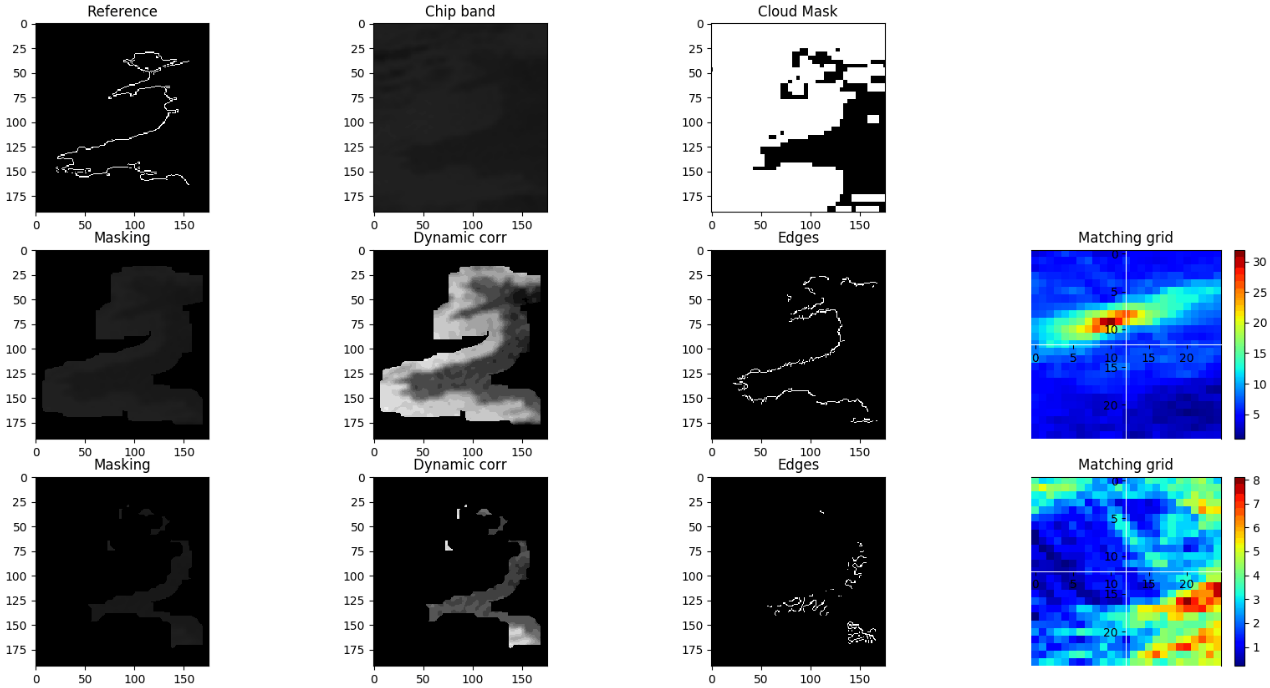



**Figure 30: Matching result after improvement of the scenarios for the band 3 of chip 45 of image MSG2-SEVI-MSG15-0100-NA-20100101011242**

Because of this new implementation of the cloud acceptance criterion, a lot of chips classified as MA_CMA by the first implementation are classified as MA_CR by the new implementation. Indeed, these chips are contaminated by clouds (but less than 50% of the chips is covered by clouds because these chips were not rejected by the cloud acceptance criterion of scenario 2 in the first implementation) the majority of which are located near the coastlines and covering more than 50% of the area near the coastlines. However, this impacts the results of scenario 2 only and does not penalize scenario 1.

Moreover, half of the chips classified as MR_CMA by the first run are classified as MA_CMA by the run after improvement of the scenarios (cf. Figure 31 in comparison with Figure 29). **This validates the use of a coastline mask** such as clouds far from the coastlines do not contaminate the matching results. Then the number of chips in the classification case MR_CMA has been a little bit lowered, reducing the positive impact that offers the cloud mask (by masking edges contaminating the matching in scenario 1) in this classification case and leading to an **improvement of scenario 1**.

Because of this new implementation of the coastline mask, the classification case MA_CR of the new implementation is the second one that receives most chips classified as MR_CR by the first implementation.

Then, because of the combination of the implementation of the coastline mask and the new cloud acceptance criterion the classification case MA_CR receives a lot of chips classified as MR_CMR by the first implementation (from MR to MA thanks to the coastline mask and from CMR to MR because of the new cloud acceptance criterion).

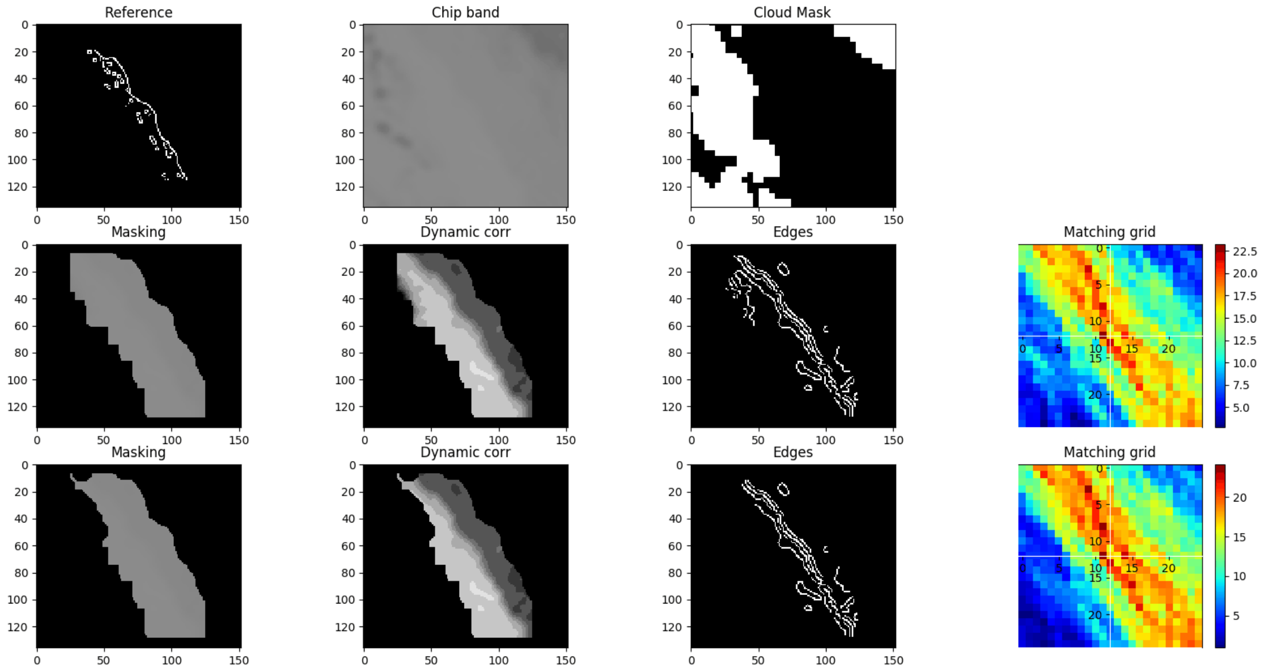



**Figure 31: Matching result after improvement of the scenarios for the band 8 of chip 5 of image MSG2-SEVI-MSG15-0100-NA-20100105051241**

Globally, the up left 3 by 3 square in Table 8 highlights that the majority of the accepted chips by scenario 1 of the first implementation are still accepted by scenario 1 of the second implementation as illustrated also in Table 9 (dedicated to the results of scenario 1 only). Indeed 17 283 chips among the 18 347 accepted chips by scenario 1 of the first run (see Table 5) are still accepted by scenario 1 of the run after improvement of the scenarios. Only 1 064 chips that were accepted by scenario 1 of the first run are rejected by scenario 1 of the run after improvement of the scenarios whereas 3 963 chips rejected by scenario 1 of the first run are accepted by scenario 1 of the run after improvement of the scenarios.

**Table 9: Repartition of the chips accepted and rejected by scenario 1 regarding the run after improvement of the scenarios and the run after improvement of the scenarios**

| | | First implementation | |
|---|---|---|---|
| | | Chip accepted | Chip rejected |
| **New implementation** | Chip accepted | 17 283 | 3 963 |
| | Chip rejected | 1 064 | 107 290 |

However the average distance to the separation hyper plane of the new accepted chips, reflecting the chip classification reliability, should be assessed in order to validate completely the improvement of scenario 1, and to a lesser extent of scenario 2. Table 10 details the average distance to the separation hyper plan for the different classification cases after improvement of the scenarios.

The average distances to the separation hyper plane in yellow in the first row of the table (corresponding to the average distance to the separation hyper plane of the accepted chips by scenario 1) have been increased in comparison with the first implementation (cf. Table 6) except for the one of the classification case MA_CMR reduced from 6.3% to 6.0%. Moreover, the low absolute values of the average distances before the improvement of the scenarios have been increased after the improvement of the both scenarios which highlights a more reliable classification with the new implementation than with the first one.

So **the improvement of the two scenarios leads to a more reliable classification of the chips, validating these improvements.**

**Table 10: Mean distance of the chips to the separation hyper plane for the different classification cases after improvement of the scenarios**

| | | Scenario 2 (with cloud information) | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| Scenario 1 (without cloud information) | Chip accepted | 19.3%<br>21.6% | -7.2%<br>7.3% | -5.0%<br>14.3% |
| | Chip rejected | 6.0%<br>-9.7% | -14.3%<br>-15.2% | -15.1%<br>-18.0% |

Moreover, omitting cloud information seems to give better results with this new implementation than with the first one, which will be verified in the following section.

### 4.3.2. Impact of clouds with the new implementation

#### 4.3.2.1. Impact of clouds on the chip acceptance

Table 11 presents the classification of the chip acceptance by the two scenarios after improvement of the scenarios.

**Table 11: Number of chips in the different classification cases after improvement of the scenarios**

| | | Scenario with cloud information | | | | |
|---|---|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds | | |
| Scenario without cloud information | Chip accepted | 11 409 | 2 543 | 7 294 | 21 246 | Total |
| | Chip rejected | 1 404 | 24 865 | 82 085 | 108 354 | |
| | | 12 813 | 27 408 | 89 379 | | |
| | | Total | | | | |

Table 12 shows the evolution of the number of chips in the different classification cases between the run before the improvement of the scenarios and the run with the same input images after the improvement.

The number of accepted chips by scenario1 has increased whereas the number of accepted chips by scenario 2 has decreased.

More precisely, as detected previously, the number of rejected chips after matching by scenario 2 has decreased whereas the number of chips rejected because of a high amount of clouds near the coastlines has increased, in comparison with the first run. This phenomenon highlights that when the coastline is cloud free, the matching process works correctly.

Moreover, the number of chips accepted by scenario 1 and rejected by the cloud acceptance criterion of scenario 2 have been multiplied by around 3, and as written previously the classification case MA_CR is a positive point in accordance to the objective of omitting cloud information. So the more chips in this classification case, the more beneficial the omission of cloud information will be. Indeed, regarding all the cases where chips are accepted by at least one scenario, it is the classification case MA_CR that implies the difference between the total numbers of accepted chips between the two scenarios considering this new implementation.

So **omitting cloud information could be more positive after the improvement of the two scenarios, especially for scenario 1**. But it should be verified through the analysis of the average distance to the separation hyper plane of the chips in the different classification cases.

**Table 12: Evolution of the different classification cases between before and after the improvement of the scenarios (percentage regarding the total number of chips)**

| | | Scenario with cloud information | | | | |
|---|---|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds | | |
| **Scenario without cloud information** | Chip accepted | -1,58% | -0,30% | 4,15% | **2,27%** | **Total** |
| | Chip rejected | -0,09% | -14,69% | 12,51% | -2,27% | |
| | | **-1,67%** | -14,99% | 16,66% | | |
| | | **Total** | | | | |

In the previous section it has been shown that the low absolute value of the average distances of the chips to the hyper plane in the different classification cases has been increased after the improvement of the scenarios (cf. Table 6 and Table 10) that highlights a more reliable classification with the new implementation.

Moreover, the analysis performed for the run before the improvement of the scenario is still the same for these new results that confirms that **omitting cloud information is more positive after the improvement of the two scenarios considering the chip acceptance reliability.**

So it has been demonstrated that the reliability of the chip classification for both scenarios is correct in the new implementation and better than the chip classification with the first implementation. Then the study of the impact of the percentage of clouds near the coastlines on the chip acceptance process is updated. Table 13 presents the mean cloud coverage near the coastlines for the different classification cases after improvement of the scenarios.

**Table 13: Mean cloud cover near the coastlines for the different classification cases after improvement of the scenarios**

| | | Scenario 2 (with cloud information) | | |
|---|---|---|---|---|
| | | Chip accepted | Chip rejected after matching | Chip rejected because of clouds |
| **Scenario 1 (without cloud information)** | Chip accepted | 17.9% | 26.3% | 72.9.6% |
| | Chip rejected | 28.5% | 23.4% | 88.8% |

First of all, for the new implementation, the mean cloud cover near the coastlines in the different classification cases is higher than the mean cloud cover in the whole chips for the same classification cases with the new implementation. This points out that the clouds are not equally distributed in the chips and are located near the coastlines principally, that is not facilitating the matching process.

Secondly the mean cloud cover near the coastlines in the different classification cases of the new implementation is higher than the mean cloud cover in the whole chips for the same classification cases in the first implementation. This means that the new implementation accepts more clouds in the interest area of the chips for the matching compared to the first implementation, whereas the classification is more reliable. So **the new implementation is less sensitive to the presence of clouds in the chips**.

Finally the numerous chips classified as MA_CR by the new implementation are in average contaminated by a lot of clouds near the coastline (71.9%) whereas the reliability of the acceptance by scenario 1 is good (the average distance to the separation hyper plane is equal to 13.6%) and even better than the reliability of acceptance by scenario 1 of the chips classified as MA_CMR. It implies two possibilities:

- The cloud mask is not correct for the chips classified as MA_CR;

- Or the clouds are transparent for the dedicated chip bands classified as MA_CR and the coastlines under the clouds can be detected. Table 14 shows that chip bands at the different wavelengths are classified as MA_CR so the non-detection of the clouds by scenario 1 is not related to a specific wavelength (please refer to Table 1 for the correspondence between the SEVIRI band number and the wavelengths).

**Table 14: Distribution over the different chip bands classified as MA_CR**

| Band number | 0 | 1 | 2 | 3 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Number of chips | 515 | 1177 | 1213 | 1366 | 825 | 641 | 801 | 623 | 133 |

**Thus the improvement of the scenarios is beneficial considering the chip acceptance process for scenario 1** as more chips are accepted by scenario 1 after the improvement, and with a higher reliability (*i.e.* average distance to the separation hyper plane)**.** In the other and, **the analysis of the components of the GQA quality does not enable to conclude if the GQA quality computed by scenario 2 after the improvement of the scenarios will be better than the GQA quality computed by the same scenario before the improvement**. Indeed for scenario 2, less chips are accepted, but with a higher reliability. It will be assessed in the next section.

Anyway, after the improvement, as more chips are accepted by scenario 1 and with a higher reliability than by scenario 2**, the quality of the GQA estimated by scenario 1 should be better than the one estimated by scenario 2**.

#### 4.3.2.2. Impact of clouds on the GQA quality

GQA results and GQA quality for the same SEVIRI images as those used in section 3.3 have been processed again with the new implementation of the scenarios.

Among the 648 (72 SEVIRI images, 9 bands per image) GQA and GQA quality values computed by each one the two scenarios, 115 are not studied because no chips have been accepted by one of the two scenarios, or by both of them. Among the remaining SEVIRI image bands, 334 GQA estimations computed by scenario 1 are higher than the corresponding GQA estimations computed by scenario 2 (against 292 in the first implementation) and 199 GQA estimations computed by scenario 2 are higher than the corresponding GQA results computed by scenario 1 (against 252 in the first implementation). Then the mean value between the GQA results computed by scenario 1 and scenario 2 is equal to 0.041 low resolution SEVIRI pixel (against -0.009 in the first implementation). Thus, on average, **the GQA results obtained by each of the two scenarios stay really close to each other, but GQA estimated by scenario 1 is in average higher than the GQA estimated by scenario 2**.

Moreover, as is described in Table 15, 514 GQA quality values computed by scenario 1 are lower than the corresponding GQA quality values computed by scenario 2 (against 524 in the first implementation) and 24 GQA quality values computed by scenario 2 are lower than the corresponding GQA quality values computed by scenario 1 (against 20 in the first implementation).

On average in case the GQA quality computed by scenario 2 is lower (better) than the GQA quality computed by scenario 1 (second row in Table 15):

- The GQA quality values given by the two scenarios are higher (worse) than in case the GQA quality computed by scenario 1 provide the most reliable GQA estimation (the best GQA quality);
- The GQA estimation provided by scenario 1 is closer to the best estimation than the GQA estimation given by scenario 2 in case scenario 1 provides the most reliable GQA estimation, regarding the average differences between the GQA quality values.

**Table 15: Analysis of GQA quality of the first implementation**

| | Number of cases | Mean GQA quality 1 | Mean GQA quality 2 | Mean difference between the GQA quality values |
|---|---|---|---|---|
| Scenario 1 is the best | 514 | 0.3202 | 0.3635 | 0.043 |
| Scenario 2 is the best | 24 | 0.4500 | 0.4295 | 0.020 |

Moreover, Table 16 details the evolutions of the mean GQA quality in case scenario 1 provides the best GQA estimation and in case scenario 2 provides the best GQA estimation, between the first and the new implementations:

- The mean GQA quality values given by scenario 1 and scenario 2 in both cases (scenario 1 is the best, and scenario 2 is the best) have been reduced (excepted the mean GQA quality value computed by scenario 2 when scenario 1 is the best, that is quite the same), implying that the new implementation provides more reliable GQA estimation in both cases;
- The mean GQA quality values given by scenario 1 and scenario 2 in case scenario 2 provides better GQA estimation have been more improved (but are still higher) than in case scenario 1 provides better GQA estimation.

**Table 16: Differences of mean GQA quality between the first and the new implementations**

| | First implementation | | New implementation | | Differences | |
|---|---|---|---|---|---|---|
| | Mean GQA quality 1 | Mean GQA quality 2 | Mean GQA quality 1 | Mean GQA quality 2 | Mean GQA quality 1 | Mean GQA quality 2 |
| **Scenario 1 is the best** | 0.3315 | 0.3629 | 0.3202 | 0.3635 | -0.0011 | 0.0006 |
| **Scenario 2 is the best** | 0.5031 | 0.4836 | 0.4500 | 0.4295 | -0.0531 | -0.0541 |

**Thus the improvement of the two scenarios induces more reliable GQA estimation. Globally, if scenario 1 provided better GQA estimations than scenario 2 with the first implementation, it is even better in the new implementation.**

#### 4.3.2.3. Impact of the cloud cover threshold

In order to assess the impact of the cloud cover threshold in Scenario 2, a sensitive study has been perform on 12 SEVIRI images acquired at different season and night/day time, by using the last version of the code.

Figure 32 shows the impact of the cloud cover threshold on the number of chips rejected by the cloud acceptance criterion. The chips with a percentage of cloudy pixel in the coastline mask higher than the threshold are rejected whereas the chips with a percentage of cloudy pixel in the coastline mask lower than the threshold are accepted by the cloud acceptance criterion. As expected, the higher the threshold, the less the number of rejected chips. However, it could be noted that more than 30% of the image chips have 95% of cloudy pixels in the coastline mask (near the coastlines) at least.

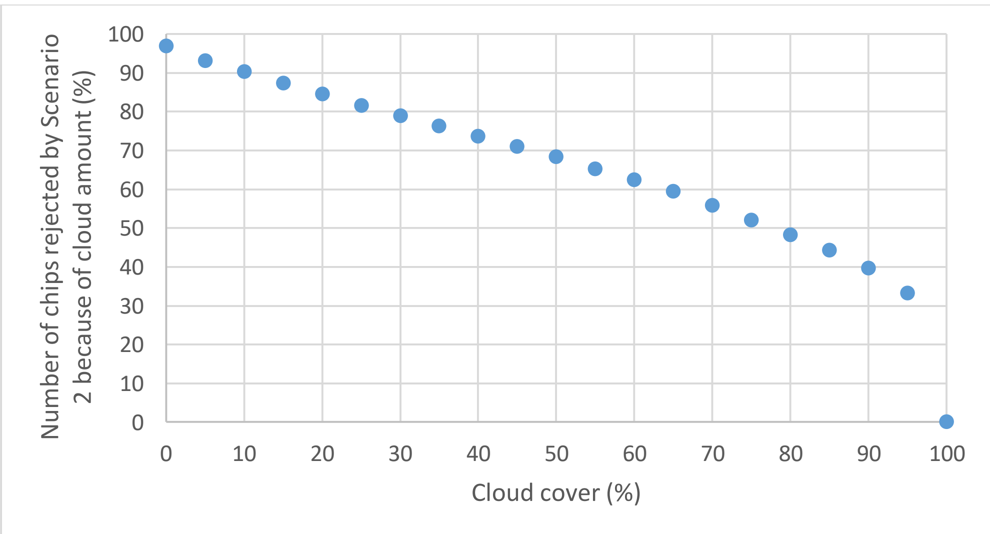



**Figure 32: Impact of the cloud cover threshold on the number of rejected chips by the cloud acceptance criterion of Scenario 2**

Figure 33 shows the impact of the cloud cover threshold on the number of accepted and rejected image chips by the two scenarios. Of course, the number of rejected and accepted image chips by Scenario 1 is independent of the cloud cover threshold.

For scenario 2, all the image chips are rejected if the cloud cover threshold is equal to 0% (by the cloud acceptance criterion), and the number of rejected image chips reduces and tends to the number of rejected image chips obtained with Scenario 1 when the cloud cover threshold tends to 100% (all the image chips are accepted by the cloud acceptance criterion).

However, even if all the image chips are accepted by the cloud acceptance criterion (cloud cover threshold equal to 100%), the cloud mask is still applied on the cloudy pixel before the matching. And, as the number of rejected image chips by Scenario 2 when all the image chips are accepted by the cloud acceptance criterion, is not smaller than the number of rejected image chips by Scenario 1, **the cloud mask has no a positive impact on the chip acceptance process**.

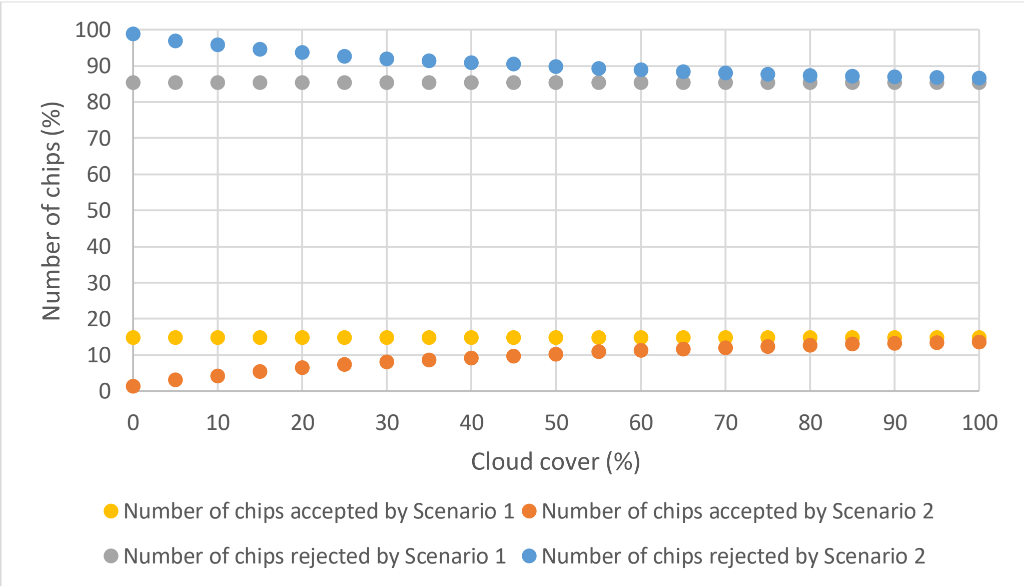



**Figure 33: Impact of the cloud cover threshold on the chips acceptance**

Figure 34 shows the impact of the cloud cover threshold on the GQA quality. Of course, the GQA quality obtained by Scenario 1 is independent of the cloud cover threshold. However, for Scenario 2, the GQA quality is very bad when the cloud cover threshold is low, then the GQA quality decreases while the cloud cover threshold increases, until to reach a minimum value for a cloud cover threshold equal to 30% before to increases and to tend to a constant value higher than the GQA quality obtained with Scenario 1.

This highlights that the choice of a cloud cover threshold equal to 50% for the study has not a crucial impact on the global result, namely that Scenario 1 provides better results than Scenario 2. However, **Scenario 2 seems to be optimized for a cloud cover threshold equal to 30%**.

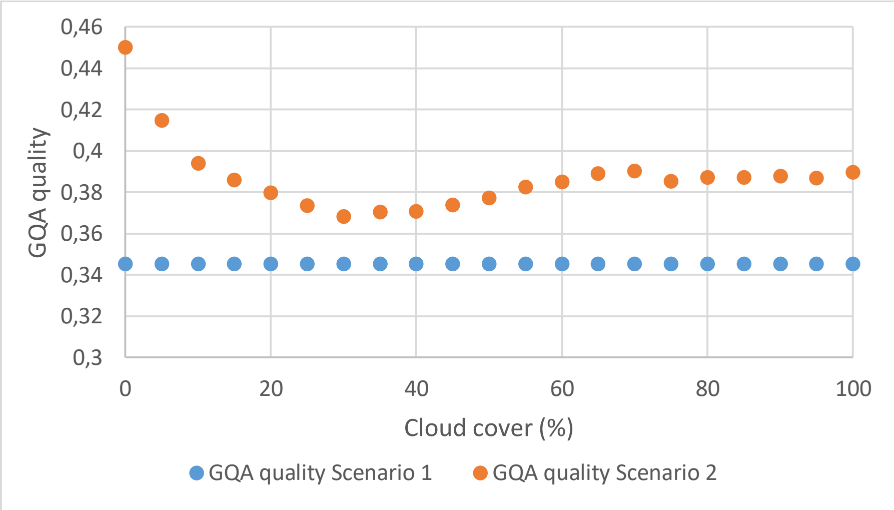



**Figure 34: Impact of the cloud cover threshold on the GQA quality (mean GQA quality of the processed SEVIRI images)**

# 5. Conclusion

Precise geolocation of Earth Observation satellite measurements in Earth coordinates is critical for users exploiting L1 and L2 satellite products. In theory, determining on-ground pixel location is possible from the information on the exact orbit position, attitude, and position of optical elements such as scanning mirrors. In practice, however, various errors, such as deviations from prescribed orbit, attitude variations, and uncertainty on the precise position of optical instrument components make this task prone to errors that are not acceptable for end-users. For this reason, in this study known fixed landmarks on the Erath are used to **estimate the GQA of the satellite images** thanks to a process including edge extraction and matching with the known landmarks of chips of the whole images. However the performance of the GQA estimation process is greatly influenced by **cloud contamination** in the images. As it is not possible to be sure to know the positions of each cloudy pixel of an image and to work only on cloud free pixels the study “Impact of undetected clouds on Image Navigation Quality Assessment” [AD1] aims at **understanding and at quantifying the influence of the impact of clouds on the GQA estimation process and at evaluating whether it is possible to omit cloud information without degrading the GQA quality**.

The study is based on two scenarios applied to the same input satellite (SEVIRI in this study) images: in the first one GQA result is computed without using cloud information and in the second one, GQA result is computed by using cloud information. First and foremost the GQA estimation process thanks to each of these two scenario has been implemented (cf. section 2).

Then a mean to **assess the quality of the estimated GQA** to define **which one of the two scenarios provides the more precise GQA estimation** for the same input image has been investigated. A mathematical formulation of the GQA quality has been proposed and its relevance has been demonstrated (cf. section 3). By lack of information, the GQA quality formulation reflects the precision of the GQA only and not its accuracy.

However, thanks to this formulation, GQA results provided by each of the two scenarios have been compared and the scenario without cloud information gives more reliable GQA estimations than the scenario with cloud information (cf. section 3.3). This result was not expected and the image chips used to estimate the GQA results have been analysed.

The analysis showed that **not using cloud information increases the chip acceptance reliability** used in the GQA estimation process and **enables to accept with good reliability some chips with a high amount of clouds** that are rejected by the cloud acceptance criterion of scenario with cloud information (cf. section 4.1). Then the both scenarios have been improved (cf. section 4.2) and the new analysis of the chips used for the GQA estimation by each of the two scenarios confirms that **omitting cloud information does not degrade the GQA quality**, actually, **the chip acceptance process is improved, so the GQA quality is improved too**. This result is mainly explained by the fact that binary cloud information has been used for this study: clouds or no clouds. By using more complex cloud information, for example related to the altitude and the nature of the clouds, the same study could lead to a different conclusion.

# Appendix A Variation of the matching method

Another method has been investigated based on the above principle. It consists to define matching patches smaller than the size of the whole reference image. For each pixel of the edges of the reference image, the corresponding pixel in the edges of the image chip band is searched by computing the cross correlation between two matching windows: one anchored in the reference image (the central pixel of which is the pixel we want to retrieve in the image chip band) and the second one moving in an anchored search window, the central pixel of which has the same location as the central pixel of the anchored matching window in the reference image). Then a matching grid as presented above is obtained for each pixel belonging to the edges of the reference image as is shown in Figure 35. In this example, the size of the matching window is (3, 3) and the maximum search direction is equal to one pixel in each direction, so the resulting matching grid size is (3, 3).

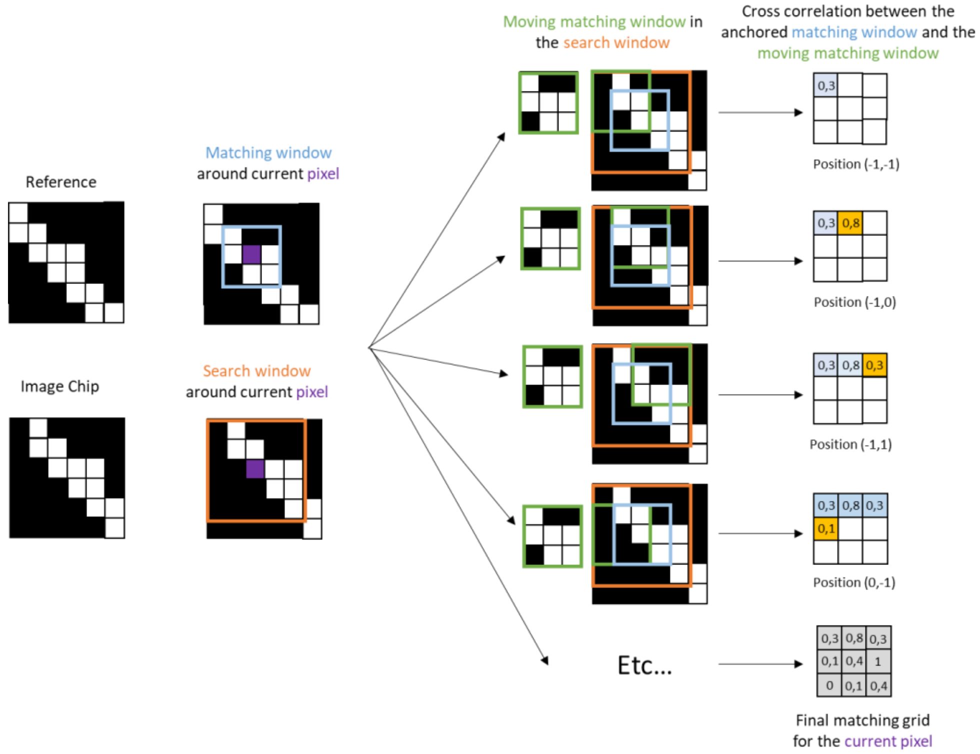



**Figure 35: Matching process with moving matching windows**

All these matching grids (one for each pixel belonging to the edges of the reference image) are summed to obtain the global matching grid for the whole image chip band.

This method allows to not take into account the clouds far from the edges of the reference image in the matching score computation. However it has been abandoned because if a number of consecutive pixels of the edges of the image reference higher than the size of the matching window are aligned in the same direction, the matching result will be high and the same for several location of the moving matching window. So the offset between the image chip band and the reference image will not be find with precision. Of course this phenomenon could occur with the retained approach but globally the selected landmarks have edges in different directions.

# Appendix B Coefficients not used in the GQA quality formulation

As mentioned in section 3.2.4, two identified coefficients have not been used for the GQA quality formulation. These coefficients are presented below.

A coefficient that could give information on the reliability of the GQA estimation on the whole SEVIRI image band is the distribution of the accepted chips. Indeed, it is assumed that the better the chips are distributed in the SEVIRI grid, the better the GQA quality on the whole image, as is illustrated in Figure 36. In the two cases, as much chips have been accepted. However, the GQA estimates for the right case is more representative of the GQA of the whole SEVIRI image band than the GQA estimates for the left cases because all the chips used for this last GQA estimation are close to each other and should give GQA information over South America only.

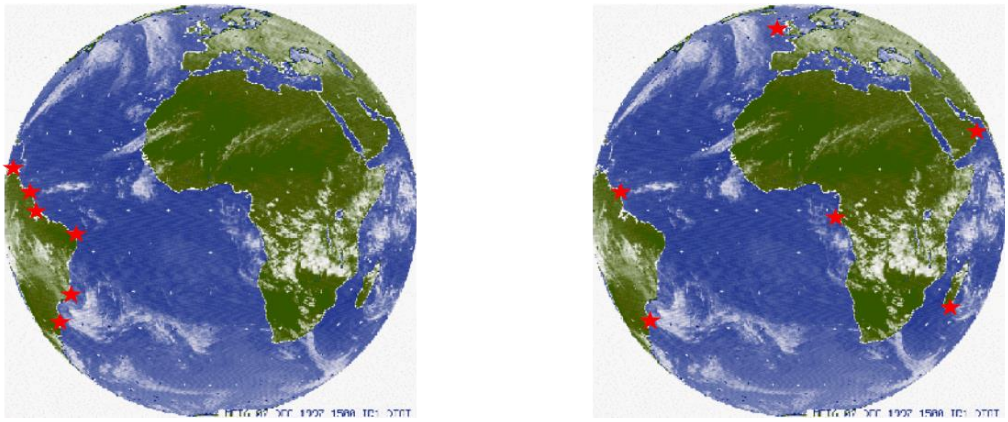


**Figure 36: Examples of different accepted chip repartition over the SEVIRI grid**

As the distribution of the accepted chips is constrained by the distribution of the 200 reference landmarks used for the study, this coefficient has been computed as the sum of the Euclidian distance between all the accepted chips, normalized by the sum of the Euclidian distance of a same number of chips giving the best distribution. Thus this normalized distribution does not depend on the location of the initial reference landmark and on the number of accepted chips. The complexity of the computation of the normalization coefficient for each number of accepted chips (from 2 to 200) is high, but these coefficients are computed once and stored in an ancillary file, so they are accessible easily.

How the normalization coefficient for 3 accepted chips is computed is presented in Figure 37.. On the left, there is an illustration of the SEVIRI grid containing 5 image chips (against 200 in reality). Then all the possible configurations of 3 chips among the 5 of the SEVIRI grid are presented in the 10 pictures on the right. For each configuration, a distance d is computed, as the sum of the distance between all the chips pairs. The best distribution is obtained when the greater distance is found. In this case, for 3 accepted chips, the last configuration gives the highest distance, so the best distribution of 3 chips on the SEVIRI grid, and the normalization coefficient for this example is equal to 13.

However, the GQA standard deviation varies randomly according to the distribution coefficient. Indeed, the distribution of the accepted chips does not affect the GQA values directly. As the impact of the distribution of the accepted chips on the GQA estimation cannot be mathematically demonstrate, it will not be used in the mathematical formulation of the GQA quality.

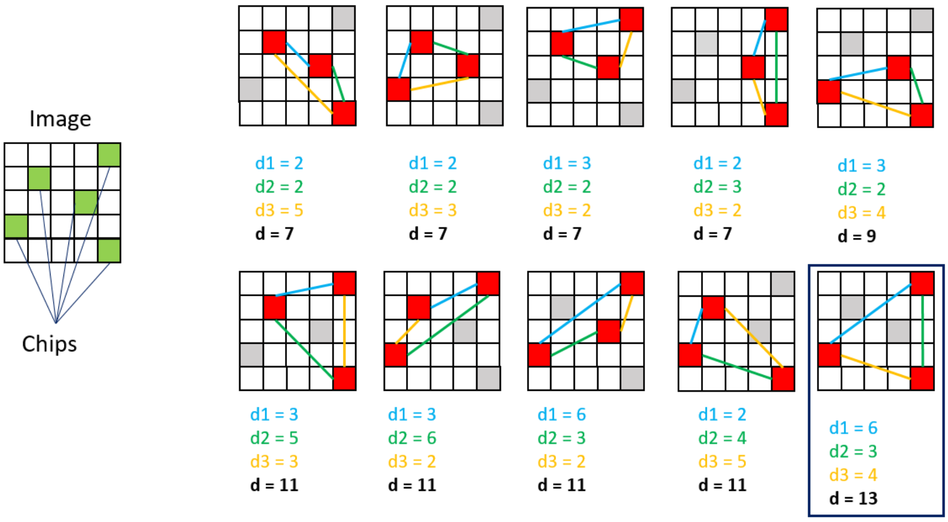



**Figure 37: Chips distribution evaluation**

The last investigated coefficients that could give information on the reliability of the GQA estimation on the whole SEVIRI image band are the normalized standard deviation of the offsets of the accepted chips in South/North and East/West directions. Indeed, if the offsets in a given direction for all the accepted chips are nearly the same, implying that the standard deviation of these offsets tends to 0, the SEVIRI image band is just translated globally from the mean offset value. And, its GQA quality should be better than a SEVIRI image band the accepted chips of which give different offset values, implying that the standard deviation of these offsets is higher than the previous one and that the SEVIRI image band has undergone more complex transformations than a simple translation.

However, even if the mean GQA decreases while the standard deviation of the offsets of the accepted chips in South/North and East/West directions decreases, the GQA standard deviation does not seem to be related to the standard deviation of the offsets.

Maybe this result could be improved if the offsets obtained for the accepted chips could take more different values. Indeed, by oversampling the image chips by a factor 4, the obtained offset are discretized by 0.25 pixels.

Nevertheless, these coefficients will not be used in the mathematical formulation of the GQA quality.