EUMETSAT NWC SAF

SUPPORT TO NOWCASTING AND VERY SHORT RANGE FORECASTING

# Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO

NWC/CDOP3/GEO/AEMET/SCI/ATBD/Precipitation, Issue1 , Rev. 0.1

Date 29th October 2021

*Applicable to*

*GEO-PC-v154 (NWC-022)*
*GEO-CRR-v402 (NWC-026)*
*GEO-PCPh-v30 (NWC-027)*
*GEO-CRRPh-v30 (NWC-084)*

**Prepared by AEMET**

## REPORT SIGNATURE TABLE

| Function | Name | Signature | Date |
|---|---|---|---|
| **Prepared by** | José Alberto Lahuerta (AEMET) | | *29th October 2021* |
| **Reviewed by** | Llorenç Lliso and Pilar Ripodas (GEO Managers) | | *29th October 2021* |
| **Authorised by** | Pilar Ripodas Agudo<br>NWC SAF Project Manager | | *29th October 2021* |

## DOCUMENT CHANGE RECORD

This is the first version of the document, which is based on the Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO

Changes describe below refer to this document.

| Version | **Date** | Pages | CHANGE(S) |
|---|---|---|---|
| 1.0 | *1st September 2021* | 83 | CRRPh and PCPh are new algorithms, both based on a Principal Component Analysis. PC and CRR keep the same from the last version of the ATBD NWC/CDOP2/GEO/AEMET/SCI/ATBD/Precipitation, Issue 2, Rev.1, 21 January 2019. |
| 1.0.1 | *29th October 2021* | 83 | As outcome of the RR/PCR review, It has been included figures : 22,23,24,25,26,27,28,37,38,39,40 and some clarifications in sections 4 ,5 and 6 |

# Table of contents

List of Tables and Figures

# 1. INTRODUCTION

The Eumetsat "Satellite Application Facilities" (SAF) are dedicated centres of excellence for processing satellite data, and form an integral part of the distributed EUMETSAT Application Ground Segment (http://www.eumetsat.int). This documentation is provided by the SAF on Support to Nowcasting and Very Short Range Forecasting, NWC SAF. The main objective of NWC SAF is to provide, further develop and maintain software packages to be used for Nowcasting applications of operational meteorological satellite data by National Meteorological Services. More information can be found at the NWC SAF webpage, http://nwc-saf.eumetsat.int. This document is applicable to the NWC SAF processing package for geostationary meteorological satellites, NWC/GEO.

## 1.1 Scope of the document

This document is the Algorithm Theoretical Basis Document for the precipitation products Precipitating Clouds (PC), Convective Rainfall Rate (CRR) and Precipitation products from Cloud Physical Properties (PPh) of the NWC/GEO software package. PPh generates two different products: Precipitating Clouds from Cloud Physical Properties (PCPh) and Convective Rainfall rate from Cloud Physical Properties (CRRPh).

The Algorithm Theoretical Basis Document describes the physics of the problem together with the mathematical description of the algorithm. It also provides information on the objectives, the needed input data and the outputs of the products.

## 1.2 Software Version Identification

This document describes the algorithms implemented in the 2021 NWC-GEO software package release (GEO-PC v1.5.4, GEO-CRR v4.0.2, GEO-PCPh v3.0 and GEO-CRRPh v3.0 ).

## 1.3 Improvement from previous version

2021 precipitation products version includes these technical improvements:

- New CRRPh and PCPh algorithms based on a Principal Component Analysis. There is only one algorithm for each CRRPh and PCPh that includes both day and night conditions.
- Microphysical properties are simulated at night time and used in the algorithm.
- More information is extracted from the SEVIRI channels.
- CRRPh incorporates a Cloud Water Path enhancement correction factor along with a lightning module.
- Adaptation to Himawari9 and GOES17. This adaptation is purely technical in order to use Himawari9 and GOES17 channels, but no objective validation has been performed for these satellites.

Note:

PC and CRR keep the same from previous 2018.1 version

## 1.4 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| AEMET | Agencia Estatal de Meteorología |
| ATBD | Algorithm Theoretical Basis Document |
| BALTRAD | Baltic Radar Network |
| CAPPI | Constant Altitude Plan Position Indicator |
| CMIC | Cloud Microphysics |
| COT | Cloud Optical Thickness |
| CRRPh | Convective Rainfall Rate from Cloud Physical Properties |
| CRR | Convective Rainfall Rate |
| CSI | Critical Success Index |
| CT | Cloud Type |
| CWP | Cloud Water Path |
| ESSL | European Severe Storm Laboratory |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| FAR | False Alarm Ratio |
| HRIT | High Rate Information Transmission |
| ICD | Interface Control Document |
| ICP | Illumination Conditions Parameter |
| IQF | Illumination Quality Flag |
| IR | Infrared |
| LI | Lightning Imager |
| MAE | Mean Absolute Error |
| ME | Mean Error |
| MRV | Maximum Reflectivity in the Vertical |
| MSG | Meteosat Second Generation |
| MTG | Meteosat Third Generation |
| NIR | Near Infrared |

| | |
|---|---|
| NWCLIB | Nowcasting SAF Library |
| NWC SAF | Satellite Application Facility for Nowcasting |
| PC | Precipitating Clouds |
| PCorr | Percentage of Corrects |
| PCA | Principal Component Analysis |
| PCPh | Precipitating Clouds from Cloud Physical Properties |
| PGE | Product Generation Element |
| POD | Probability of Detection |
| PoP | Probability of Precipitation |
| PPh | Precipitation from Cloud Physical Properties |
| PWRH | Moisture Correction Factor |
| $R_{eff}$ | Effective Radius |
| RLR | Rainfall-Lightning Ratio |
| RMSE | Root Mean Square Error |
| RR | Rain Rate |
| SAF | Satellite Application Facility |
| SEVIRI | Spinning Enhanced Visible and Infrared Imager |
| SW | Software |
| $T_{surf}$ | Surface Temperature |
| 2-V | 2-Variable |
| 3-V | 3-Variable |
| VIS | Visible |
| VIS-N | Normalized Visible |
| WV | Water Vapour |

## 1.5 REFERENCES

### 1.5.1 Applicable Documents

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD.X].

For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the current edition of the document referred applies.

Current documentation can be found at the NWC SAF Helpdesk web: https://www.nwcsaf.org.

| Reference | Title | Code | Vers | Date |
|---|---|---|---|---|
| [AD. 1] | Proposal for the Third Continuous Development and Operations Phase (CDOP-3) March 2017-February 2022 | NWC SAF: CDOP-3 proposal | 1.0 | 11/04/16 |
| [AD. 2] | NWCSAF Project Plan | NWC/CDOP3/SAF/AEMET/MGT/PP | 1.5 | 15/04/21 |
| [AD 3] | Configuration Management Plan for the NWC SAF | NWC/CDOP3/SAF/AEMET/MGT/CMP | 1.1 | 15/04/20 |
| [AD 4] | NWCSAF Product Requirements Document | NWC/CDOP3/SAF/AEMET/MGT/PRD | 1.3 | 29/05/20 |
| [AD 5] | NWCSAF CDOP-3 Service Specifications | NWC/CDOP3/SAF/AEMET/MGT/SSD | 1.2 | 20/03/19 |

*Table 1: List of Applicable Documents*

### 1.5.2 Reference Documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable ones, and can be looked up to enhance the information included in this document if it is desired. They are referenced in this document in the form [RD.X].

For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the current edition of the document referred applies.

Current documentation can be found at the NWC SAF Helpdesk web: http://www.nwcsaf.org

| Reference | Title | Code | Vers | Date |
|---|---|---|---|---|
| [RD 1] | Algorithm Theoretical Basis Document for SAFNWC/MSG “Precipitating Cloud” (PC-PGE04 v1.5) | SAF/NWC/CDOP2/SMHI/SCI/ATBD/4 | 1.5.4 | 15/07/13 |
| [RD 2] | Scientific and Validation Report for the Precipitation Product Processors of the NWC/GEO | NWC/CDOP3/GEO/AEMET/SCI/VR/Precipitation | 1.1 | 18/12/19 |
| [RD 3] | Data Output Format for the NWC/GEO | NWC/CDOP3/GEO/AEMET/SW/DOF | 1.1 | 01/10/19 |
| [RD 4] | Interface Control Document for Internal and External Interfaces of the NWC/GEO | NWC/CDOP3/GEO/AEMET/SW/ICD/1 | 1.1 | 01/10/19 |
| [RD 5] | User Manual for the Precipitation Product Processors of the NWC/GEO | NWC/CDOP3/GEO/AEMET/SCI/UM/Precipitation | 1.0 | 21/01/19 |
| [RD 6] | Algorithm Theoretical Basis Document for the Cloud Product Processors of the NWC/GEO | NWC/CDOP2/GEO/MFL/SCI/ATBD/Cloud | 2.1 | 21/01/19 |
| [RD 7] | Software User Manual of the Parallax Correction Processor of the NWC/GEO | NWC/CDOP2/GEO/AEMET/SW/SUM/PLAX | 1.0 | 18/05/15 |
| [RD 8] | Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO | NWC/CDOP2/GEO/AEMET/SCI/ATBD/Precipitation | 2.1 | 21/01/19 |

*Table 2: List of Referenced Documents*

# 2. DESCRIPTION OF PRECIPITATING CLOUDS (PC) PRODUCT

PC product keeps the same from the last "Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO" v2018 2.1 [RD 8]

Refer to Algorithm Theoretical Basis Document for SAFNWC/MSG "PrecipitatingCloud" (PC-PGE04 v1.5)[RD 1].

## 2.1 PRECIPITATING CLOUDS (PC) OVERVIEW

The relatively weak coupling between spectral features in the visible and infrared channels with precipitation rate for all situations except for strong convection makes it in most cases doubtful to try to assign precipitation rates from SEVIRI data alone. However it is possible to statistically determine the likelihood of precipitation from visible and infrared spectral signatures in a SEVIRI scene. The PC product for MSG is thus to be seen as a complement of the convective rain rate product, which specifically addresses convective situations, and the rapidly developing thunderstorm product, which also takes into account the time evolution of systems. The precipitating cloud product can serve as a general tool for Nowcasting of precipitation, especially for areas where no surface radar data is available. It should however be noted, that the nature of the input data usually leads to an overestimation of the precipitating area.

## 2.2 PRECIPITATING CLOUDS (PC) ALGORITHM DESCRIPTION

### 2.2.1 General algorithm design

The precipitating clouds product gives the likelihood of precipitation. Validation and prototyping for earlier software versions have shown that there is no skill in trying to stratify Total precipitation likelihood into light to moderate precipitation and strong precipitation. As a consequence only the total precipitation likelihood is now reported as class 1:

- Class 1: precipitation >0.1 mm/h
- Class2: obsolete, set to zero

A linear combination of those spectral features, which have the highest correlation with precipitation, is used to construct a Precipitation index PI. For each value of the PI, the probability of precipitation in the respective classes is then determined from a comprehensive dataset of co-located satellite data, precipitation rates from rain gauge measurements and surface temperatures from NWP.

In the calculation of the PI special attention has been given to spectral features in the visible, which implicitly contain information on cloud microphysical properties at the cloud top, such as effective radius and cloud phase. The algorithm employed is cloud type dependent in the sense that mapping from PI to precipitation likelihood makes use of cloud type dependent lookup tables. For the PI calculation a day and a night version exists, where the night version only makes use of IR channels not influenced by sunlight.

### 2.2.2 Data used for algorithm development and tuning

Tuning had to be performed over the central Europe. Since not enough systematic radar data sets were available for tuning of the SEVIRI algorithm to radar, SYNOP current weather reports (October2003 – August 2004) and French rain gauge data (January 2004 – December 2004) were used for tuning. The current default configuration for algorithm version 1.5 has been unchanged since version 1.3 (released spring 2007) and uses a cloud type dependent tuning based on French Gauge data. There is also the option to configure the algorithm for a previous version of tuning released with v1.2. This older tuning is independent of cloud type, and based on European SYNOP reports for current weather. It is however not recommended to change to this option because of inferior algorithm performance. Validation results for version 1.2 and the versions identical to the current default algorithm version are reported in the validation report Scientific and Validation Report for the Precipitation Product Processors of the NWC/GEO [RD 2].

A precipitation index PI is calculated using the same formula for all cloud types, but mapping to likelihood is performed cloud type dependent. Validation results are reported in Scientific and Validation Report for the Precipitation Product Processors of the NWC/GEO [RD 2].

| | Precipitation Frequency [%] | Precipitation Frequency [%] | Algorithm number used for potentially raining cloud types ( |
|---|---|---|---|
| Cloud type | French gauges | Hungarian gauges | |
| **1 –cloud free land** | **0,6** | **0,1** | |
| **2 –cloud free sea** | **0,1** | **-** | |
| **3 –snow/ice land** | **4,0** | **3,9** | |
| **4 –snow/ice sea** | **-** | **-** | |
| **5 –very low Cu** | **-** | **-** | |
| **6 –very low St** | **1,9** | **0,5** | |
| **7 –low Cu** | **-** | **-** | |
| **8 –low St** | **7,7** | **4,3** | **Not considered raining, might need adjustment** |
| **9 -medium Cu** | **-** | **-** | **Algorithm 1** |
| **10 -medium St** | **21,5** | **13,3** | **Algorithm 1** |
| **11 -high&opaque Cu** | **-** | **-** | **Algorithm 2** |
| **12 - high&opaque St** | **30,3** | **31,4** | **Algorithm 2** |
| **13 -v. high&op Cu** | **-** | **-** | **Algorithm 2** |
| **14 -v. high&op St** | **43,6** | **35,4** | **Algorithm 2** |
| **15 – thin Ci** | **1,4** | **0,9** | |

| | | | |
|---|---|---|---|
| **16 - mod. Thick Ci** | **3,6** | **2,4** | |
| **17 – thick Ci** | **13,1** | **11,0** | **Algorithm 3** |
| **18 –Ci above lower cloud** | **5,1** | **3,9** | **Algorithm 4** |
| **19[2)] Fractional cloud** | **1,2** | **0,5** | |

*Table 3: Total likelihood of precipitation for different cloud types as compared to collocated French rain gauge data and Hungarian gauge data for Jan-Dec 2004. Rain gauge data averaged over 30 minutes*

### 2.2.3 Algorithm details

It was investigated which spectral features of SEVIRI were most correlated with precipitation. The Precipitation Index PI is constructed as a linear combination of those spectral features which are most correlated with precipitation as to maximise the correlation of PI and precipitation.

We have chosen a Precipitation Index of the form:

PI=a0 +a1*Tsurf +a2*T108+a3*(T108-T120) +a4*abs(a5-R06/R16)

+a6*R06+ a7*R16+a8*T062+a9*T073+a10*T039) *Eq.1*

This formulation will allow to specify different day and night algorithms and to easily tune the algorithm by just providing different coefficient files, for example for different cloud types. In the current implementation however, algorithms for different cloud type groups are using the same set of coefficients, but a cloud type specific mapping of PI to precipitation likelihood.

Cloud type groups are defined as follows:

Algorithm 0: all cloudtypes. In version 1.2 PI coefficients from tuning to synop current weather observations are supplied, together with tables for matching PI to precipitation likelihood. Use of this algorithm is not recommended. Instead the cloud type dependent algorithms tuned on French gauge data and outlined below should be used, as supplied in the standard configuration since version 1.3.

Algorithms 1 to 4 are tuned on French gauge data (average over 30 minutes). Coefficient sets for these algorithms are identical since version 1.3, but mapping of the resulting PI to precipitation likelihood is cloud type dependent:

Algorithm 1: cloud type 9,10 = medium level opaque cloud

Algorithm 2: cloud types 11 – 14 = high and very high opaque cloud

Algorithm 3: cloud type 17 = thick cirrus

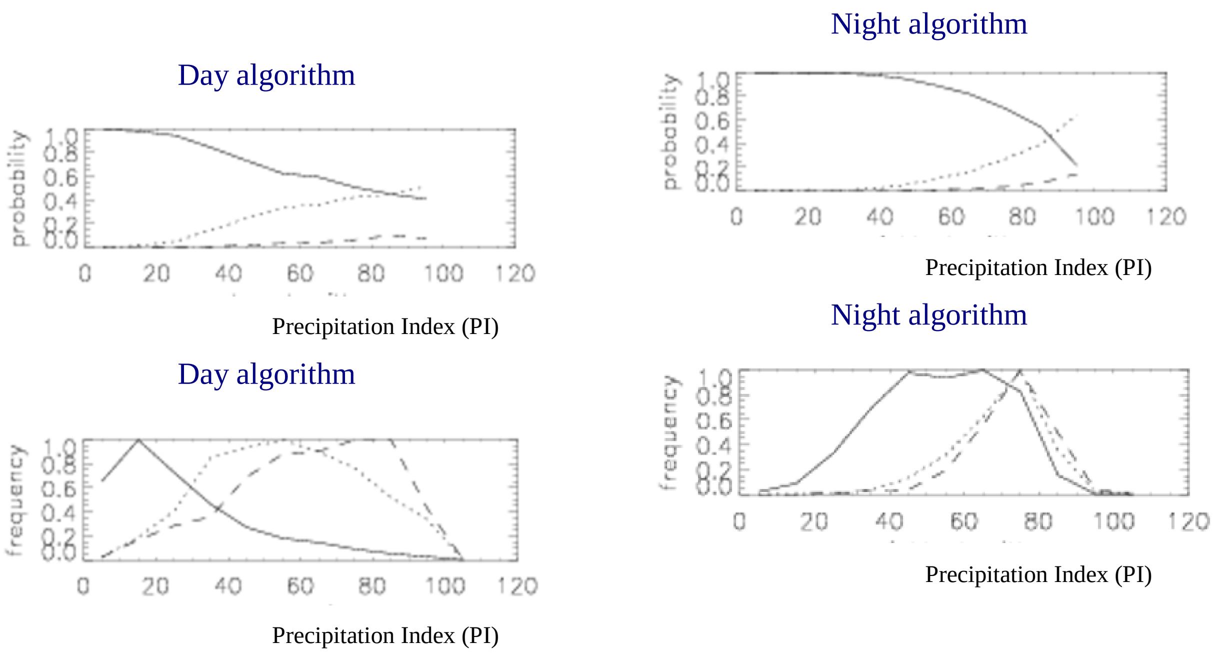
Algorithm 4: cloud type 18 = cirrus above lower clouds

Table 4 shows PC algorithm coefficients for day and night for all clouds, both for synop based tuning (Alg0), and for tuning to French gauge data which is currently identical to Algorithms 1 to 4.

| | a0 | a1 | a2 | a3 | a4 | a5 | a6 | a7 | a8 | a9 | a10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Alg0-day**<br>Tuned on SYNOP<br>(use not recommended) | 130.0 | -1.17841 | 0.193517 | 1.34862 | -0.403661 | 3.2 | 1.21913 | -1.14646 | 1.0137 | -0.729214 | 0.482047 |
| **Alg0-night**<br>Tuned on synop (use not recommended) | 130. 0 | -0.808931 | -0.660192 | -1.3209 | 0.0 | 0.0 | 0.0 | 0.0 | 1.56148 | -1.46149 | 0. |
| **Alg1,2,3,4-day**<br>Tuned on rain gauges (default) | **230.0** | **-1.35** | **-0.63** | **-2.59** | **63.79** | **0.0** | **-0.40** | **0.0** | **-0.92** | **0.32** | **1.52** |
| **Alg1,2,3,4-night**<br>Tuned on rain gauges(default) | **460.0** | **-0.90** | **-0.91** | **-5.34** | **0.0** | **0.0** | **0.0** | **0.0** | **-0.27** | **0.65** | **0.0** |

*Table 4: Coefficients a0 to a10 for current day and night algorithm according to tuning with current weather observations from synop (algorithm 0) and tuned on French gauge data (algorithms1 to 4)*

How the PI maps to probability for different intensity classes is illustrated in Figure 1. The normalised frequency distribution for different intensity classes as observed by gauge data is given in the lower panel. The total likelihood of precipitation is split into two intensity classes in these plots and would be represented by the sum of likelihood for light and heavy precipitation for each value of PI respectively. The likelihood that a certain value of the PI falls into a certain precipitation class is determined from the (not normalised) frequency distribution under the constraint that the total likelihood has to be 100% (upper panel). There seems to be no potential to differentiate intensity classes for the large majority of cases. A substantial overlap of all precipitating classes with the no-precipitation class is apparent in the normalized frequency distribution. This is especially true for the night algorithm. Generally better precipitation discrimination can be performed at day time since the daytime algorithm is strongly dependent on the R6/R16 feature, discontinuities between day and night algorithms could not be avoided. When deriving the probabilities that a given PI belongs to a certain precipitation class, the resulting distribution suffers from the fact that there is a wide overlap between the precipitating and non-precipitating classes, as well as from the generally much larger number of non-precipitating cases.




**DAY** **NIGHT**

*Figure 1: algorithm1-4 for all potentially precipitating cloud types tuned on French gauge data. Left: day algorithm, right: night algorithm. Lower panels: normalised histogram for different precipitation classes (solid line: no precipitation, dotted: light to moderate precipitation, dashed: heavy precipitation). Upper panels: same as lower, but for probability, total precipitation likelihood would be the sum of the dotted and dashed lines in the upper panel X-axis: Precipitation Index PI*

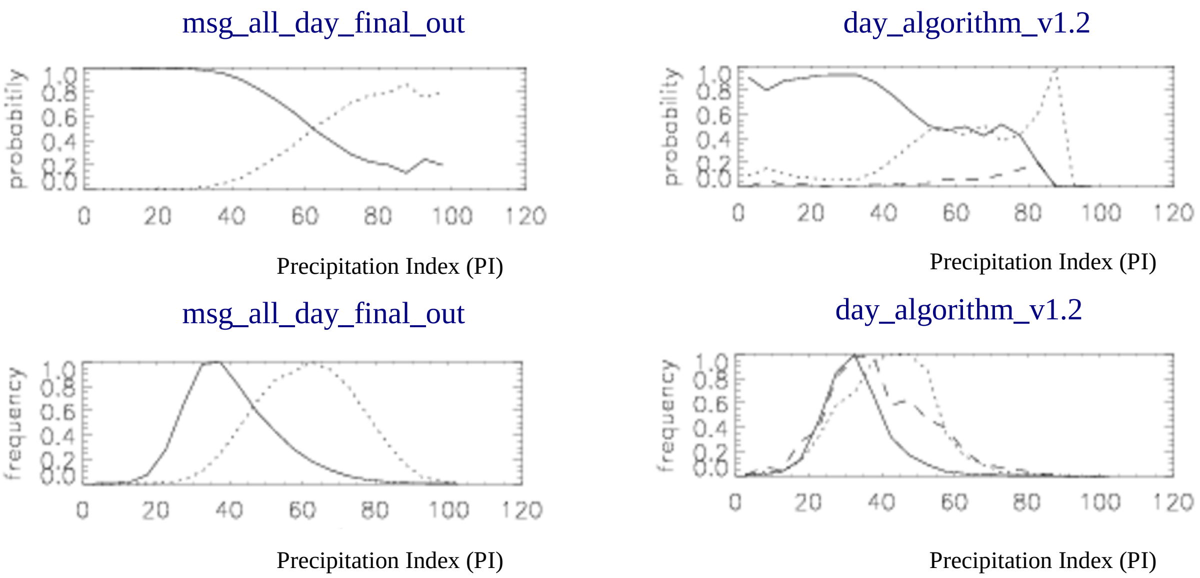


*Figure 2 : left: day algorithm tuned on synop collocation data set (dotted: all precipitation). Right: same algorithm, but applied on rain gauge collocation data set (dotted light to moderate precipitation, dashed heavy precipitation). X-axis: Precipitation Index PI*

An example of the precipitating clouds product is given in Figure 3.

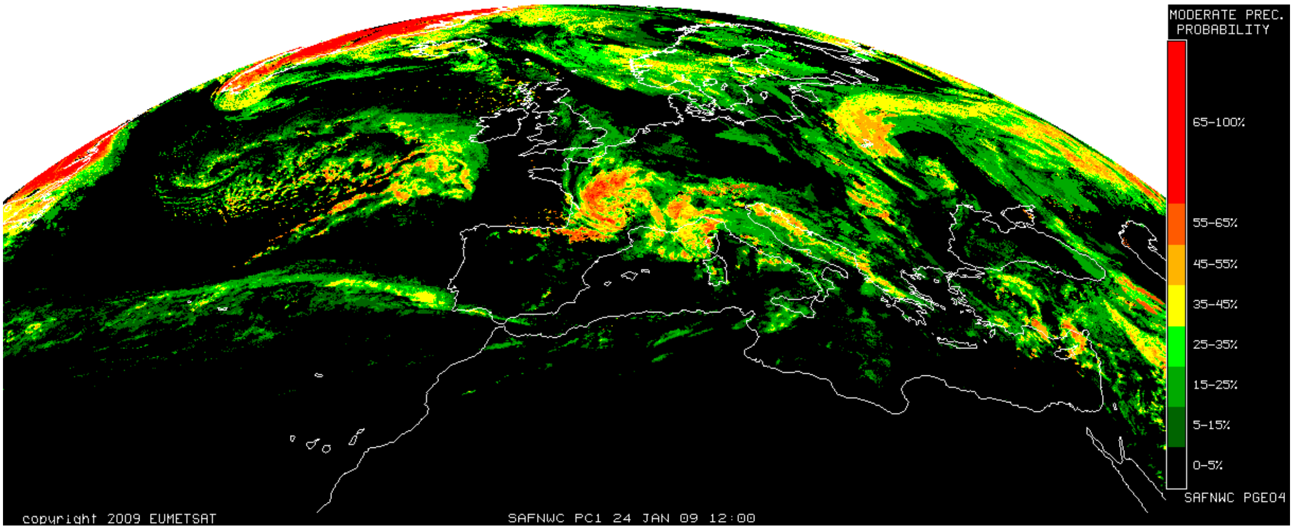



*Figure 3: 200901241200 precipitating clouds product over MSG-N, configured for day algorithm. Dark green hues present precipitation likelihood 5%-25%,light green 25%-35%, yellow hues 35%-45% and orange/red red 45% and higher*

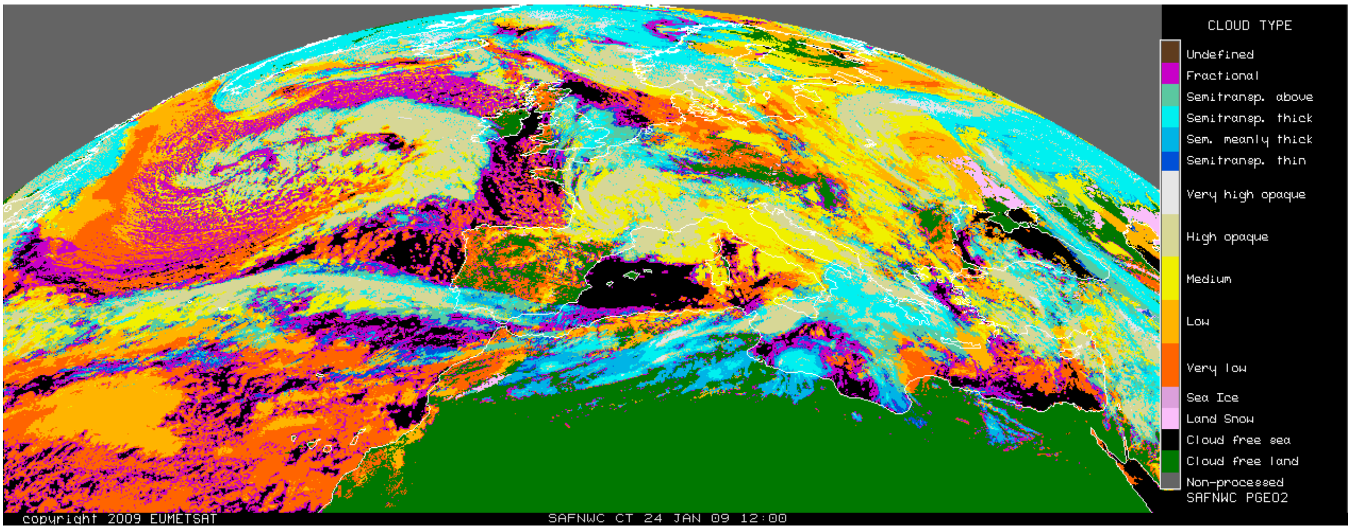


*Figure 4: Cloud type input 200901241200 over MSG-N.*

### 2.2.4 Practical considerations

#### *2.2.4.1 List of Precipitating Clouds (PC) inputs*

Satellite imagery:

The following SEVIRI brightness temperatures and visible reflectances are needed at full IR spatial resolution:

| *VIS0.6* | *NIR1.6* | *IR3.8* | *IR6.2* | *IR7.3* | *IR10.8* | *IR12.0* |
|---|---|---|---|---|---|---|
| Day-time | Day-time | Day-time | Day-time and Night-time | Day-time and Night-time | Day-time and Night-time | Day-time and Night-time |

*Table 5. PC SEVIRI inputs*

The SEVIRI channels are input by the user in HRIT format and extracted on the desired region by NWC-GEO software package.

Cloud type (CT) product output:

CT output, in NetCDF format, is mandatory input to PC.

NWP parameters:

NWP surface temperature is a mandatory input for PC.

Sun and satellite angles associated to satellite imagery

This information is mandatory. It is computed by the PC software itself, using the definition of the region and the satellite characteristics.

#### *2.2.4.2 Description of the Precipitating Clouds (PC) output*

The content of the PC output is described in the Data Output Format Document [RD 3].A summary is given below:

| Container | Content |
|---|---|
| PC | NWC GEO PC Total Precipitation Likelihood: |

| **Class** | **Total Precipitation Likelihood (%)** |
|---|---|
| 0 | 0 |
| 1 | 10 |
| 2 | 20 |
| 3 | 30 |
| 4 | 40 |
| 5 | 50 |
| 6 | 60 |
| 7 | 70 |
| 8 | 80 |
| 9 | 90 |
| 10 | 100 |
| FillValue | No data or corrupted data |

### *Geophysical Conditions*

| Field | Type | Description |
|---|---|---|
| Space | Flag | Set to 1 for space pixels |
| Illumination | Parameter | Defines the illumination condition<br><br>0: N/A (space pixel)<br>1: Night<br>2: Day<br>3: Twilight |
| Sunglint | Flag | Set to 1 if Sunglint |
| Land_Sea | Parameter | 0: N/A (space pixel)<br>1: Land<br>2: Sea<br>3: Coast |

### *Processing Conditions*

| Field | Type | Description |
|---|---|---|
| Satellite_input_data | Parameter | Describes the Satellite input data status<br><br>0: N/A (space pixel)<br>1: All satellite data are available<br>2: At least one useful satellite channel is missing<br>3: At least one mandatory satellite channel is missing |
| NWP_input_data | Parameter | Describes the NWP input data status<br><br>0: N/A (space pixel or NWP data not used)<br>1: All NWP data are available<br>2: At least one useful NWP field is missing<br>3: At least one mandatory NWP field is missing |
| Product_input_data | Parameter | Describes the Product input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All input Product data are available<br>2: At least one useful input Product is missing<br>3: At least one mandatory input Product is missing |
| Auxiliary_input_data | Parameter | Describes the Auxiliary input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All Auxiliary data are available<br>2: At least one useful Auxiliary field is missing<br>3: At least one mandatory Auxiliary field is missing |

***Quality***

| Field | Type | Description |
|---|---|---|
| Nodata | Flag | Set to 1 if pixel is NODATA |
| Internal_consistency | Flag | Set to 1 if an internal consistency check has been performed. Internal consistency checks will be based in the comparison of the retrieved meteorological parameter with physical limits, climatic limits, neighbouring data, NWP data, etc. |
| Temporal_consistency | Flag | Set to 1 if a temporal consistency check has been performed Temporal consistency checks will be based in the comparison of the retrieved meteorological parameters with data obtained in previous slots. |
| Quality | Parameter | Retrieval Quality<br>0: N/A (no data)<br>1: Good<br>2: Questionable<br>3: Bad<br>4: Interpolated |

Another file is generated including statistical information related to the product generation. It contains histograms of precipitation probability and processing flags, and it is generated in ascii format. This file may be useful to get statistics on general algorithm performance.

#### *2.2.4.3 Example of Precipitating Clouds (PC) visualisation*

Examples of both day-time and night-time PC product can be found below:

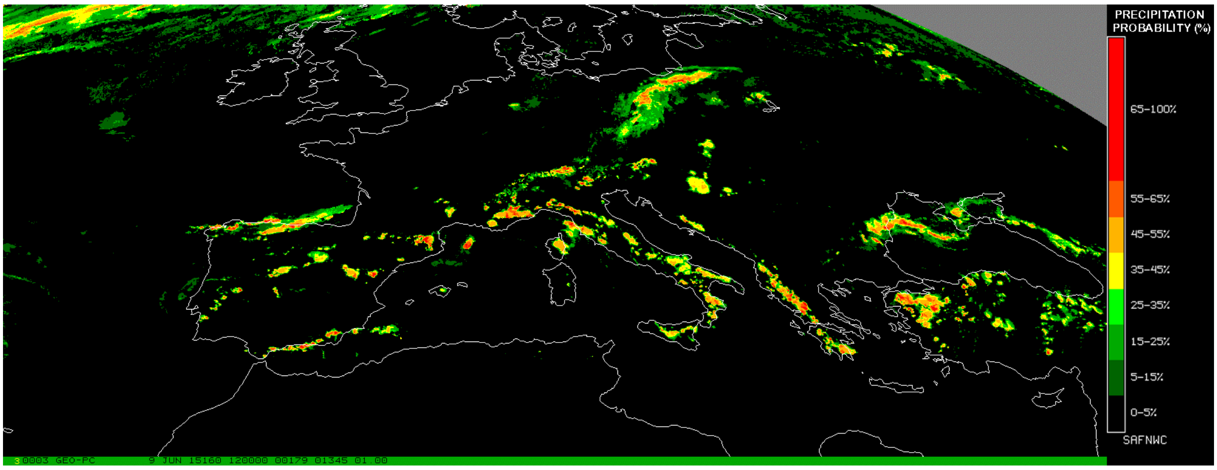



*Figure 5. Example of the precipitating clouds product over a day-time scene on 9th June 2015 at 12:00 UTC*

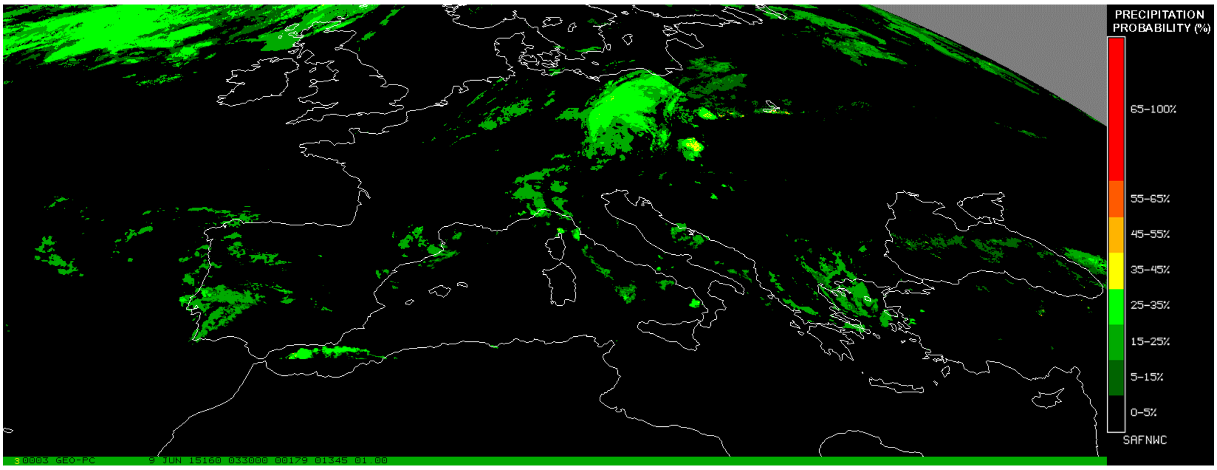



*Figure 6. Example of the precipitating clouds product over a night-time scene on 9th June 2015 at 03:30 UTC*

## 2.3 ASSUMPTIONS AND LIMITATIONS

- The current version of the product contains a certain dependence on sun zenith angle.
- There is also a clear jump in algorithm performance between day and night algorithm, which cannot be totally avoided.
- The product degrades considerably at high viewing angles and use for viewing angles greater than 60 degrees is not recommended.
- The algorithm does currently not detect any precipitation from low clouds

# 3. DESCRIPTION OF CONVECTIVE RAINFALL RATE (CRR) PRODUCT

CRR product keeps the same from the last "Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO" v2018 2.1 [RD 8]

## 3.1 CONVECTIVE RAINFALL RATE (CRR) OVERVIEW

Convective Rainfall Rate (CRR) product is a Nowcasting tool that provides information on convective, and stratiform associated to convection, instantaneous rain rates and hourly accumulations.

In the processing of the product, CRR uses some calibration analytical functions that have been calibrated taking as "truth" the radar data. There are two types of functions:

- 2-Variable (2-V) function that depends on 10.8IR and (10.8IR - 6.2WV) SEVIRI data
- 3-Variable (3-V) function that depends on 10.8IR, (10.8IR - 6.2WV) and 0.6VIS-N SEVIRI data

The 3-V calibration analytical function gives better results but there are some situations in which it can't be used, for instance, during the night time. The type of calibration to be used can be chosen by the user through the CRR model configuration file.

The analytical functions have been calibrated using radar data from:

- Baltic radar network
- Hungarian radar network
- Spanish radar network

To take into account the influence of environmental and orographic effects on the precipitation distribution, some corrections can be applied to the basic CRR value. The possible corrections are the moisture correction, the cloud top growth/decaying rates or evolution correction, the cloud top temperature gradient correction and the orographic correction.

At this stage, the CRR precipitation pattern computed in the previous step is combined with a precipitation pattern derived through a lightning algorithm.

At the end of the process CRR product produces five different outputs.

In one of them, the CRR value in mm/h is converted into classes. There are 12 classes that divide the rain rates in some different ranges and each pixel of the output image has a rain class assigned.

There exists an output that contains the information on the instantaneous rain rate in mm/h in each pixel of the image. The hourly accumulation output gives information about the precipitation occurred during the last hour.
The classes, the instantaneous rain rate in mm/h and the hourly accumulation outputs have the same colour palette.
Information on the corrections applied and the processing status is available on the CRR_QUALITY and CRR_DATAFLAG outputs respectively.

## 3.2 CONVECTIVE RAINFALL RATE (CRR) ALGORITHM DESCRIPTION

### 3.2.1 Theoretical description

In this section the theoretical basis and practical implementation of the algorithm are described.

#### *3.2.1.1 Physics of the problem*

All visible and infrared precipitation estimation schemes are necessary indirect because the radiation does not penetrate through the cloud. The cloud's brightness temperature and visible reflectance may be related to the rain falling from it, but the raindrops themselves are not directly sensed (Kidder and Vonder Haar, 1995).

The empirical relationship that the higher and thicker are the clouds the higher is the probability of occurrence and the intensity of precipitation is used in the CRR algorithm. Information about cloud top height and about cloud thickness can be obtained, respectively, from the infrared brightness temperature (IR) and from the visible reflectances (VIS) (Scofield, 1987) (Vicente and Scofield, 1996).

10.8IR-6.2WV brightness temperature difference is a useful parameter for extracting deep convective cloud with heavy rainfall (Kurino, 1996). Negatives values of the 10.8IR-6.2WV brightness temperature difference have been shown to correspond with convective cloud tops that are at or above the tropopause (Schmetz et al., 1997).

Some observable features (like environmental moisture, cloud growth, cloud top structure, topography underneath, etc.) affect to convective precipitation rates more than the stratiform rain cases (Vicente, 1998) (Vicente, 1999).

It is stated that convective phenomena are related to the electrical activity in the clouds. The lightning algorithm is based on the assumption that the higher is the spatial and temporal density of lightning occurrence, the stronger is the convective phenomenon and the higher is the probability of occurrence and the intensity of convective precipitation.

#### *3.2.1.2 Mathematical Description of the Convective Rainfall Rate (CRR) algorithm*

##### 3.2.1.2.1 Convective Rainfall Rate (CRR) algorithm outline

The CRR algorithm developed within the NWC SAF context estimates rainfall rates from convective systems, using 10.8IR, 6.2WV and 0.6VIS-N MSG SEVIRI imagery and calibration analytical functions generated by combining satellite and Radar data.

The calibration functions, which have been calibrated through a statistical process, try to connect satellite multi-band imagery with rain rates. In the calibration process composite radar data are compared pixel by pixel with geographically matched satellite data with the same resolution. Rainfall rate RR is obtained, as a function of two or three variables (10.8IR brightness temperature, 10.8IR-6.2WV brightness temperature differences and normalised 0.6VIS reflectances):

RR = f (10.8IR, 10.8IR-6.2WV, 0.6VIS-N), for 3-V calibration

RR = f (10.8IR, 10.8IR-6.2WV), for 2-V calibration

The basic CRR mm/h value for each pixel is obtained from the calibration functions. If in a pixel the sun zenith angle is lower than a threshold and the solar channel is used, the basic CRR data is obtained from a 3-V analytical function which uses 10.8IR, 6.2WV and 0.6VIS-N imagery. If in a pixel the sun zenith angle is higher than the threshold or lower, but the solar channel is not going to be used, the basic rain rate values are obtained from 2-V analytical function which only uses 10.8IR and 6.2WV imagery. The threshold that decides, depending on the sun zenith angle, whether the solar channel can be used or not, is chosen by the user through the CRR model configuration file. The name of this threshold in the configuration file is DAY_NIGHT_ZEN_THRESHOLD and its default value is 80º.

When the solar channel is used, the normalised visible reflectances are obtained dividing by the cosine of the solar zenith angle. The option of using the solar channel in the computation of the CRR values can be chosen by the user through the CRR model configuration file.

In the retrieval of basic CRR values from 3-V calibration function, some pixels could occasionally present normalised visible reflectances greater than 100. In those cases the CRR values will be retrieved using the 2-V calibration function. This occurs in few instances and has been observed mainly under very low sun illumination conditions. Those pixels can be easily identified as they will have assigned a value as a missing data in some channel in the CRR_DATAFLAG output.

A filtering process is performed in order to eliminate stratiform rain data which are not associated to convective clouds: the obtained basic CRR data are set to zero if all the pixels in a grid of a selected semisize (def. value: 3pix) centred on the pixel have a value lower than a selected threshold (def. value: 3mm/h). The threshold and the size of the grid can be modified by the user by means of the model configuration file.

To take into account the temporal and spatial variability of cloud tops, the amount of moisture available to produce rain and the influence of orographic effects on the precipitation distribution, several correction factors can be applied to the basic CRR value. Therefore, the possible correction factors are the moisture correction, the cloud top growth/decaying rates or evolution correction, the cloud top temperature gradient correction and the orographic correction.

Lightning activity can provide valuable information about convection. A lightning algorithm can be applied to derive a precipitation pattern that will be combined with the CRR one computed in the previous step in order to complement it.

At the end of the process the final values of the CRR rainfall rates in mm/h are used in order to obtain three different outputs:

- CRR rainfall rates in mm/h
- CRR classes: rainfall rate in mm/h is divided into twelve classes.
- CRR hourly accumulations: A trapezoidal integration is performed in order to compute the hourly accumulations. The description of this process can be found in ANNEX C: Hourly accumulations.

### 3.2.1.2.2 Convective Rainfall Rate (CRR) calibration analytical functions procedure

The analytical functions have been built taking the previous calibration matrices as starting point. The calibration matrices obtaining method can be read in ATBD for CRRv3.1.1.

The calibration matrices were modelled and described by the analytical functions that best fitted them. An example of this modelling can be seen in Figure3.

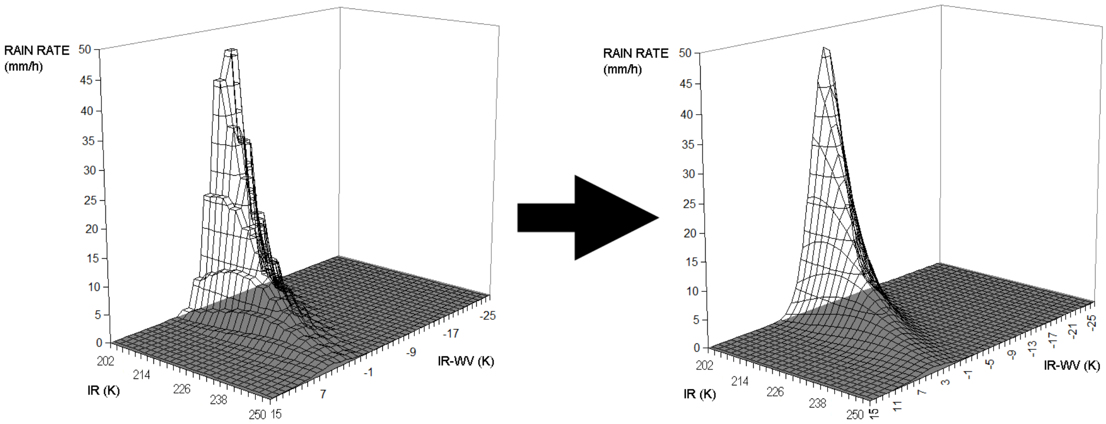



*Figure 7. From calibration matrices to analytical functions*

The perfect matching between matrices and functions is impossible to reach; also, the calibration process over a function is easier than over a matrix. For these reasons a new calibration process was done over the functions.

3.2.1.2.2.1 <u>Analytical functions calibration process</u>

The calibration process was done using the following radar data:

| Radar network | Type of radar | Frequency Scanning Radar | Dataset used | Type of product used | MSG scans over the radar area | Matching time |
|---|---|---|---|---|---|---|
| **Baltrad network** | C- Band | 15 minutes | 21 rainy days June-August 2004 | Pseudo-CAPPI at 2Km | About 11 min later than the MSG time slot | MSG time slot 15 min later than the radar one. |
| **Hungarian radar network** | C- Band | 15 minutes | 18 rainy days May-September 2009 | Maximum reflectivity in the vertical (MRV) and Echotop | About 11 min later than the MSG time slot | MSG time slot 15 min later than the radar one. |
| **Spanish radar network** | C- Band | 10 minutes | 111 rainy days throughout 2009 | PPI and Echotop | About 10 min later than the MSG time slot | 0 and 30 min MSG slots have been matched to 10 and 40 |

| | |
|---|---|
| | min radar images respectively. |

*Table 6. Description of the radar calibration data*

For a better matching of radar – satellite images, the radar products were converted into MSG projection using a bi-linear interpolation scheme.

A quality control has been used for the Spanish radar dataset taking advantage of the quality image generated for the radar national composite products (Gutierrez and Aguado, 2006). No quality control methods have been used for Baltrad and Hungarian radar datasets.

Ground echoes, like anomalous propagation echoes, were removed in Pseudo-CAPPI, MRV and PPI scenes. To that end 10.8IR SEVIRI imagery were used together with the basic AUTOESTIMATOR algorithm (Vicente et al., 1998).

Considering that CRR is a specific product for convective situations, only images with convective echoes, as far as possible, were used during the calibration process. To that end, Echotop product was used when available. Only scenes where the ratio between the number of echoes greater than 6 Km and the ones greater than 0 Km was lower than 15% in the Echotop image were selected.

Since images with convective situations can also include non convective echoes, a calibration area was selected. This selection included the area corresponding to 15x15 pixels boxes centred on that ones that reached a top of 6 km and a rain rate of 3 mm/h simultaneously.

Since the perfect matching is not possible a smoothing process in 3x3 pixels boxes was done for a better radar-satellite matching.

Once the radar calibration dataset was prepared, CRR was run using the analytical functions applying small shifts to the coefficients. Also a smoothing process in 3x3 pixels boxes was done over CRR imagery. Then several comparisons between CRR rain rates and radar rain rates were done computing accuracy and categorical scores. Special attention was paid to RMSE, POD and FAR. The coefficients of the functions were adjusted and the ones which got the best scores were chosen.

###### 3.2.1.2.2.2 Analytical functions description

An analytical function is easier to handle and to analyze than a big matrix. Two calibration functions were obtained:

**2-V calibration function**: RR (10.8IR, 10.8IR-6.2WV)

The function independent variable is (10.8IR-6.2WV) SEVIRI data and its coefficients have a dependence on 10.8IR SEVIRI data. The mathematical formulation of this function is the following:

$$RR(mm/h) = H(IR) * \exp\left[ -0.5 * \left( \frac{(IR - WV) - C(IR)}{W(IR)} \right)^2 \right]$$

Where RR is the rain rate in mm/h, and H(IR), C(IR) and W(IR) are coefficient functions depending on 10.8IR SEVIRI data.

Looking at the formula of this function it can be observed that it is a symmetric bell-shaped curve where H(IR) is the height, C(IR) is the position of the symmetry axis and W(IR) is related to the width of the curve. All these parameters, depending on 10.8IR data, have a meaning.

The mathematical formula of the coefficient function related to the height of the 2-V calibration function, H(IR), is the following:

$$H(IR) = a * \exp[b * IR]$$

Where the coefficients are: a = $8*10^8$ and b = - 0.082

According to these coefficients the graph of this curve is shown in Figure 8.

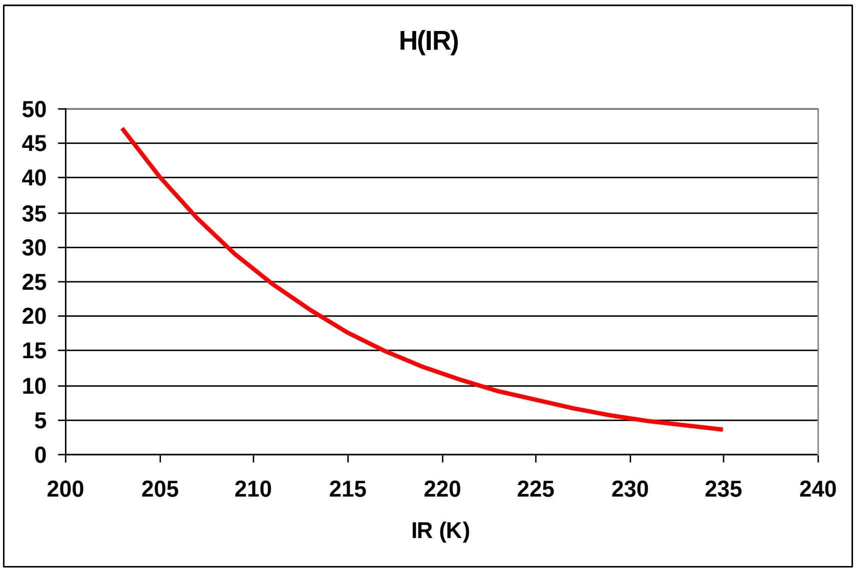



*Figure 8. Height of the 2-V function plotted between 205K and 235K*

It is clear from the curve that the lower the IR brightness temperature the higher H(IR), so the higher are the estimated rain rates.

Regarding the position of the symmetry axis C(IR), the formula is:

$$C(IR) = c * IR + d$$

Where the coefficients are: c = 0.2 and d = - 45.0

This function is plotted in Figure 9.

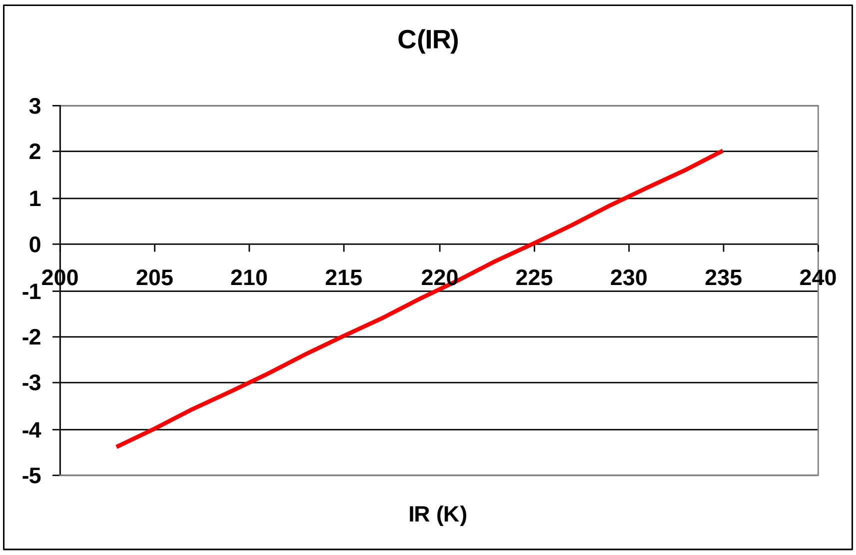



*Figure 9. Coefficient related to the position of the symmetry axis of the 2-V function*

As it has been seen, the 2-V calibration function is a symmetric bell-shaped curve whose independent variable is (10.8IR-6.2WV) and whose coefficients depend on IR. The symmetry axis of the "bell curve" is given by C(IR). Looking at Figure 9 it can be deduced that the highest rain rates are estimated for 10.8IR-6.2WV values close to zero; and the lower are the 10.8IR brightness temperatures, the lower the value of 10.8IR-6.2WV that provides the highest rain rates estimations.

Finally, the equation that provides information on the width of the bell-shaped curve is:

$$W(IR) = f * \exp\left[-0.5\left(\frac{IR+g}{h}\right)^2\right] + j$$

Where f = 1.5; g = - 215.0; h = 3.0 and j = 2.0

The graph of the W(IR) is plotted in Figure 10.

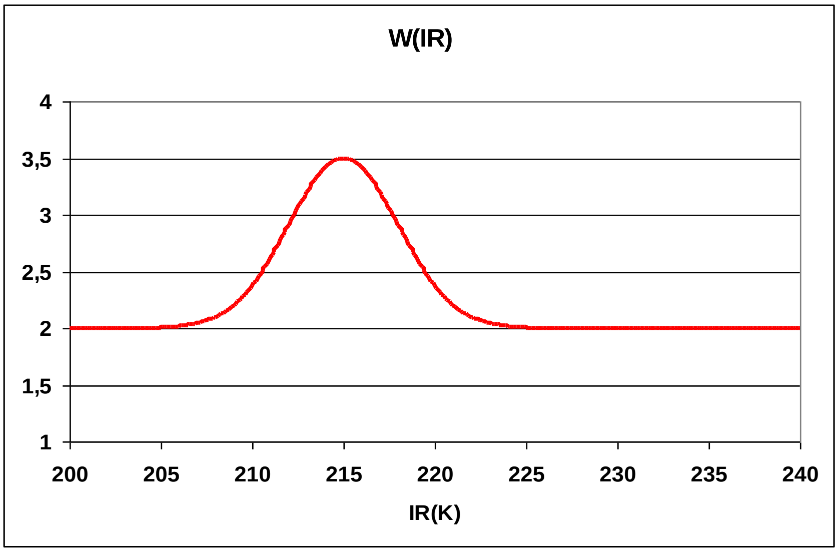



*Figure 10. Coefficient that provides information on the width of the 2-V function*

W(IR) is also a symmetric bell-shaped curve whose symmetry axis is centred in 215K. This means that for this brightness temperature the curve gets wider so it could be deduced that for IR=215K, there is a higher likelihood of precipitation occurrence although the rain rates are not the highest.

**3-V calibration function**: RR(10.8IR, 10.8IR-6.2WV, 0.6VIS)

The 3-V function independent variables are 10.8IR-6.2WV and 0.6VIS-N SEVIRI data and its coefficients have dependence on 10.8IR SEVIRI data and on latitude. Its mathematical formulation is the following:

$$RR(mm/h) = \underbrace{\exp\left[-0.5*\left(\frac{VIS\_N - C\_Vis(Lat)}{8.5}\right)^2\right]}_{(Factor\_VIS-N)} \underbrace{* H(IR) * \exp\left[-0.5*\left(\frac{(IR-WV)-C(IR)}{W(IR)}\right)^2\right]}_{(Factor\_IRWV)}$$

The 3-V calibration function is the product of two symmetric bell-shaped curves, Factor_VIS-N and Factor_IRWV. The Factor_IRWV one is similar to the 2-V function and Factor_VIS-N depends on the VIS-N imagery.

The interpretation of the bell-shaped curve Factor_IRWV is the same as in the case of the 2-V function. For the 3-V function the H(IR), C(IR) and W(IR) coefficients are the following:

$$H(IR) = a * \exp[b * IR]$$

Where: a = $1.25*10^8$ and b = - 0.073

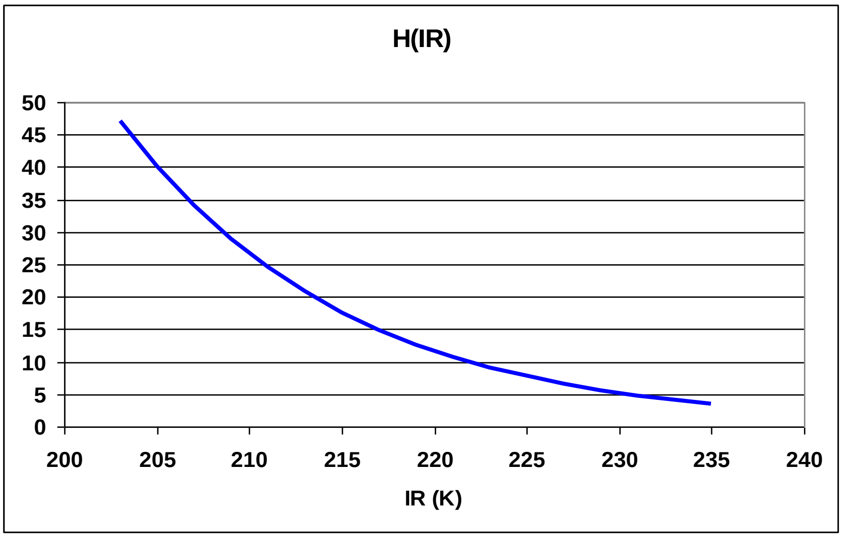



*Figure 11. Height of the 3-V function plotted between 205K and 235K .*

$$C(IR) = c * IR + d$$

Where: c = 0.25 and d = - 53.75

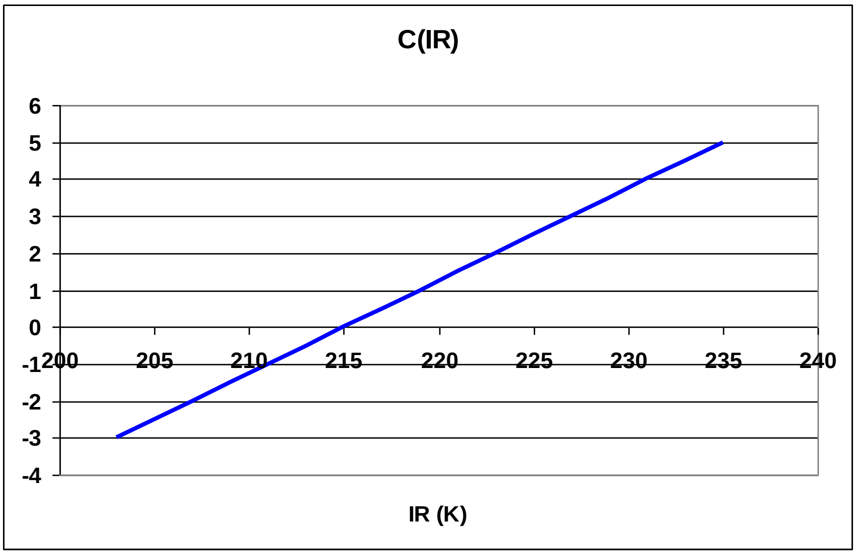



*Figure 12. Coefficient related to the position of the symmetry axis of the 3-V function .*

$$W(IR) = f * \exp\left[-0.5\left(\frac{IR+g}{h}\right)^2\right] + j$$

Where: f = 1.5; g = - 227.0; h = 14.0 and j = 4.0

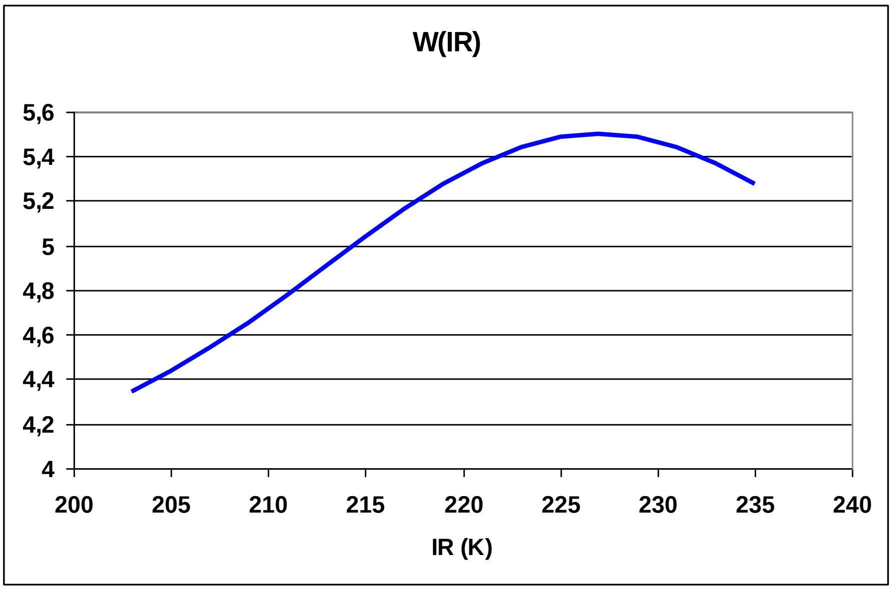



*Figure 13. Coefficient that provides information on the width of the 3-V function*

Regarding the H(IR) coefficients for 2-V and 3-V functions, both the shape and the maximum rain rates estimated are very similar.

As for the position of the symmetry axis, the lower the 10.8IR brightness temperatures, the lower the value of 10.8IR-6.2WV that provides the highest rain rates estimations for both 2-V and 3-V functions. The difference is that in 3-V case, the (10.8IR-6.2WV) values that provide the highest rain rates are a bit higher than in the case of 2-V function.

In the case of the coefficient that provides information on the width of the 2-V and 3-V functions, the difference is higher. It can be observed that the 3-V function is always much wider and the 10.8IR brightness temperature for which there is a higher likelihood of precipitation occurrence is warmer (227K) than in the case of the 2-V function. This means that 3-V function rain rates estimations are higher for the same range of 10.8IR brightness temperatures and (10.8IR-6.2WV) differences than 2-V function rain rates estimations. 2-V function limits the rain rate estimations to lower IR brightness temperatures.

It must be taken into account that 3-V function is also composed of other symmetric bell-shaped curve Factor_VIS-N that depends on the VIS-N imagery. It can be interpreted that Factor_IRWV is the height of Factor_VIS-N, so the highest estimations given by 3-V function will be given by Factor_IRWV, and Factor_VIS-N filters these estimations depending on the normalized visible reflectances.

The higher is the 0.6VIS-N reflectance, the higher is the optical thickness of the cloud so the higher should be the rain rate assigned. This can be seen in Figure 14.

It has been seen that for Spanish latitudes the highest rain rates are obtained for VIS reflectances of about 82%, for different years. According to the other radar-satellite datasets (Hungary and Baltrad) reflectances that provide the highest rain rates decrease with latitude. The quantity of solar energy that reaches higher latitudes is lower than the ones that reach latitudes closer to the equator and normalization process is not good enough to fix this problem. This dependence on the latitude could be a corrective effect additional to the normalization.

To take account of this fact a latitude dependency has been included in the 3-V function. As can be observed in Figure 14, the lower is the latitude the higher is the reflectance for which 3-V function assigns higher rain rates. This latitude dependence can be observed in Figure 15.

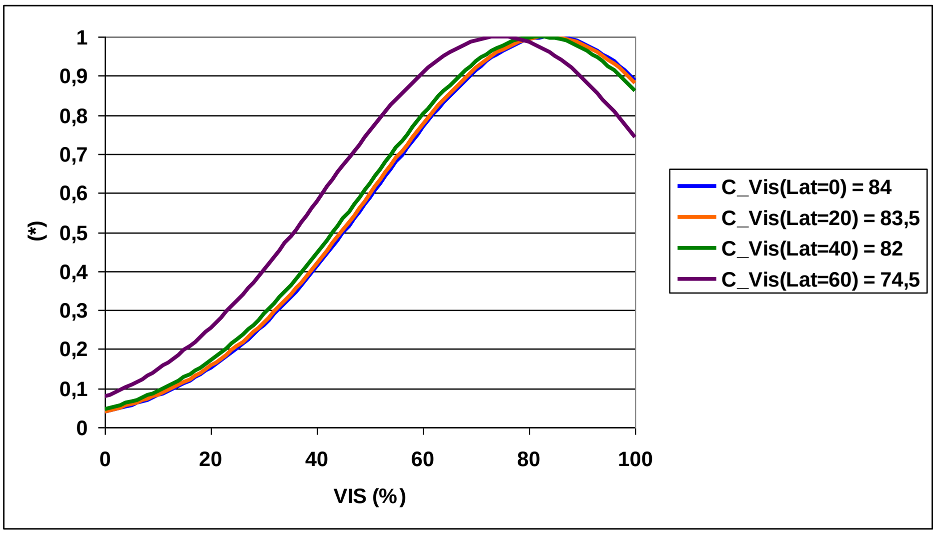



*Figure 14. Dependence of the 3-V function on the Normalized Visible Reflectances.*

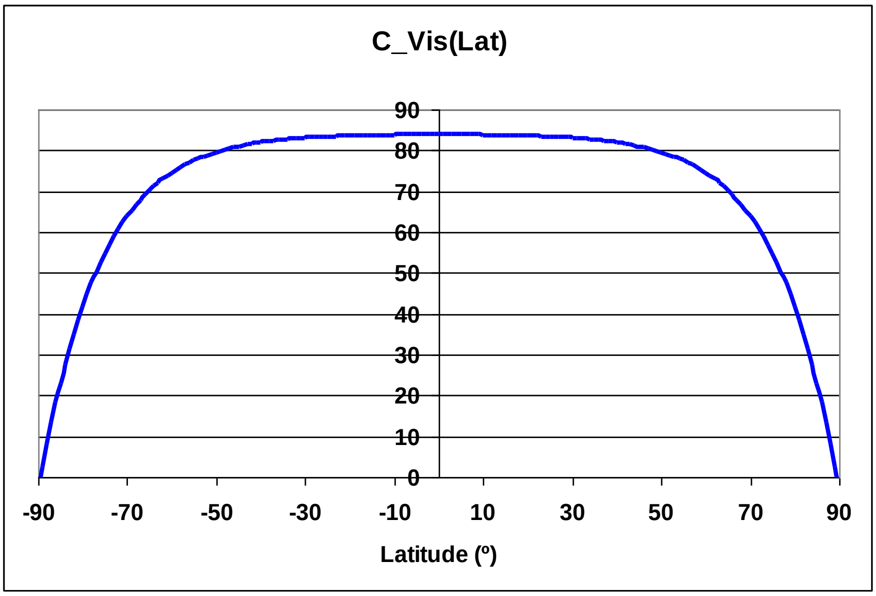



*Figure 15. Dependence of the Normalized Visible reflectances on Latitude*

#### 3.2.1.2.3 Convective Rainfall Rate (CRR) correction factors description

##### 3.2.1.2.3.1 Moisture Correction Factor

When thunderstorms take place in quite moist environments the computed rainfall rate should be greater than when they occur in dry air masses. To take into account this effect a moisture correction factor has been developed. It adjusts the estimates when the air is dry or quite moist. This factor has been defined as the product of the total precipitable water, PW, in the layer from surface to 500 hPa. by the relative humidity, RH, (mean value between surface and 500 hPa. level), obtained from a numerical model.

In order to compute the PWRH factor, the precipitable water is expressed in inches of water and the relative humidity in percentage. This factor takes values between 0.0 and 2.0. An environment is considered to be dry if PWRH is significantly below 1.0 and quite moist if PWRH is greater than 1.0.

The PWRH factor decreases rainfall rates in very dry environments and increases them in very moist ones. However, for high latitudes where convective systems can contain hail (so that radar rainfall is unrealistically high), if IR cloud top temperature is lower than 215K , there is no need to increase the rainfall rates, but instead, it is necessary to decrease them whenever the environment is dry (PWRH<1.0). Based on this justification, the following criterion is applied:

If latitude >55°N, T10.8 < 215 K and PWRH >1.0 the computed rainfall rate should not be multiplied by the PWRH correction factor.

Otherwise, the computed rainfall rate is multiplied by the PWRH correction factor.

##### 3.2.1.2.3.2 Cloud Growth Rate Correction Factor

Convective rain is assumed to be associated with growing clouds exhibiting overshooting tops. Consecutive satellite IR images are used to indicate vertically growing and decaying cloud systems.

A convective system is more active and produces greater rainfall rates when the tops are becoming colder and expanding. Based on the conclusion that decaying clouds with cold tops that are becoming warmer produce little or no rainfall, the output is modified according to the following:

· If a IR pixel in the second scene is colder than in the first one, convection is intensifying, so rainfall rate computed in that pixel with the information from the second scene remains the same.

· If a IR pixel in the second scene is warmer than in the first one, convection is weakening. In this case, rainfall rate computed with the information from the second scene is multiplied by a coefficient. The coefficient value can be modified by the user through the Keyword COEFF_EVOL_GRAD_CORR_00 in the model configuration file (Default value for Normal Mode (0.35) is set in the configuration file. Recommended value for Rapid Scan mode is 0.55).

· If there is no change in the cloud-top temperature in the two consecutive scenes (no growth or decay), rainfall rate computed from the second scene stays the same.

Therefore, the cloud growth correction factor, also designated as evolution correction factor, is only applied if the analysed pixel becomes warmer in the second image.

##### 3.2.1.2.3.3 Cloud-top Temperature Gradient Correction Factor

When consecutive IR scenes are not available, cloud growth rate correction factor can not be applied. Then cloud-top temperature gradient correction is used instead.

This alternative correction method is based on the fact that much information can be extracted from cloud-top structure on a single IR image.

Cloud-top temperature gradient correction factor, also designated as gradient correction factor, is based on a search of the highest (coldest) and lowest (less cold) cloud tops. The concept of finite difference is used to locate the maximum and minimum local temperature within grids of 3x3 or 5x5 pixels centred on the point $P_0=(x_0,y_0)$. The idea is to search for the pixels that are below the average cloud top surface temperature (local temperature minima) and assume that these pixels indicate active convection connected to precipitation beneath.

Cloud-top temperature can be named as $T=T(x,y)$, where T is the cloud-top temperature as a function of the x and y co-ordinates. For those pixels whose T is lower than 250K, the following analysis is done:

Maxima and minima can be found studying the first and second derivative of T. The process is the following:

Second derivative of T in the point $P_0=(x_0, y_0)$:

$$T''_x(x = x_0) = \left.\frac{\partial^2 T}{\partial x^2}\right|_{x=x_0}$$

$$T''_y(y = y_0) = \left.\frac{\partial^2 T}{\partial y^2}\right|_{y=y_0}$$

$$T''_{xy}(x = x_0, y = y_0) = \left.\frac{\partial^2 T}{\partial x \partial y}\right|_{x=x_0, y=y_0}$$

Hessian in $P_0=(x_0, y_0)$:

$$H = (T''_x(x = x_0)) \cdot (T''_y(y = y_0)) - (T''_{xy}(x = x_0, y = y_0))^2$$

$P_0$ is characterized in the following way:

- *$H > 0$ and $T''_x(x=x_0) < 0$ $\Rightarrow$ maximum*
- *$H > 0$ and $T''_x(x=x_0) > 0$ $\Rightarrow$ minimum*
- *$H < 0$ $\Rightarrow$ no maximum, no minimum*
- *$H = 0$ $\Rightarrow$ not known*

Once this analysis has been done in a grid of 3x3 pixels, the previous derived rainfall rate is adjusted in the following way:

- If the pixel $P_0$ has a temperature maximum, indicating a relatively low cloud top with $P_0$ warmer than its surrounding, the previous rainfall rate is multiplied by a coefficient whose value can be modified by the user through the keyword COEFF_EVOL_GRAD_CORR_01 in the model configuration file (Default value: 0.25).
- If the pixel $P_0$ has a temperature minimum, which means that $P_0$ is colder than the surrounding indicating a high cloud top, the previous rainfall rate stays the same.
- If $P_0$ has not a temperature maximum or minimum, which means that $P_0$ is at the same height and temperature as the surrounding pixels, the previous rainfall rate is multiplied by a coefficient whose value can be modified by the user through the keyword COEFF_EVOL_GRAD_CORR_02 in the model configuration file (Default value: 0.50).
- If $P_0$ temperature can not be defined as a maximum or a minimum, the whole process is repeated using pixels within a 5x5 pixel's grid.
- Finally, if $P_0$ temperature remains undefined as a maximum or a minimum within the 5x5 pixel's grid, the original rainfall rate value is not modified.

###### 3.2.1.2.3.4 Orographic Correction Factor

Local topography has long been recognised to have an effect on the distribution and intensity of precipitation. However, the rain induced by orographic forcing is a complex process associated with complicated flows. Rainfall amounts are dependent on the atmospheric flow over the mountains and on the characteristics of the flow disturbances created by the mountains themselves.

This correction factor uses the interaction between the wind vector (corresponding to 850 hPa level from the NWP) and the local terrain height gradient in the wind direction to create a multiplier that enhances or diminishes the previous rainfall estimate, as appropriate.

The wind direction for the 48-km grid cell containing the location being tested is assumed to be constant in magnitude and direction. A one-dimensional cross-section of the terrain, determined by the wind direction, is extracted from the elevation map. The wind path length, D pixels, is variable from 3 km (pixel resolution) to 24 km (8 pixels), depending upon wind speed. Accordingly, D is determined by a 15-minute fetch (converted into units of pixels) of the wind speed U:

$$D = U * \frac{900s}{3000m/pixel}$$

The extracted terrain cross-section extends D pixels upwind and downwind from the reference site, giving a total length of 2D+ 1 pixel. The height of the test location can be denoted as $Z_{D+1}$; the location farthest upwind is $Z_1$, the location farthest downwind is $Z_{2D+1}$. The slope between a point A and a downwind point B can be defined as

$$S_{AB} = \frac{(Z_B - Z_A)}{(B - A)}$$

For each pixel, A, upwind of the site and the site itself (i.e., from 1 to D+1), the slope between it and each point B within D pixels downwind is calculated (i.e., from A+1 to A+D). The maximum slope found for each point A is retained as the slope $S_A$. The net slope S, used for the correction, is equal to the mean of the $S_A$ values.

Finally, we can define a rainfall rate enhancement parameter, M, as the result of the vertical velocity induced by a wind with horizontal speed U blowing over a surface with slope of S. Since M should not have effect on the rainfall amounts on a flat terrain, it can be written as:

$$M = 1 + S * U$$

M is limited to be between 0.2 and 3.5. Every CRR rain point is multiplied by the co-located M values. The eight pixels all around the image edge can not be corrected.

##### 3.2.1.2.3.5 Parallax Correction Factor

For a better convective precipitation area location a parallax correction [ANNEX A: Parallax Correction] can be applied to this product. This option is chosen by the user through the product model configuration file and it is applied by default.

#### 3.2.1.2.4 Lightning algorithm

As lightning activity is related with convection, an option to use this information to improve precipitation estimates has been added to the product.

An algorithm for rainfall estimation using lightning information has been developed. Its description can be found in ANNEX B: Lightning algorithm.

### 3.2.2 Practical considerations

#### 3.2.2.1 List of Convective Rainfall Rate (CRR) inputs

Satellite imagery:

The following SEVIRI brightness temperatures and normalized visible reflectances are needed at full IR spatial resolution:

| T10.8µm | TPrev10.8µm | T6.2µm | VIS0.6µm |
|---|---|---|---|
| Mandatory | Optional* | Mandatory | Optional |

*Table 7. CRR SEVIRI inputs*

The SEVIRI channels are input by the user in HRIT format and extracted on the desired region by NWC-GEO software package.

* If TPrev10.8µm is not available, the Cloud Growth Rate Correction Factor cannot be computed but the Cloud-top Temperature Gradient Correction Factor is computed instead as an alternative.

Numerical model:

This information is mandatory for moisture and orographic corrections. When this information is not available, CRR is computed without applying these two corrections.

Parallax correction can run without the NWP parameters using the climatic profile.

For moisture correction:

Relative Humidity at 1000, 925, 850, 700 and 500 hPa

Dew Point temperature at 2 m

Temperature at 2 m

Temperature at 1000, 925, 850, 700, 500 hPa

Surface Pressure

For parallax correction:

Temperature at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa

Geopotential at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa

For orographic correction:

U and V wind components in 850 hPa

Lightning information file for CRR:

A file with information on every lightning strike occurred in a time interval is mandatory to choose the option of adjusting the CRR precipitation pattern with the lightning information provided by ground based lightning detection networks. Information about this lightning information file structure can be found in the Interface Control Document for Internal and External Interfaces of the NWC/GEO [RD 4].

Sun angles associated to satellite imagery

This information is mandatory for normalising the VIS image when the solar channel is used. It is also used to choose whether to run day-time or night-time algorithm.

Ancillary data sets:

All this information is included in the software package:

- Saturation Vapour table is mandatory for Humidity correction and is located in the $SAFNWC/import/Aux_data/ CRR directory.
- Saturation Vapour Polynomial Coefficients table is mandatory for Humidity correction and is located in the $SAFNWC/import/Aux_data/CRR directory.
- Elevation mask is mandatory for orographic correction and is located in the $SAFNWC/import/Aux_data/ Common directory.
- Climatic profile is necessary as a backup for Parallax correction in case NWP is not available. This information is located in the $SAFNWC/import/Aux_data/CRR directory

Model configuration file for CRR:

The CRR model configuration file contains configurable system parameters in the product generation process related to algorithm thresholds, ancillary datasets, numerical model data, corrections to be applied, etc. The complete list of these parameters and the explanation of the most useful ones is available in the User Manual for the Precipitation Product Processors of the NWC/GEO [RD 5].

#### *3.2.2.2 Description of the Convective Rainfall Rate (CRR) output*

The content of the CRR output is described in the Data Output Format Document [RD 3]. A summary is given below:

<table>
<tr><th>Container</th><th>Content</th></tr>
<tr><td>crr</td><td>NWC GEO CRR Convective Rainfall Rate Class:<br><br>
<table>
<tr><th>Class</th><th>Rainfall Intensity (mm/h)</th></tr>
<tr><td>0</td><td>[ 0.0, 0.2)</td></tr>
<tr><td>1</td><td>[ 0.2, 1.0)</td></tr>
<tr><td>2</td><td>[ 1.0, 2.0)</td></tr>
<tr><td>3</td><td>[ 2.0, 3.0)</td></tr>
<tr><td>4</td><td>[ 3.0, 5.0)</td></tr>
<tr><td>5</td><td>[ 5.0, 7.0)</td></tr>
<tr><td>6</td><td>[ 7.0, 10.0)</td></tr>
<tr><td>7</td><td>[10.0, 15.0)</td></tr>
<tr><td>8</td><td>[15.0, 20.0)</td></tr>
<tr><td>9</td><td>[20.0, 30.0)</td></tr>
<tr><td>10</td><td>[30.0, 50.0)</td></tr>
<tr><td>11</td><td>[50.0, )</td></tr>
<tr><td>FillValue</td><td>No data or corrupted data</td></tr>
</table>
</td></tr>
<tr><td>crr_intensity</td><td>NWC GEO CRR Convective Rainfall Intensity:<br><br>crr_intensity(mm/h) = scale_factor * counts + add_offset<br><br>where:<br>scale_factor = 0.1<br>add_offset = 0.0</td></tr>
<tr><td>crr_accum</td><td>NWC GEO CRR Convective Hourly Rainfall Accumulation:<br><br>crr_accum(mm) = scale_factor * counts + add_offset<br><br>where:<br>scale_factor = 0.1<br>add_offset = 0.0</td></tr>
<tr><td>crr_status_flag</td><td>13 bits indicating<br>Applied Corrections:<br>Bit 0: Humidity correction applied<br>Bit 1: Evolution correction applied<br>Bit 2: Gradient correction applied<br>Bit 3: Parallax correction applied<br>Bit 4: Orographic correction applied<br>Use of optional data:<br>Bit 5: Solar channel used<br>Bit 6: Lightning data used<br>Processing information<br>Bit 7: crr_intensity set to 0 due to filtering process<br>Bit 8: crr_intensity was a hole because of the parallax correction, and then was filled by the median filter<br>Bit 9,10, 11: Use of bands for accumulation<br>1: All required bands were available<br>2: One previous CRR band is missing<br>3: At least two previous CRR bands are missing (no consecutive)<br>4: At least two previous CRR bands are missing (some are consecutive)<br>Bit 12: Accumulation quality flag. Set to 1 if:<br>not all crr values are available to perform the accumulation,<br>OR<br>any of the crr_intensity values was set to 0 due to filtering process<br>OR<br>Any of the crr_intensity values was a hole because parallax correction</td></tr>
</table>

### *Geophysical Conditions*

| Field | Type | Description |
|---|---|---|
| Space | Flag | Set to 1 for space pixels |
| Illumination | Parameter | Defines the illumination condition<br><br>0: N/A (space pixel)<br>1: Night<br>2: Day<br>3: Twilight |
| Sunglint | Flag | Set to 1 if Sunglint |
| Land_Sea | Parameter | 0: N/A (space pixel)<br>1: Land<br>2: Sea<br>3: Coast |

### *Processing Conditions*

| Field | Type | Description |
|---|---|---|
| Satellite_input_data | Parameter | Describes the Satellite input data status<br><br>0: N/A (space pixel)<br>1: All satellite data are available<br>2: At least one useful satellite channel is missing<br>3: At least one mandatory satellite channel is missing |
| NWP_input_data | Parameter | Describes the NWP input data status<br><br>0: N/A (space pixel or NWP data not used)<br>1: All NWP data are available<br>2: At least one useful NWP field is missing<br>3: At least one mandatory NWP field is missing |
| Product_input_data | Parameter | Describes the Product input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All input Product data are available<br>2: At least one useful input Product is missing<br>3: At least one mandatory input Product is missing |
| Auxiliary_input_data | Parameter | Describes the Auxiliary input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All Auxiliary data are available<br>2: At least one useful Auxiliary field is missing<br>3: At least one mandatory Auxiliary field is missing |

### *Quality*

| Field | Type | Description |
|---|---|---|
| Nodata | Flag | Set to 1 if pixel is NODATA |
| Internal_consistency | Flag | Set to 1 if an internal consistency check has been performed. Internal consistency checks will be based in the comparison of the retrieved meteorological parameter with physical limits, climatic limits, neighbouring data, NWP data, etc. |
| Temporal_consistency | Flag | Set to 1 if a temporal consistency check has been performed Temporal consistency checks will be based in the comparison of the retrieved meteorological parameters with data obtained in previous slots. |
| Quality | Parameter | Retrieval Quality<br>0: N/A (no data)<br>1: Good<br>2: Questionable<br>3: Bad<br>4: Interpolated |

### 3.2.2.3 Example of Convective Rainfall Rate (CRR) visualisation

#### 3.2.2.3.1 Instantaneous Rates

Below is shown an image corresponding to CRR classes output. It has been obtained at full resolution and all corrections have been applied.

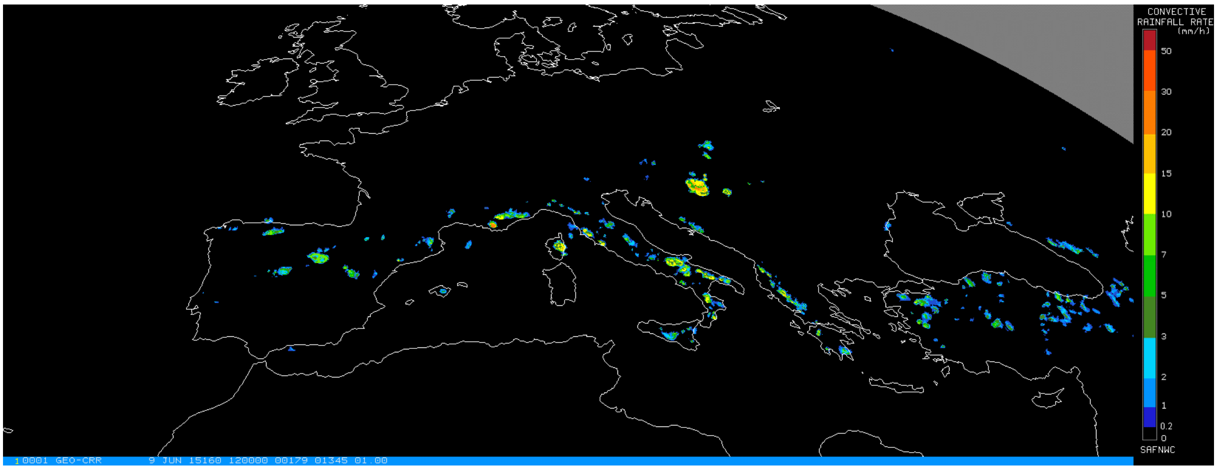



*Figure 16. CRR instantaneous intensities output corresponding to 9th June 2015 at 12:00Z*

#### 3.2.2.3.2 Hourly Accumulations

Below is shown an image corresponding to CRR hourly accumulations output. It has been obtained at full resolution and all corrections have been applied.

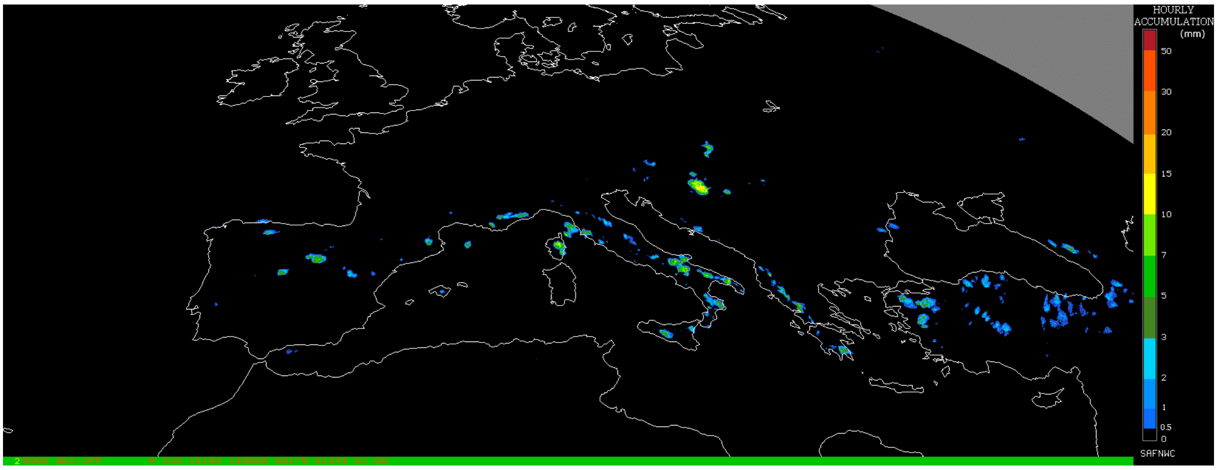



*Figure 17. CRR hourly accumulations output corresponding to 9th June 2015 at 12:00Z*

## 3.3 ASSUMPTIONS AND LIMITATIONS

The CRR product is based on a calibration method which requires the availability of a training set of precipitation data derived from radar information, to be used as ground truth to derive the relationship between satellite information and rainfall rate.

Regarding the radar data:

- The drop size distribution, used to obtain the radar rainfall rates (mm/h) from the radar reflectivity (dBZ), has been assumed to be the Marshall Palmer type throughout the calibration and validation procedures.

- No online operational method has been applied in order to adjust the radar rainfall intensities using rain gauge measurements.

- The limited availability of radar data at the time of carrying out the CRR calibration caused that three different radar datasets, with different radar products, had to be used. In the case of the Spanish radar data, PPI product were used and a quality control, taking advantage of a quality image generated for the radar national composite products (Gutierrez and Aguado, 2006), was used. In the case of the Hungarian radar data, rain rates based on Maximum reflectivity in the vertical were used, while in the case of Baltrad network, Pseudo-CAPPI at 2Km were used to derive rain rates. Is should be borne in mind that no quality control methods were used for Baltrad and Hungarian radar datasets.

- Data from the radar networks in different areas were not compared to an independent reference.

Regarding the lightning algorithm:

- The CRR lightning algorithm and the coefficients applied have been derived for Spain using the lightning information from the AEMET lightning detection network. Concerning this particular, it is important to highlight that ground based lightning detection networks provide information with different performances in detection efficiency and location accuracy. For this reason, in the model configuration file the keyword APPLY_LIGHTNING is set to 0 and by default the lightning information is not used.

- Before to use the lightning algorithm it is highly recommended to the user to adapt the coefficients to the specific performances of the lightning detection network serving that information.

- This issue can be solved in a satisfactory manner with the use of lightning information provided by MTG Lightning Imager which still requires of a technical adaptation and calibration.

## 3.4 References

Algorithm Theoretical Basis Document for "Convective Rainfall Rate" (CRR-PGE05 v3.1.1). SAF/NWC/CDOP/INM/SCI/ATBD/05.

Gutierrez, J. M. and Aguado, F.: Quality image for the Spanish Radar National Composite, Proceedings of ERAD 2006, 318-320.

Jorge Sánchez-Sesma and Marco Antonio Sosa: EPPrePMex, A Real-time Rainfall Estimation System Based on GOES-IR Satellite Imagery. IPWG, October 2004, Monterey, California, USA.

Kidder, S.Q., and T.H. Vonder Haar, 1995: Satellite Meteorology: An Introduction. Academic Press

Kurino, T., 1996: A Rainfall Estimation with the GMS-5 Infrared Split-Window and Water Vapour Measurements, Tech. Rep., Meteorological Satellite Centre, Japan Meteorological Agency.

Schmetz J., S. S. Tjemkes, M. Gube and L. van de Berg, 1997: Monitoring deep convection and convective overshooting with METEOSAT. Adv. Space Res., Vol. 19, pp433-441.

Scofield, R.A., 1987: The NESDIS operational convective precipitation estimation technique, Mon. Wea. Rev., Vol.115, pp.1773-1792.

Tapia, A., Smith, J. A., Dixon, M., 1998: Estimation of Convective Rainfall from Lightning Observations, Bull. American Meteorological Society, Vol. 37, pp. 1497-1509.

Vicente, G.A. and R.A. Scofield, 1996: Experimental GOES-8/9 derived rainfall estimates for flash flood and hydrological applications, Proc. 1996 Meteorological Scientific User's Conference, Vienna, Austria, EUM P19, pp.89-96.

Vicente, G.A., Davenport, J.C. and Scofield, R.A., 1999: The role of orographic and parallax corrections on real time high resolution satellite rainfall estimation, Proc. 1999 Eumetsat Meteorological Satellite Data User's Conferences, EUM P26, pp. 161-168.

Vicente, G.A., Scofield, R.A. and Menzel W.P. 1998: The Operational GOES Infrared Rainfall Estimation Technique, Bull. American Meteorological Society, Vol. 79, No. 9, pp. 1883-1898.

# 4. DESCRIPTION OF CONVECTIVE RAINFALL RATE FROM CLOUD PHYSICAL PROPERTIES (CRRPH) AND PRECIPITATING CLOUDS FROM CLOUD PHYSICAL PROPERTIES (PCPH)

## 4.1 CRRP$_H$ AND PCP$_H$ GENERAL OVERVIEW

Convective Rainfall Rate from Cloud Physical Properties (CRRPh) product, developed within the NWC SAF context, is a Nowcasting tool that provides information on convective and stratiform associated to convection, instantaneous rain rates and hourly accumulations.

Precipitating Clouds from cloud physical properties, developed within the NWC SAF context, is a Nowcasting tool that provides estimation on the probability of precipitation (PoP) occurrence.

In our contest, PoP is defined as the instantaneous probability that a rain rate greater or equal 0.2 mmh$^{-1}$ occurs at the pixel level.

The PoP and rain intensities estimations are done collecting information from the Cloud Phase, the Effective Radius (Reff) and the Cloud Optical Thickness (COT). All these three parameters are part of the CMIC retrieval [RD 6]. Based on COT and Reff, the Cloud Water is computed (CWP = 2/3 COT * REFF). The Cloud Phase is available during the whole day whereas Reff and COT are only available at day time. Along with the microphysical information, SEVIRI channels are also used: five infrared channels ($IR_{8.7}$,$IR_{9.7}$,$IR_{108}$,$IR_{120}$,$IR_{134}$), one visible channel ($VIS_{0.6}$) that has been normalized and corrected with the sun-earth distance and two water vapour channels ($WV_{6.2}$,$WV_{7.3}$).

Both CRRPh and PCPh use algorithms based on a Principal Component Analysis (PCA) which is a statistical procedure that uses an orthogonal transformation which converts a set of correlated variables into a set of uncorrelated ones. This way a complex problem with many dimensions to deal with is compressed and reduced into a lower number of variables keeping the maximal amount of information.

It has to be stressed that this first conceptual step of the algorithm ends up with a dimension reduction of the predictors. Only the first two Principal Components, which explains the 95% of the variance, are kept. This dimension reduction also potentially implies a noise reduction which will provide smoother data for fitting. Additionally there are more benefits. By reducing the dimension of the original problem the computation is also being accelerated. To complete the first picture of this conceptual part, it should be taken into account that these Principal Components are used to develop a LUT.

When choosing the number of variables to develop the Look Up Table, a balance between loss of information and potential noise reduction has to be found. This balance has been searched in the testing phase of the developing algorithm.

Since the products (CRRPh and PCPh) make use of the same inputs during the whole day and the Visible channels and the Cloud Water Path (CWP = 2/3 COT * REFF) derived from CMIC are only available at day time, it is necessary to create a psedo-VIS06 and a pseudo-CWP to be used at night time. The way of creating this pseudo-channels is explained later in section 6.

One positive aspect of the CRRPh and PCPh algorithms relies on the fact that make use of the same inputs for the whole day including the microphysics. Although the input information is different between day and night, the fact of using the same mathematical relationship in day and night cases effectively provides a more continuous product.

## 4.2 PHYSIC OF THE PROBLEM

Reflected IR solar radiation by the cloud tops can be useful to obtain information on microphysics and rain processes near cloud tops (Pilewskie and Twomey, 1987). The radiative properties of a cloud can be characterized through the Effective Radius ($R_{eff}$) and Cloud Optical Thickness (COT).

The most relevant measure that indicates the possibility of occurrence of rain formation processes in observed clouds is the effective radius (Rosenfeld and Gutman, 1994). The effective radius is defined as the ratio of the third to second moments of the droplet size distribution.

$$R_{eff} = \frac{\int_0^\infty N(r) r^3 dr}{\int_0^\infty N(r) r^2 dr}$$

Where N(r) is the concentration of particles having radius r.

Cloud optical thickness depends on the moisture density as well as the vertical thickness of the cloud. The higher is the COT, the higher is the possibility of occurrence of rain formation processes. It is possible to retrieve COT values from satellite data (Roebeling et al., 2006).

Two SEVIRI channels are used, together with a radiative transfer model, in order to retrieve $R_{eff}$ and COT. The cloud reflectance at VIS0.6 channel is directly related with COT while $R_{eff}$ is connected with the reflectance variations measured in near infrared channels like $NIR_{1.6}$ and $IR_{3.8}$. Due to the number of disadvantages that IR3.8 channel presents (Roebeling et al., 2006), NIR1.6 has been used.

Several studies that connect cloud physical properties and rain occurrence have been developed (Nauss and Kokhanovsky, 2007; Roebeling and Holleman, 2009).

The Effective Radius and the Cloud Optical Thickness used by this algorithm are retrieved within the CMIC algorithm [RD 6] and are available at day time.

The NWC SAF CMIC product contains information relevant to the cloud top (thermodynamical phase, cloud particle size) or integrated on the full vertical extent (optical depth, liquid and ice water path).

Apart from the information derived from the Cloud Microphysics properties on top of the clouds the SEVIRI channels add relevant and useful information:

$VIS_{0.6}$ is essential for cloud detection and it helps to tracking with fine detail.

$IR_{8.7}$ provides quantitative information on thin cirrus clouds and support the discrimination between ice and water clouds. It is also necessary for scene identification and the atmospheric instability product.

$IR_{9.7}$ is a channel focused on ozone detection. Thunderstorm's downdrafts may carry ozone from higher latitudes to lower levels.

$IR_{10.8}$ and $IR_{12.0}$ are known as infrared window channels, as they view the surface of the Earth and the cloud tops, being little affected by gaseous absorption within the atmosphere.

$IR_{13.4}$ detects the $CO_2$ absorption and leads to infer the height of semi-transparent clouds.

$WV_{6.2}$ measures the upper-troposphere water vapour content and along with the $WV_{7.3}$ which provides information on the humidity at mid-troposphere, help to identify convection and severe storms.

A combination of SEVIRI channels and cloud microphysics is done via a Principal Component Analysis.

Principal Component Analysis belongs to a family of techniques known as unsupervised learning. PCA only extracts information about the predictors (cloud microphysics, visible, infrared and water vapour channels) and does not try to relate them to the output (rainfall rates).

When dealing with a very large dataset, a natural instinct is to try to reduce its size, whilst minimising any loss of information, in order to better understand and interpret the structure of the data. This way a set of variables that are highly correlated are reduced into a smaller set of uncorrelated ones.

PCA's technical were first developed by Pearson in the late nineteen century and later by Hotelling in the 30's of the twenty century. However it was not until the emergence of computers and digital technologies that were more spread and finally used.

PCA's have been widely used in weather and climate research to explain precipitating patterns, climatic variability, to compute climatic indices. It has also been used in remote sensing to extract information of the land, flood mapping, etc.

As physics equations do not distinguish between day and night it has been developed a unique algorithm for the whole day with the same inputs.

Visible channels, in particular $VIS_{0.6}$ add accuracy while identifying clouds and the water content add very useful information that should be used during the whole day. That it is the reason why a simulation of both channels is trained at day time with the infrared and water vapour channels and used at night time.

## 4.3 THEORETICAL DESCRIPTION

As it has been said in previous section, PCA method is a statistical method that uses an orthogonal transformation which converts a set of correlated variables into a lower set of uncorrelated ones.

Since the variable range is different, to avoid numerical computation artefacts, a renormalization process is required such that all variables are within a limited range. A normal transformation is applied, in which the final variables have a mean value of zero and a standard deviation of one.

Eigenvectors, with the biggest eigenvalues, portray the directions where the data variance is bigger. They are oriented towards the largest variance. Eigenvalues define their magnitude.

In probability theory the variance of a random variable is a dispersion measure.

For this reason, In order to find the Principal Components that gather all the dataset information it is necessary to compute first the covariance matrix since it provide us with a joint dispersion among variables.

Since it is wanted to reduce the dimensionality of the dataset, loosing as lower information as possible, eigenvectors with the lower eigenvalues will be discarded, because they add little information to the problem, and potentially could lead in model overfitting.

Data are projected into a lower dimension space, transforming the original dataset (nine dimensions) into a two dimension problem centred in their two principal components.

CRRPh basic output can be modified by optionally applying three main correction factors: the content of water vertically integrated through the cloud layer, the lighting activity and a stability correction. PCPh basic output can be modified by optionally applying the stability correction.

Further recalibrations will be needed to adapt it to MTG.

**DATASET AND CALIBRATION AREA:**

CRRPh and PCPh have been calibrated with a list of days throughout 2015 that accomplished at least one of two criteria.

First criterion: A particular day is included in the calibration list if the percentage of pixels with ET (Echotop)>6km. with respects to pixels with ET>0km. exceeds the threshold of 65% at least for one slot of this day. Echotop values in km. correspond with the maximum height that echoes bigger than 12dBz are able to reach.

Second criterion: It is calculated the proportion of radar pixels with RFR>=0.2mm/h with respect to the whole image. Whenever at least one slot of a day reaches the percentage limit of 8% the day was also included in the calibration and validation list.

RFR in mm/h is obtained from the lowest Plan Position Indicator (PPI) of the radar using the Marshal-Palmer relation, $Z=200R^{1.6}$, where Z ($mm^6\,mm^{-3}$) is the reflectivity factor and R($mm\ h^{-1}$) is the rainfall rate.

PPI composite of the C-Band Spanish Radar network have been used. Since the radar outputs are available every 10 minutes and the MSG scanning over Spain takes place about 8 minutes later than the MSG slot time, 0 and 30minutes MSG slots have been matched to the 10 and 40 minutes radar images respectively.

For a better matching of radar and satellite images, the radar products have been converted into MSG projection using a bi-linear interpolation scheme. In order to compare radiances and reflectance from SEVIRI channels with radar, information from SEVIRI channels and CWP have been parallax corrected.

A radar quality image has been used as a filter image to get rid of spurious echoes, such as windmill echoes. Anomalous propagation echoes have been removed through the 10.8IR scene. A rain image has been obtained from the 10.8IR data using the basic AUTOESTIMATOR algorithm (Vicente et al., 1998).

With the aim of having a more homogeneous rain rate radar database to calibrate the CRRPh precipitating product, a box size of 25*25 pixels centred in those pixels with ET>6km and rain rates>10 $mmh^{-1}$ have been selected. If the calibrating area includes a similar proportion of radar pixels with high rain rates, medium rain rates, lower rain rates and no rainy pixels the calibrating process with Principal Components will find better connections between radar, SEVIRI channels and the Cloud Water Path (CWP).These choices have proven more useful than the ones from the previous CRRPh version (ET>6km, RR>3$mmh^{-1}$, box size 15*15).To visualize this difference, Figure 18 below shows the former calibrating box area, which was used in the previous CRRPh version, on the left, along with the current one in the middle.

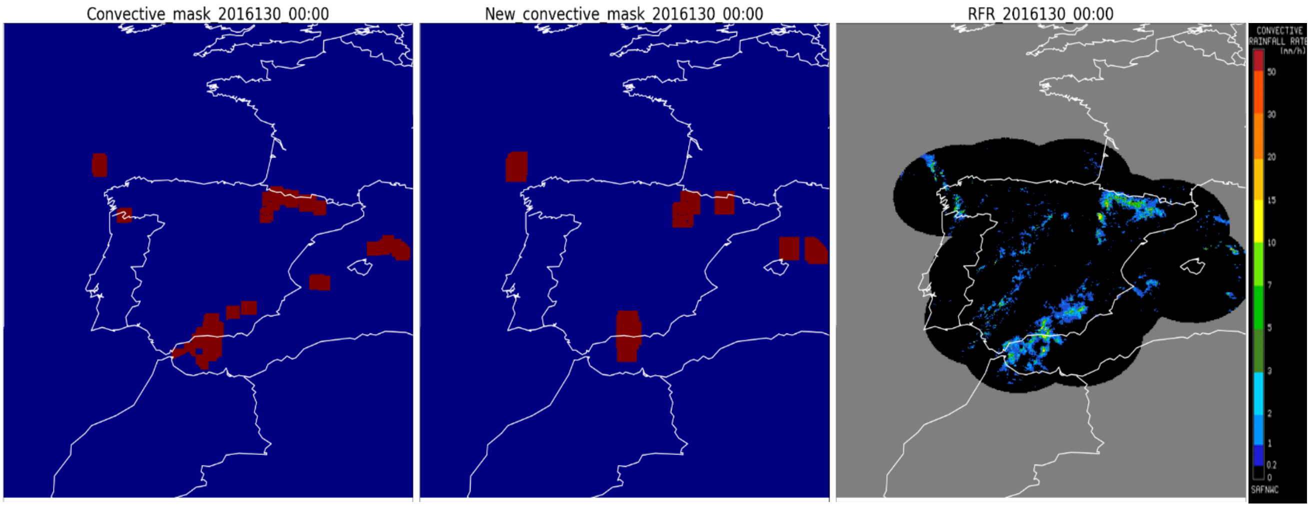



*Figure 18. Calibrating convective box areas, for previous CRRPh version(left),and for the current CRRPh (middle).Spanish composite radar on the right side.*

Unlike CRRPh, which is focused on convective events, PCPh provide us the probability of rain in any circumstances, not giving priority to convection. That is the reason behind the calibration area gathers other standards. Boxes size of 25*25 pixels centred in those pixels with rain rates > 0.2 $mmh^{-1}$ have been selected. No restriction to echotops have been set. With the aim of having a manageable dataset, a limit of 10 random boxes, gathering the criterion explained before, have been set at every time slot.

It has not been considered necessary to show PCPh calibration mask in a figure because it involves almost every pixel in the Spanish composite radar where there is precipitation. As the criteria in this case is to choose 25*25 boxes around rainy pixels higher to only 0.2 $mmh^{-1}$ (0.2$mmh^{-1}$ is the Spanish radar precipitation threshold).

As far as the calibration area is concerned, every pixel has a size of 3*3 km at the subsatellite point. At Spain latitude it would be 3.4*3.4 km more or less. Box sizes are consequently 85*85 km. The idea of choosing a higher box is to have a more homogenous dataset. One of the problems related with the radar dataset relies on the fact the proportion of zeros and low values of rain rates is very high. By choosing a higher threshold in the rain rate for calibrating the CRRPh product, the proportion of high values of rain rates is increased and the proportion of low values of radar rain rates decrease.

### 4.3.1 Inputs

**SATELLITE**:

- $IR_{8.7}$, $IR_{9.7}$, $IR_{10.8}$,$IR_{12.0}$,$IR_{13.4}$ (Brightness temperature)
- $VIS_{0.6}$ (Normalized reflectance and corrected with Sun distance)
- $WV_{6.2}$,$WV_{7.3}$ (Brightness temperature)

**NWC/GEO software:**

- GEO-CMIC ($CMIC_{COT}$,$CMIC_{REFF}$,$CMIC_{Phase}$)

**Numerical model:**

- Temperature at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa.
- Geopotential at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa.

This information is used by default for parallax correction along with $IR_{108}$. In case of lack of NWP parameters parallax correction will be run using a climatic profile.

**Ancillary data sets:**

- Climatic profile is necessary as a backup for Parallax correction in case NWP is not available. This information is included in the software package and is located in the $SAFNWC/import/Aux_data directory.

**Model configuration file for CRRPh**:

CRRPh model configuration file contains configurable system parameters in the product generation process related to algorithm thresholds, ancillary datasets, numerical model data, corrections to be applied, etc. The complete list of these parameters and the explanation of the most useful ones is available in the User Manual for the Precipitation Product Processors of the NWC/GEO [*RD 5*].

#### 4.3.1.1 *Inputs pre-processing*

The illumination conditions based on the solar zenith angle play a role in the day/night pixel transition. The DAY_NIGHT_ZEN_THRESHOLD is set to 70 degrees in the configuration file by default [RD 5].

A) **DAY**

CWP=2/3*COT*REFF ($gm^{-2}$)

If COT or REFF are not available (Fill Value) CWP is set to NODATA

B) **NIGHT**

Since the CRRPh algorithm uses the same inputs during the whole day and $CMIC_{COT}$, and $CMIC_{REFF}$ are not available at night , $VIS_{0.6}$ normalized and CWP are simulated. See section number 6.

#### 4.3.1.2 *Inputs pre-checking*

CRRPh is set to NODATA if satellite inputs are not available, CMIC Phase is undefined or is not available or $CMIC_{COT}$ and $CMIC_{REFF}$ derived from the NWC/GEO software are Fill Values.

CRRPh is set to zero in the following cases:

- CMIC Phase output indicates cloud free.
- CWP < 350 $gm^{-2}$ for both day and night time.
- $VIS_{0.6}$ normalized reflectivity < 30 for both day and night.

The threshold value for the Cloud Water Path to establish rainy/no rainy pixels is especially important to detect shallow rain. Different authors have established different thresholds, for example Goddard profiling algorithm (GPROF) has employed the rain/no-rain threshold value of 0.3 kg $m^{-2}$. Judith A. Curry, Christopher D. Ardeel and Lin Tian established a threshold of 350 $gm^{-2}$ for middle clouds and 500 $gm^{-2}$ for the onset of precipitation in low clouds.

### 4.3.2 Normalization

Normalizing every pixel consists of subtracting a fixed value (mean value) from every input channel and divide by another fixed value (standard deviation).

This table have been obtained from the training dataset.

Normalized value$_{CHANNEL}$=(Pixel value $_{CHANNEL}$ –Mean value $_{CHANNEL}$)/Standard Deviation $_{CHANNEL}$

| Channel | Mean value | Standard Deviation |
|---|---|---|
| CWP ($gm^{-2}$) | 685.16662125 | 970.72914466 |
| IR108 (K) | 242.03643436 | 21.27664634 |
| IR120 (K) | 240.61806287 | 20.58345066 |
| IR134 (K) | 234.23087442 | 13.51909372 |
| IR87 (K) | 242.08655814 | 20.78070428 |
| IR97 (K) | 235.34225756 | 10.67359157 |
| VIS06 (K) | 58.34996219 | 19.11695598 |
| WV62 (K) | 227.09770767 | 6.63554842 |
| WV73 (K) | 233.72949906 | 12.15742079 |

*Table 8. Normalizing Parameters to compute the CRRPh*

### 4.3.3 Projections

Every pixel should be processed as follows:

$P_1$=CWP normalized *$v_{11}$+$IR_{10.8}$ normalized *$v_{12}$+IR120*$v_{13}$+.......+$WV_{7.3}$normalized*$v_{19}$

$P_2$=CWP normalized *$v_{21}$+$IR_{10.8}$ normalized *$v_{22}$+IR120*$v_{23}$+.......+$WV_{7.3}$normalized*$v_{29}$

| Eigenvectors | | | |
|---|---|---|---|
| $v_{11}$ | -0.19426604 | $v_{21}$ | -0.83583786 |
| $v_{12}$ | 0.3590499 | $v_{22}$ | -0.09086839 |
| $v_{13}$ | 0.35865646 | $v_{23}$ | -0.10932737 |
| $v_{14}$ | 0.35813887 | $v_{24}$ | -0.13268772 |

| | | | |
|---|---|---|---|
| $v_{15}$ | 0.35933997 | $v_{25}$ | -0.07710212 |
| $v_{16}$ | 0.35286535 | $v_{26}$ | -0.08497276 |
| $v_{17}$ | -0.29268894 | $v_{27}$ | -0.44038809 |
| $v_{18}$ | 0.35407908 | $v_{28}$ | -0.14455189 |
| $v_{19}$ | 0.3343171 | $v_{29}$ | -0.18856688 |

*Table 9. Eigenvectors to compute the projections*

### 4.3.4 Look up calibration table

The calibration Look up table has the following characteristics:

- 2Dimensions with 200 bins.
- X Axis goes from -8 to 8, Y Axis goes from -10 to 10, both with 200 divisions.
- Central value of the first division is -7.96 along X axis with a 0.08 increase. Highest value is 7.96.
- Central value of the first division is -9.95 along Y axis with a 0.1 increase. Highest value is 9.95.

Once P1 and P2 have been calculated for every pixel CRRPh is computed as follows:

$$CRRPh=LUT(x=p1,y=p2)$$

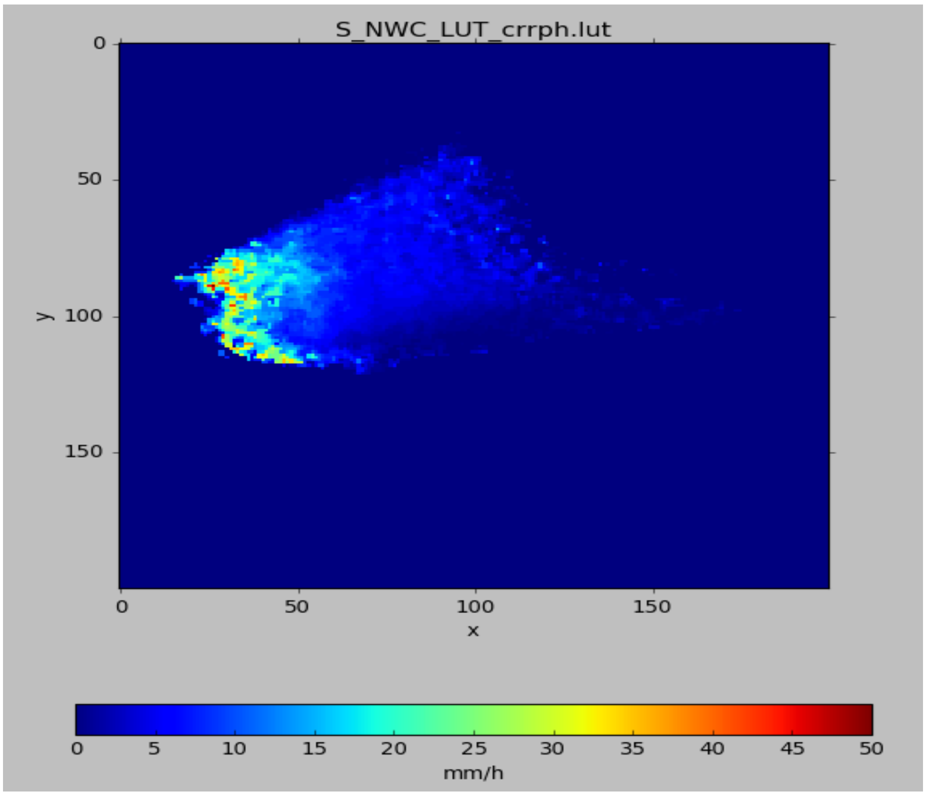



*Figure 19. 200*200 calibration LUT for CRRPh*

The colour palette is the Z axis and it represents the ninety percentile of the radar intensity over the calibration year.

For every pixel of the image, there are different brightness temperatures and radiances. After having applied equation in section 4.3.3, all this information have been compressed in p1 and p2. Now, it is time to connect p1 and p2 for every pixel with the rain intensities provided by the Spanish radar. Since it has been processed 70 days throughout a year, for every (p1, p2) there are a large amount of rain intensities values. Within this radar dataset there are many zeros. It has been chosen the ninety percentile of the radar dataset associated to every (p1, p2) not to finally obtain in the CRRPh algorithm low rain intensities values.

The LUT has been smoothed in 3*3 boxes with a median filter to reduce some noise.

Next step consists of rescaling the signal by doing a linear interpolation. 3 mm/h is set to zero and the maximum value, which is 36 mm/h is set to 50 mm/h.

It has been observed in different tests conducted with MSG data that establishing these empirical thresholds gives reasonable good results. Whenever MTG data be available a final tuning of this thresholding will be done.

Figure 19 represents the final calibration LUT for the CRRPh.

### 4.3.5 CRRPh CORRECTION FACTORS

A) Cloud Water Path Correction Factor

An enhancement correction factor based on the Cloud Water Path (CWP) has been incorporated, being available for the whole day. This way, the CRRPh output have been modified, providing with more rainfall rate in those areas with more content of water.

This CWP enhancement only applies to areas that exceed a specific threshold, being convective cores much more prone to be affected by this correction than areas with stratiform clouds. This correction factor applies to day time and also to night time. Since at night time the simulation of the Cloud Water Path hardly reach values higher than 1500 gm$^{-2}$ the limit to start applying the correction factor starts at 650 gm$^{-2}$ and this correction factor increases with higher values of the pseudo-CWP. The threshold at day time starts at 4000 and it is also not steady, it increases with higher values of the Cloud Water Path.

Day Time:

**CWP >=5000** gm$^{-2}$

crrp_intensity=crrph_intesity***3**

**CWP >= 4500** gm$^{-2}$ **y CWP < 5000** gm$^{-2}$

crrph_intensity=crrph_intensity***2.5**

**CWP>=4000** gm$^{-2}$ **y CWP<4500** gm$^{-2}$

crrph_intensity=crrph_intensity***2**

Night Time:

**CWP_simulated>=1500** gm$^{-2}$

crrph_intensity=crrph_intensity * **3**

**CWP_simulated>=1250** gm$^{-2}$ y **CWP_simulado <1500** gm$^{-2}$

crrph_intensity=crrph_intensity ***2**

**CWP_simulated= 1000** gm$^{-2}$ y **CWP_simulado** < **1250** gm$^{-2}$

crrph_intensity=crrph_intensity * **1.75**

**CWP_simulated>= 650** gm$^{-2}$ y **CWP_simulado** < **1000** gm$^{-2}$

crrph_intensity=crrph_intensity ***1.5**

These thresholds have been empirically chosen and they are optimized for summer convection. Day time and night time thresholds are different. Night time thresholds are lower than day time ones because the CWP simulated underestimates real CWP at night time.

As it can be checked later in Figure 40, section 6, the number of pixels to apply the enhancement at night time is much larger than the number of pixels to apply the enhancement at day time. However, it has been observed that many of these pixels that exceed the night threshold of 650 gm$^{-2}$ are eliminated by the stability correction (explained below).

B) Lightning module

As lightning activity is related with convection, an option to use this information to improve precipitation estimates has been added to the product.

An optional algorithm for rainfall estimation using lightning information has been developed. Its description can be found in ANNEX B: Lightning algorithm

C) Stability correction

Like other NWC SAF products do, such as the RDT-CI (Rapidly Developing Thunderstorm – Convection Warning) a stability mask is used. This mask make use of the NWP data to compute several convective indexes: K Index (KI), Showalter (SHW) and Lifted Index (LI). The combination of such indexes allow to identify stable regions where convection is unlikely to happen.

If pixel value of LI index stable (>0) and pixel value of SHW index stable (>3) and pixel value of KI index stable (< 20), then it will considered full stable case at pixel level.

Precipitation output for CRRPh and PCPh is removed in those stable regions. This stable mask is an optional parameter and configurable by the user. It is set to use it by default.

## 4.4 EXAMPLE OF VISUALIZATION

**DAY**

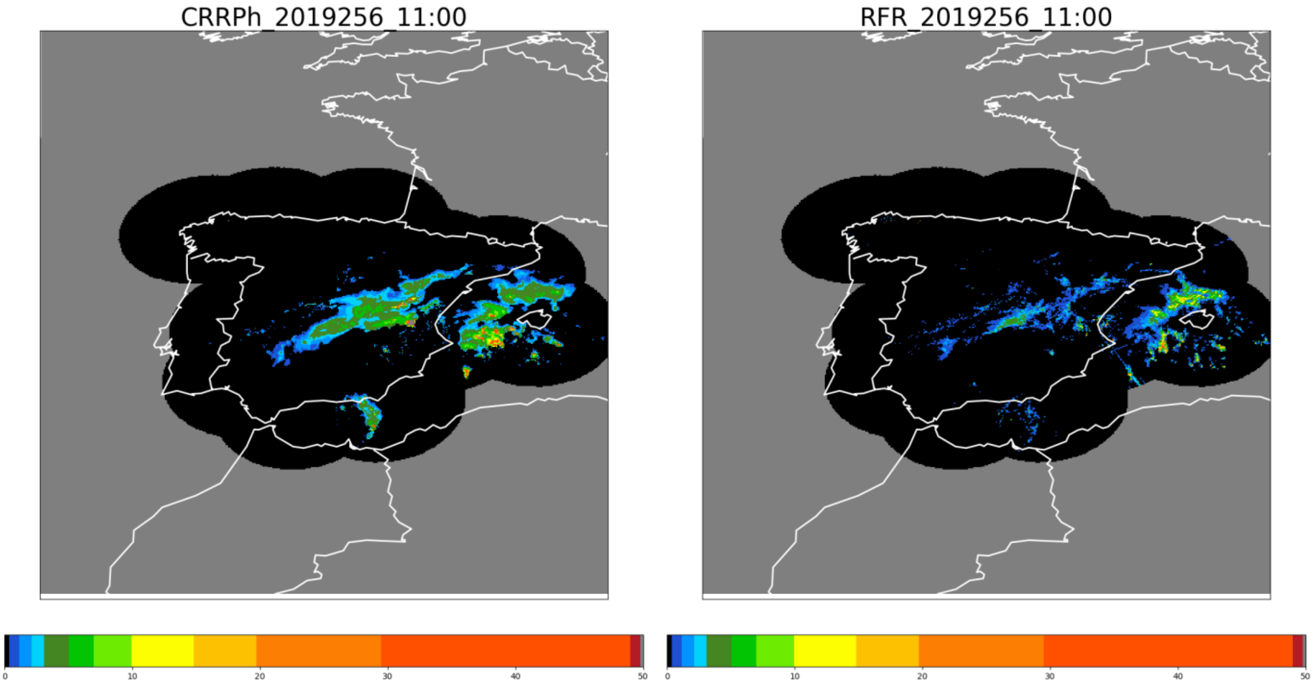



*Figure 20. CRRPh instantaneous rain rate intensities (mmh$^{-1}$) over Spain the 13$^{th}$ September 2019 at 11:00Z on the left side. Spanish composite radar (mmh$^{-1}$) on the right side.*

**NIGHT**

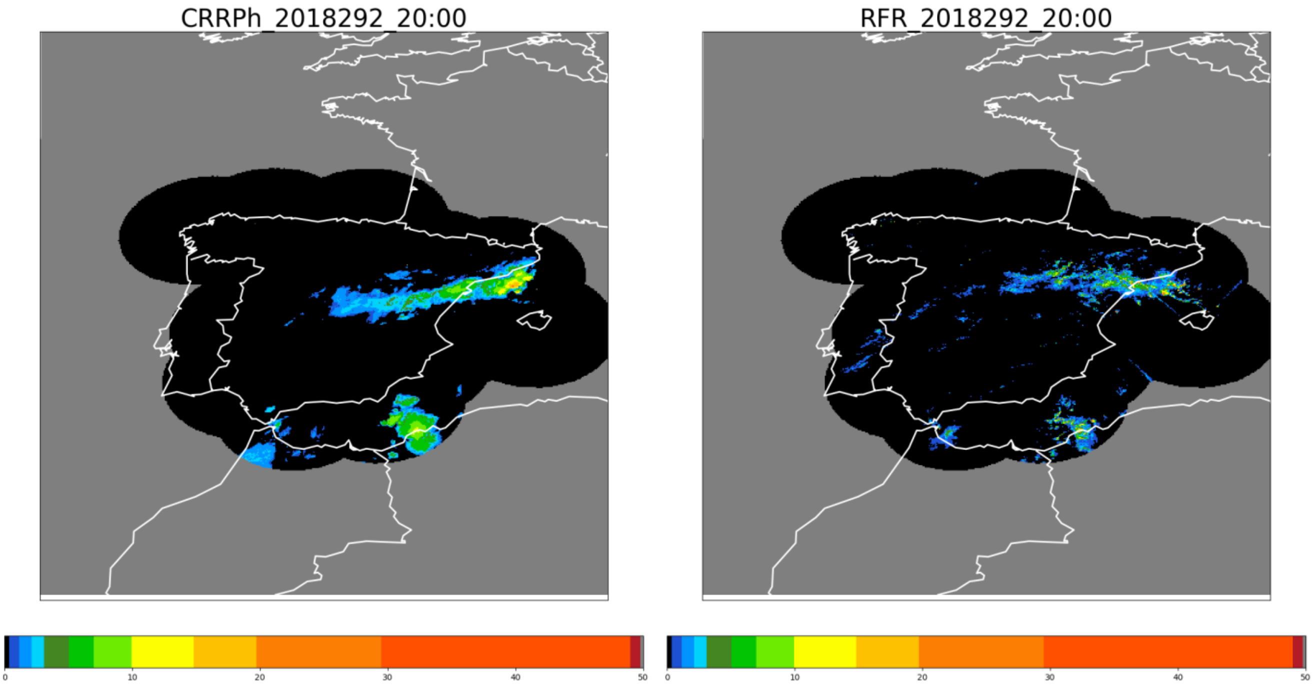



*Figure 21 CRRPh instantaneous rain rate intensities (mmh$^{-1}$) over Spain the 19$^{th}$ October 2018 at 20:00Z on the left side. Spanish composite radar (mmh$^{-1}$) on the right side.*

## CLOUD WATER PATH CORRECTION FACTOR

### DAY

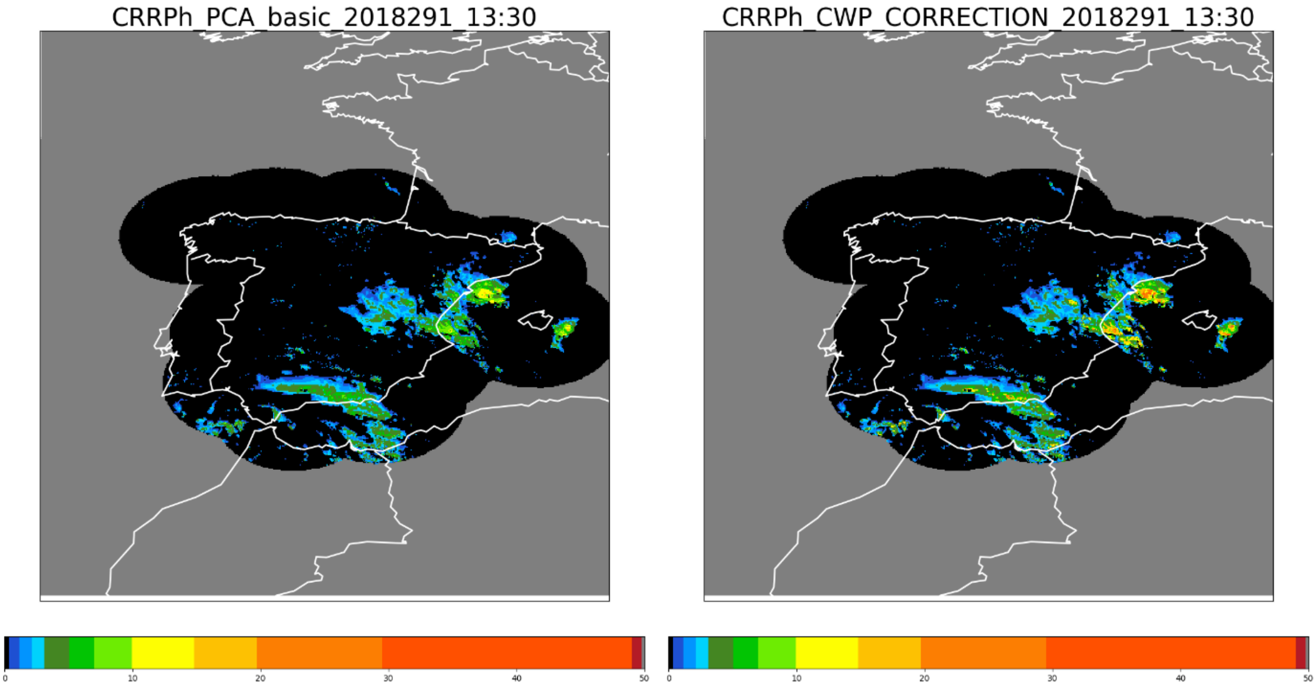



*Figure 22 CRRPh instantaneous rain rate intensities (mmh$^{-1}$) over Spain the 18$^{th}$ October 2018 at 13:30Z. Basic output on the left and having applied the CWP correction factor on the right.*

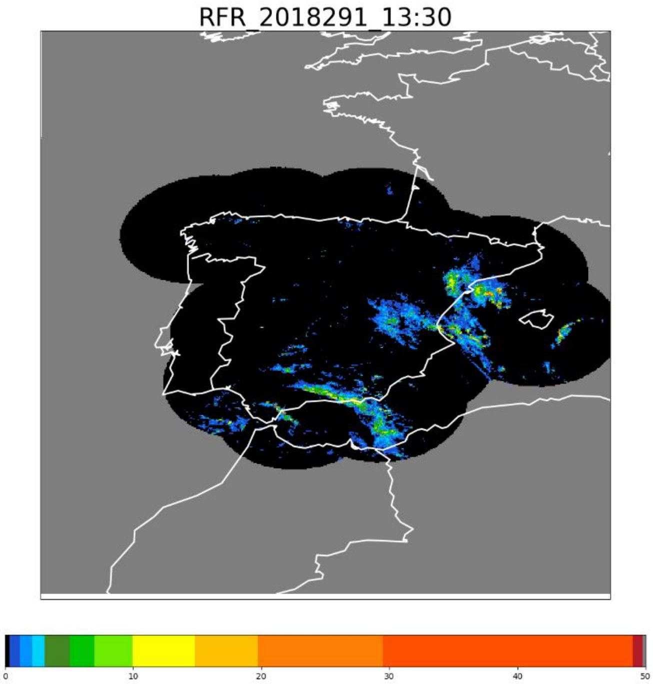



*Figure 23. Spanish composite radar (mmh$^{-1}$) over Spain the 18$^{th}$ October 2018 at 13:30Z*

Looking at Figure 23, it can be shown that several active convective cores with high rain intensities are approaching the cape of Tortosa. There is another line of cells a little bit more to the south, near the coastline of Valencia, also with high rain rates values. By comparing left side of Figure 22 and right side of the same figure with Figure 23 it can be noticed that the cloud water path correction factor enhances the CRRPh basic output and lead to higher rain intensity values in those convective nuclei. There are some other little convective cells near Balearic Islands, at both sides of the Strait of Gibraltar and offshore the Atlantic sea that the cloud water path correction factor highlights.

**NIGHT**

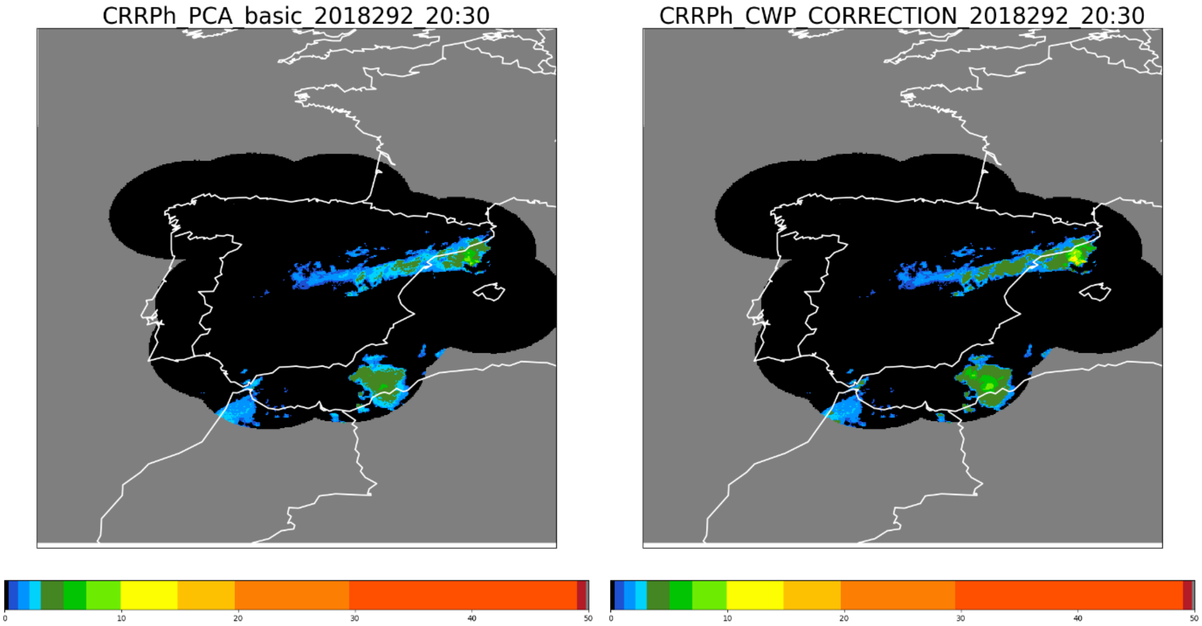



*Figure 24. CRRPh instantaneous rain rate intensities (mmh$^{-1}$) over Spain the 19$^{th}$ October 2018 at 20:30Z. Basic output on the left and having applied the CWP correction factor on the right.*

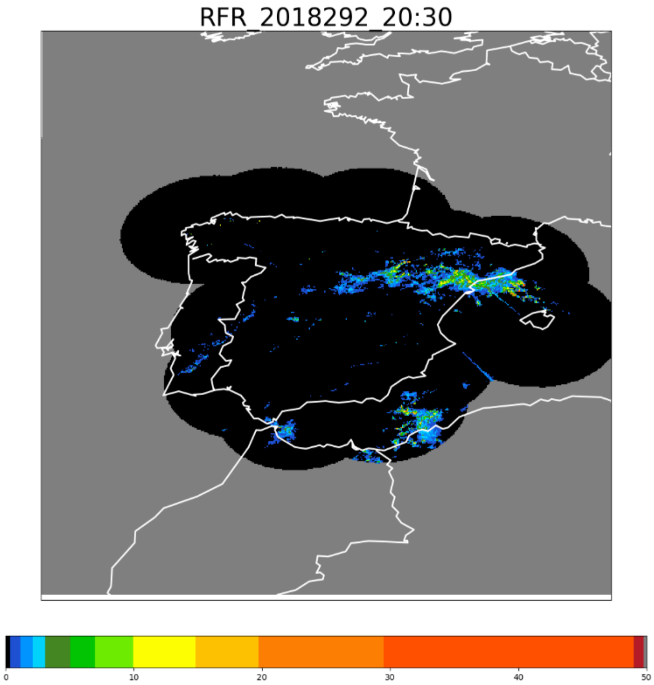



*Figure 25. Spanish composite radar (mmh$^{-1}$) over Spain the 19$^{th}$ October 2018 at 20:30Z*

In this second example, there was a very active convective nuclei near the coast of Barcelona and Tarragona and another one at the East of Algeria. Both of them were increased by the cloud water path correction factor.

## 4.5 ASSUMPTIONS AND LIMITATIONS

CRRPh and PCPh have been calibrated over the Iberian Peninsula with a 2015 database and does not cover the same output area (Europe). It is important to take into account that the relationship between what the GEO satellite can see and precipitation in the surface of the Earth is not direct. There is an uncertainty related to the estimation of probability of rain and rain intensities. This uncertainties turn into larger precipitation areas than the real precipitation areas shown by the Spanish radar. The way to control the extent of precipitation areas is done by establishing Cloud Water Path thresholds at day and night time. Rain intensities and probability of rain are not depicted if CWP is below 350gm$^{-2}$ and 250gm$^{-2}$ respectively.

The method is based on a PCA analysis keeping only the first two components that explain the 94.9 % of the variance. The more components the algorithm includes the more variance will be explained.

$VIS_{0.6}$ and CWP are simulated with infrared and water vapour channels with two PCA's that explained the 99.3% of the variance. Pseudo-$VIS_{0.6}$ and pseudo-CWP are artificial tools created to supply the lack of information at night. There is no way to verify if the simulation of the CWP and VIS06 are good at night time because there is not the sun. Checking the simulation at day time can be done by comparing the real value of the CWP against the CWP simulated, but that do not guarantee a perfect match at night. At day time, it can also be checked those pseudo-channels by simulating them and introduce those inputs in the CRRPh algorithm, but again this do not guarantee a perfect performance at night. Since these simulated inputs are not possible to be validated at night what it has been validated is the CRRPh final product.

PCPh and CRRPh use the same pseudo-variables at night. Taking for granted that the "training dataset is large and representative enough, the pseudo-variables can be common for the two algorithms.

In order to reduce some noise, the calibrating LUT has been smoothed in 3*3 boxes with the mean value.

So far CRRPh and PCPh are based on a LUT. It will be investigated the possibility to introduce neural networks or other artificial intelligence techniques ensuring that the final products take advantage of these novel techniques.

The product have been calibrated with the Spanish composite radar using the Marshal-Palmer relation, $Z=200R^{1.6}$, however there are other Z-R equations that distinguish between convective and stratiform rain.

During last summer, precipitating products were texted in the European Severe Storm Laboratory (ESSL). Some of the feedbacks it have been received from them are:

- Comparing the two microphysical products (CRRPh v2018 and the present one), participants considered the new one to be more accurate. That said, the older one was found to better highlight the convective cores.
- It was noted that microphysical products performed better in case of stratiform rainfall.

Since it has been noted this new microphysical version tends to underestimate high rainfall rates in convective cores, it has been enhanced the rainfall rates by multiplying the rainfall rate output by a correction factor in terms of the cloud water content. This way, the more content of water there is in the cloud, the bigger the correction factor is.

It is also important to remind that this product should be tuned again, along with the enhancing correction factor during the commissioning phase.

CRRPh and PCPh algorithms at night time depend on the reconstructed pseudo-VIS06 channel and the pseudo-CWP product that have been built extracting information from $IR_{8.7}$, $IR_{9.7}$, $IR_{108}$, $IR_{120}$, $IR_{134}$,$WV_{6.2}$ and $WV_{7.3}$ channels. Therefore, it may be considered that somehow the same information be used twice. In addition to that, Cloud Water Path and $VIS_{0.6}$ values processed throughout a year have large standard deviations compared to the mean value. It is a fact that the method to retrieve the psedo-variables and hence the CRRPh and PCPh products have uncertainties related to them. However, all these a priori inconveniences do not traduce in additional problems. The way to build CRRPh and PCPh products could have been tackled in two different ways:

a) Using two different algorithms. One for day time and another different one for night time. Day time algorithm can make use of $VIS_{0.6}$ and Cloud Water Path information since this information is available at day time. Night time algorithm only can be developed with infrared and water vapour channels, using the Principal Component Analysis technique.

b) Using only one algorithm for all instances, trying to unify all different scenarios in only one algorithm, giving a sense of continuity.

CRRPh output at night time for cases a) and b) have been previously studied and the results are quite similar, with case b) providing a better continuity between day and night time product. This is the reason why case b) has been finally chosen for the operational algorithm.

This is going to be illustrated by means of different examples:

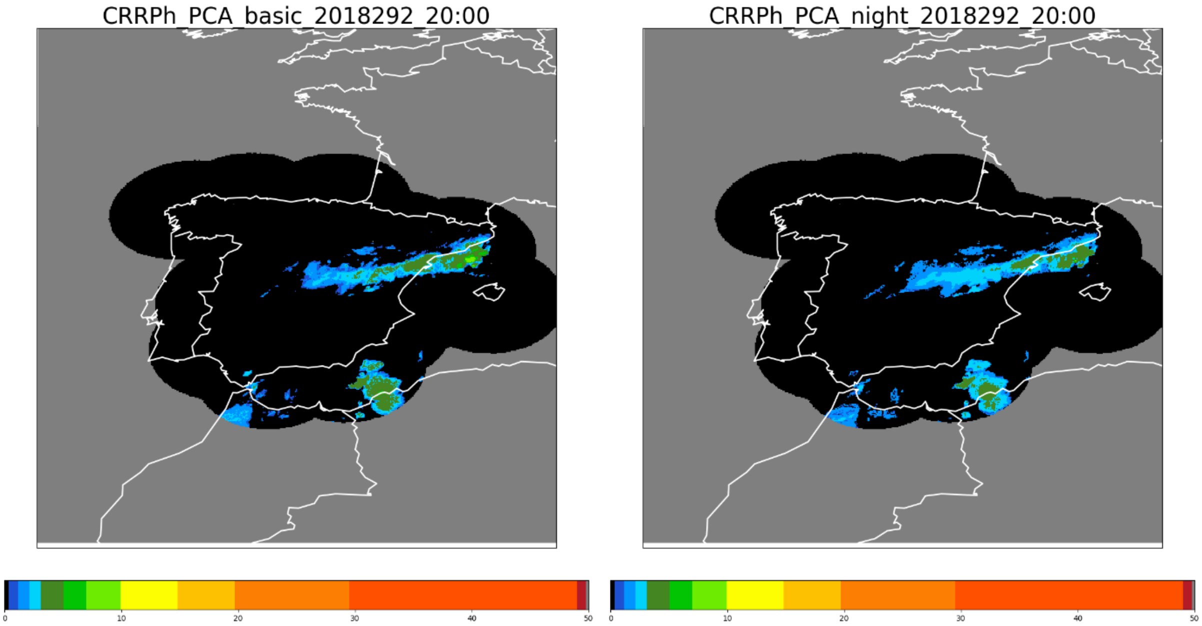



*Figure 26. CRRPh night algorithm instantaneous rain rate intensities (mmh$^{-1}$) basic output, over Spain the 19$^{th}$ October 2018 at 20:00Z with 9 inputs, including the pseudo-VIS06 and pseudo-CWP on the left. CRRPh night algorithm with 7 inputs, not including the additional pseudo-variables on the right.*

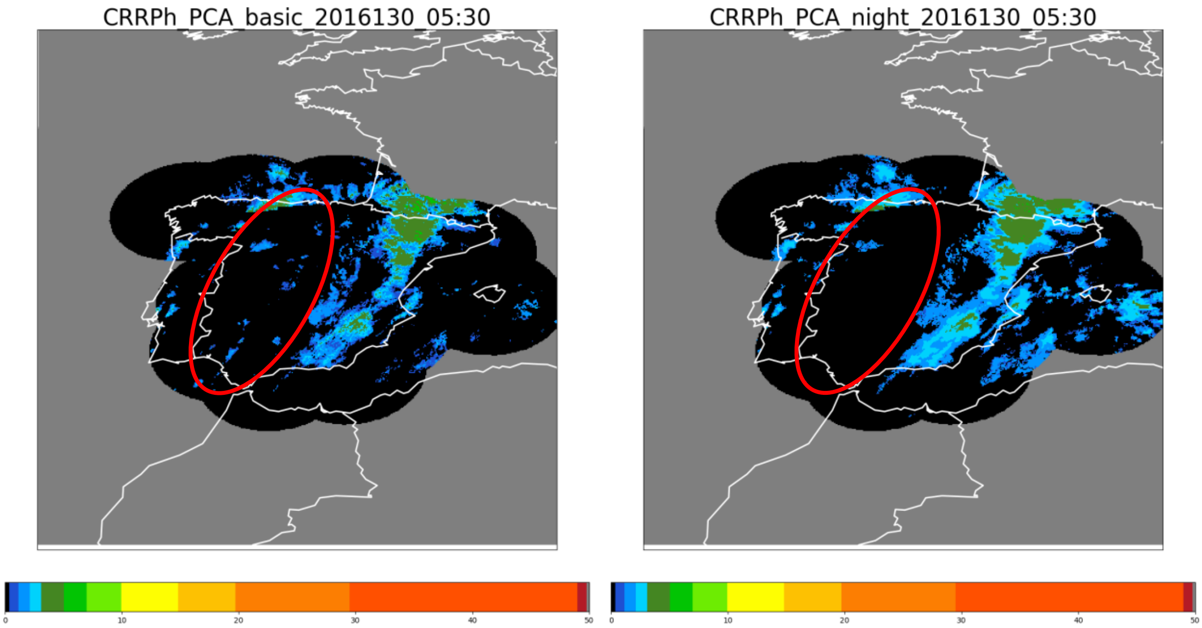



*Figure 27. CRRPh night algorithm, instantaneous rain rate intensities (mmh$^{-1}$) basic output, over Spain the 9$^{th}$ May 2016 at 05:30Z with 9 inputs, including the pseudo-VIS06 and pseudo-CWP on*

*the left. CRRPh night algorithm with 7 inputs, not including the additional pseudo-variables on the right.*

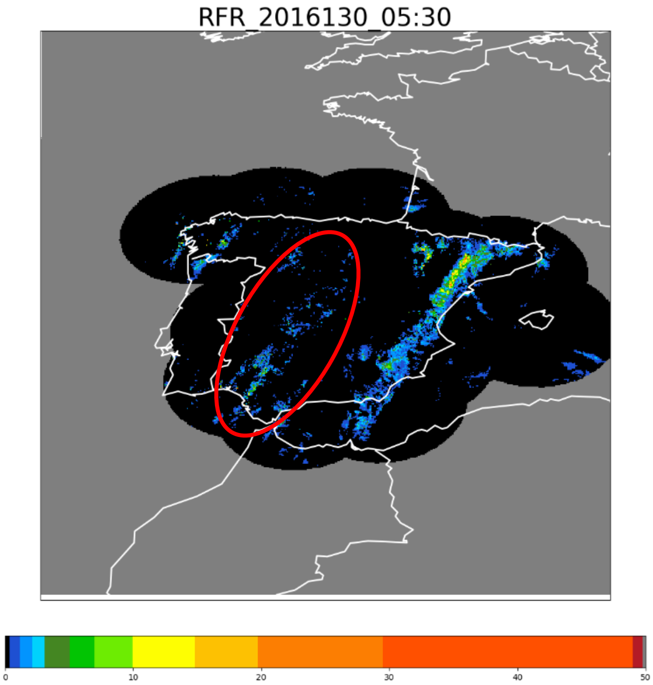



*Figure 28. Spanish composite radar (mmh$^{-1}$) over Spain the 9$^{th}$ May 2016 at 05:30Z*

It can be noticed in Figure 26, that both CRRPh algorithm with 9 inputs and with 7 inputs (not using the pseudo-variables) have a similar aspect. By using the pseudo-variables at night time more additional noise is not introduced to the signal. In this occasion, the Cloud Water Path correction has not been introduced on the CRRPh with night inputs so that a better comparison be made.

As far as Figures 27 and 28 are concerned, not only do the CRRPh product with 9 inputs not differ from the one with 7 inputs but also it better catch the precipitation pattern in the middle of the Spanish Peninsula marked inside an oval circle in red colour. It also can be noticed that precipitation area is larger at the North East of the Spanish Peninsula than the Spanish radar composite. It is probably due to the anvil of cumulonimbus that spread at high levels when reaching the limit of the Troposphere. This uncertainty is probably introduced by infrared channels.

## 4.6 CRRPH OUTPUT

The content of the CRRPh output is described in the *Data Output Format Document* [RD 8].

A summary is given below:

| **GLOBAL ATRIBUTES** | |
|---|---|
| **Product completeness** | Percentage of pixels within the region containing data |
| **Product quality** | Weighted mean of the quality of all pixels with data, using the following weights: 1: Good pixels, 0.5: Questionable quality |

<table>
<tr><th>Container</th><th>Content</th></tr>
<tr><td>crrph_intensity</td><td>
NWC GEO CTMP-CRR Convective Rainfall Intensity
<table>
<tr><th colspan="4">DAY</th></tr>
<tr><td>GEO-CMIC-PHASE INPUT</td><td>GEO-CMIC PHASE INPUT CLASS</td><td>COTT OR REFF FROM CMIC OR IR108, OR IR87, OR IR97, OR IR120, OR IR134, OR VIS06, OR WV62, OR WV73</td><td>CRR-PH OUTPUT</td></tr>
<tr><td rowspan="2">Liquid</td><td rowspan="2">1</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Ice</td><td rowspan="2">2</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Mixed</td><td rowspan="2">3</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td>Cloud-free</td><td>4</td><td>NOT APPLICABLE</td><td>0</td></tr>
<tr><td>Undefined</td><td>5</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
<tr><td>No data or corrupted data</td><td>FillValue</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
</table>
<table>
<tr><th colspan="4">NIGHT</th></tr>
<tr><td>GEO-CMIC-PHASE INPUT</td><td>GEO-CMIC PHASE INPUT CLASS</td><td>IR108, OR IR87, OR IR97, OR IR120, OR IR134, OR WV62, OR WV73</td><td>CRR-PH OUTPUT</td></tr>
<tr><td rowspan="2">Liquid</td><td rowspan="2">1</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Ice</td><td rowspan="2">2</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Mixed</td><td rowspan="2">3</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>crrph_intensity(mm/h) = scale_factor * counts + add_offset</td></tr>
<tr><td>Cloud-free</td><td>4</td><td>NOT APPLICABLE</td><td>0</td></tr>
<tr><td>Undefined</td><td>5</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
<tr><td>No data or corrupted data</td><td>FillValue</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
</table>
crrph_intensity(mm/h) = scale_factor * counts + add_offset<br>
where:<br>
scale_factor = 0.1<br>
add_offset = 0.0
</td></tr>
<tr><td>crrph_accum</td><td>
NWC GEO CTMP-CRR Convective Hourly Rainfall Accumulation<br>
crrph_accum(mm) = scale_factor * counts + add_offset<br>
where:<br>
scale_factor = 0.1<br>
add_offset = 0.0
</td></tr>
</table>

| Container | Content |
|---|---|
| crrph_status_flag | 10 bits indicating<br>Data Availability:<br>Bit 0: $R_{eff}$ or COT not computed (out of cloud, night time, phase not defined)<br>Bit 1: Phase not computed or undefined<br>Bit 2: IR band missing (used in parallax correction)<br>Applied Correction:<br>Bit 3: Parallax correction applied<br>Use of optional data:<br>Bit 6: Not used<br>Bit 8: crrph_intensity was a hole because of the parallax correction, and then was filled by the median filter<br>Other information<br>Bit 8: crrph_intensity was a hole because of the parallax correction, and then was filled by the median filter<br>Bit 9, 10, 11:Use of bands for accumulation<br>1: All required bands were available<br>2: One previous CRRPh band is missing<br>3: At least two previous CRRPh bands are missing (no consecutive)<br>4: At least two previous CRRPh bands are missing (some are consecutive)<br>Bit 12: Accumulation quality flag. Set to 1 if:<br>not all crr-ph values are available to perform the accumulation,<br>OR<br>any of the crrph_intensity values was set to 0 due to filtering process<br>OR<br>Any of the crrph_intensity values was a hole because parallax correction<br><br>Bit 13: Accumulation illumination flag:<br>1: Accumulation computed only with day algorithm.<br>2: Accumulation computed only with night algorithm<br>3: Accumulation computed with mixed algorithms. |

### *Geophysical Conditions*

| Field | Type | Description |
|---|---|---|
| Space | Flag | Set to 1 for space pixels |
| Illumination | Parameter | Defines the illumination condition<br><br>0: N/A (space pixel)<br>1: Night<br>2: Day<br>3: Twilight |
| Sunglint | Flag | Set to 1 if Sunglint |
| Land_Sea | Parameter | 0: N/A (space pixel)<br>1: Land<br>2: Sea<br>3: Coast |

***Processing Conditions***

| Field | Type | Description |
|---|---|---|
| Satellite_input_data | Parameter | Describes the Satellite input data status<br><br>0: N/A (space pixel)<br>1: All satellite data are available<br>2: At least one useful satellite channel is missing<br>3: At least one mandatory satellite channel is missing |
| NWP_input_data | Parameter | Describes the NWP input data status<br><br>0: N/A (space pixel or NWP data not used)<br>1: All NWP data are available<br>2: At least one useful NWP field is missing<br>3: At least one mandatory NWP field is missing |
| Product_input_data | Parameter | Describes the Product input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All input Product data are available<br>2: At least one useful input Product is missing<br>3: At least one mandatory input Product is missing |
| Auxiliary_input_data | Parameter | Describes the Auxiliary input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All Auxiliary data are available<br>2: At least one useful Auxiliary field is missing<br>3: At least one mandatory Auxiliary field is missing |

***Quality***

| Field | Type | Description |
|---|---|---|
| Nodata | Flag | Set to 1 if pixel is NODATA |
| Internal_consistency | Flag | Set to 1 if an internal consistency check has been performed. Internal consistency checks will be based in the comparison of the retrieved meteorological parameter with physical limits, climatic limits, neighbouring data, NWP data, etc. |
| Temporal_consistency | Flag | Set to 1 if a temporal consistency check has been performed<br>Temporal consistency checks will be based in the comparison of the retrieved meteorological parameters with data obtained in previous slots. |
| Quality | Parameter | Retrieval Quality<br>0: N/A (no data)<br>1: Good<br>2: Questionable<br>3: Bad (REMOVE)<br>4: Interpolated |

## 4.7 REFERENCES

PRINCIPAL COMPONENT ANALYSIS: A BEGINNER'S GUIDE - I. Introduction and application By IAN T. JOLLIFFE Institute of Mathematics, University of Kent, Canterbury.

Judith A. Curry Christopher D. Ardeel Lin Tian: Liquid water content and precipitation characteristics of stratiform clouds as inferred from satellite microwave measurements. Journal of Geophysical Research: Atmospheres, Volume 95, Issue D10, pp. 16,659-16,671 DOI: 10.1029/JD095iD10p16659

Gutierrez, J. M. and Aguado, F.: Quality image for the Spanish Radar National Composite, Proceedings of ERAD 2006, 318-320.

Nauss, T. and Kokhanovsky, A. A., 2007: Advances in GeosciencesAssignment of rainfall confidence values using multispectral satellite data at mid-latitudes: first results, Advances in Geosciences, 10, 99–102.

Roebeling, R. and Holleman, I., 2009: SEVIRI rainfall retrieval andvalidation using weather radar observations, J Geophys Res,114, D21 202, doi: 10.1029/2009JD012 102.

Rosenfeld, D. and G. Gutman, 1994. Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data, Atmos. Res., 34, 259–283, doi:10.1016/0169-8095(94)90096-5.

Rosenfeld, D., William L. Woodley, Amit Lerner, Guy Kelman, Daniel T. Lindsey, 2008. Satellite detection of severe convective storms by their retrieved vertical profiles of cloud particle effective radius and thermodynamic phase. J. Geophys. Res. D4, 113.

Tapia, A., Smith, J. A., Dixon, M., 1998: Estimation of Convective Rainfall from Lightning Observations, Bull. American Meteorological Society, Vol. 37, pp. 1497-1509.

# 5. DESCRIPTION OF PRECIPITATING CLOUDS FROM CLOUD PHYSICAL PROPERTIES (PCPH)

## 5.1 PRECIPITATING CLOUDS FROM CLOUD PHYSICAL PROPERTIES (PCPH) OVERVIEW

### 5.1.1 Inputs

**SATELLITE**:

- $IR_{8.7}$, $IR_{9.7}$, $IR_{10.8}$,$IR_{12.0}$,$IR_{13.4}$ (Brightness temperature)
- $VIS_{0.6}$ (Normalized reflectance and corrected with Sun distance)
- $WV_{6.2}$,$WV_{7.3}$ (Brightness temperature)

**NWC/GEO software:**

- GEO-CMIC ($CMIC_{COT}$,$CMIC_{REFF}$,$CMIC_{Phase}$)

**Numerical model:**

Temperature at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa.

- Geopotential at 1000, 925, 850, 700, 500, 400, 300, 250 and 200 hPa.

This information is used by default for parallax correction along with $IR_{108}$. In case of lack of NWP parameters parallax correction will be run using a climatic profile.

**Ancillary data sets:**

- Climatic profile is necessary as a backup for Parallax correction in case NWP is not available. This information is included in the software package and is located in the $SAFNWC/import/Aux_data directory.

**Model configuration file for PCPh**:

PCPh model configuration file contains configurable system parameters in the product generation process related to algorithm thresholds, ancillary datasets, numerical model data, etc. The complete list of these parameters and the explanation of the most useful ones is available in the User Manual for the Precipitation Product Processors of the NWC/GEO [*RD 5*].

#### *5.1.1.1 Inputs pre-processing*

The illumination conditions based on the solar zenith angle play a role in the day/night pixel transition. The DAY_NIGHT_ZEN_THRESHOLD is set to 70 degrees in the configuration file by default [RD 5].

**C) DAY**

CWP=2/3*COT*REFF ($gm^{-2}$)

If COT or REFF are not available (Fill Value) CWP is set to NODATA

**D) NIGHT**

Since the PCPh algorithm uses the same inputs during the whole day and $CMIC_{COT}$, and $CMIC_{REFF}$ are not available at night , $VIS_{0.6}$ normalized and CWP are simulated. See section number 6.

#### 5.1.1.2 Inputs pre-checking

PCPh is set to NODATA if satellite inputs are not available, CMIC Phase is undefined or is not available or $CMIC_{COT}$ and $CMIC_{REFF}$ derived from the NWC/GEO software are Fill Values.

PCPh is set to zero in the following cases:

- CMIC Phase output indicates cloud free.
- CWP < 250 $gm^{-2}$ for both day and night time.
- $VIS_{0.6}$ normalized reflectivity < 30 for both day and night.

The threshold value for the Cloud Water Path to establish rainy/no rainy pixels is especially important to detect shallow rain. Different authors have established different thresholds, for example Goddard profiling algorithm (GPROF) has employed the rain/no-rain threshold value of 0.3 kg $m^{-2}$. Judith A. Curry, Christopher D. Ardeel and Lin Tian established a threshold of 350 $gm^{-2}$ for middle clouds and 500 $gm^{-2}$ for the onset of precipitation in low clouds.

In this case the threshold to detect rain is slightly lower than in the case of CRRPh. It could have been kept the same criterion, with low impact in summer season, but improving during the rest of the year.

Cloud Water Path threshold to delimit the probability of rain and CWP threshold to depict rain intensities have been empirically obtained. It would have been nice to have the same threshold for both products, because it would have given a sense of continuity. CRRPh and PCPh are different products. The best CWP threshold found for PCPh is 250$gm^{-2}$ and 350$gm^{-2}$ for CRPPh.

### 5.1.2 Normalization

Normalizing every pixel consists of subtracting a fixed value (mean value) from every input channel and divide by another fixed value (standard deviation)

Normalized $value_{CHANNEL}$=(Pixel value $_{CHANNEL}$ –Mean value $_{CHANNEL}$)/Standard Deviation $_{CHANNEL}$

| Channel | Mean value | Standard Deviation |
|---|---|---|
| CWP ($gm^{-2}$) | 520.48193446 | 780.39100519 |
| IR108 (K) | 242.76576216 | 17.20011957 |
| IR120 (K) | 241.32026712 | 16.7090834 |
| IR134 (K) | 234.68889863 | 11.11713272 |
| IR87 (K) | 242.86897636 | 16.73599863 |
| IR97 (K) | 234.46443828 | 9.0608213 |
| VIS06 (K) | 56.93809613 | 17.59992087 |
| WV62 (K) | 227.46449281 | 5.44519917 |
| WV73 (K) | 234.96183994 | 9.87543324 |

*Table 10. Normalizing Parameters for PCPh*

### 5.1.3 Projections

Every pixel should be processed as follows:

$P_1$=CWP normalized $*v_{11}+IR_{10.8}$ normalized $*v_{12}+IR120*v_{13}+.......+WV_{7.3}$normalized$*v_{19}$

$P_2$=CWP normalized $*v_{21}+IR_{10.8}$ normalized $*v_{22}+IR120*v_{23}+.......+WV_{7.3}$normalized$*v_{29}$

| **Eigenvectors** | | | |
|---|---|---|---|
| $v_{11}$ | 0.16196726 | $v_{21}$ | -0.79885475 |
| $v_{12}$ | -0.36737917 | $v_{22}$ | -0.06545106 |
| $v_{13}$ | -0.3668939 | $v_{23}$ | -0.0886929 |
| $v_{14}$ | -0.36577753 | $v_{24}$ | -0.1257116 |
| $v_{15}$ | -0.36739541 | $v_{25}$ | -0.04760894 |
| $v_{16}$ | -0.35333574 | $v_{26}$ | -0.0402052 |
| $v_{17}$ | 0.27229012 | $v_{27}$ | -0.50034629 |
| $v_{18}$ | -0.36048641 | $v_{28}$ | -0.14551587 |
| $v_{19}$ | -0.32630803 | $v_{29}$ | -0.24181205 |

*Table 11. Eigenvectors to compute the projections*

### 5.1.4 Look up calibration table

The calibration Look up table has the following characteristics:

- 2Dimensions with 200 bins.
- X Axis goes from -10 to 10, Y Axis goes from -10 to 10, both with 200 divisions.
- Central value of the first division is -9.95 along X axis with a 0.1 increase. Highest value is 9.95.

- Central value of the first division is -9.95 along Y axis with a 0.1 increase. Highest value is 9.95.

Once P1 and P2 have been calculated for every pixel PCPh is computed as follows:

PCPh=LUT(x=p1,y=p2)

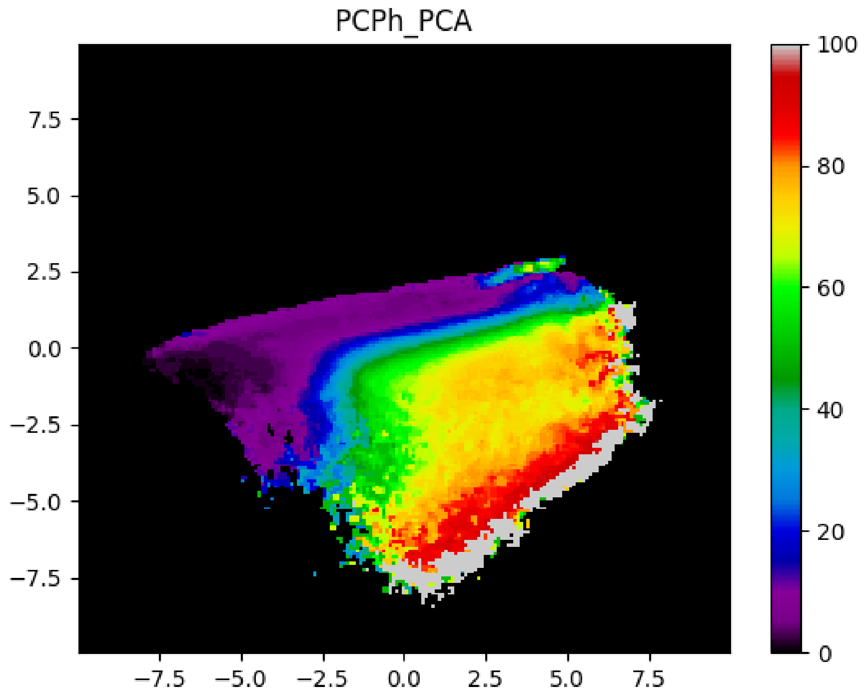



*Figure 29. 200*200 calibration LUT for PCPh*

Z axis is the colour palette and represents the probability of rain.

Radar pixels with rain rates greater than or equal to 0.2 mm/h have been considered as rainy.

Once the two projections have been calculated, it is necessary to associate them with a probability of rain. Then, for every pair of points (p1,p2) the proportion of radar rainy pixels is evaluated by dividing the number of rainy pixels among all the radar pixels.

The LUT has been smoothed in 3*3 boxes with a median filter to reduce some noise.

## 5.2 Example of visualization

### DAY

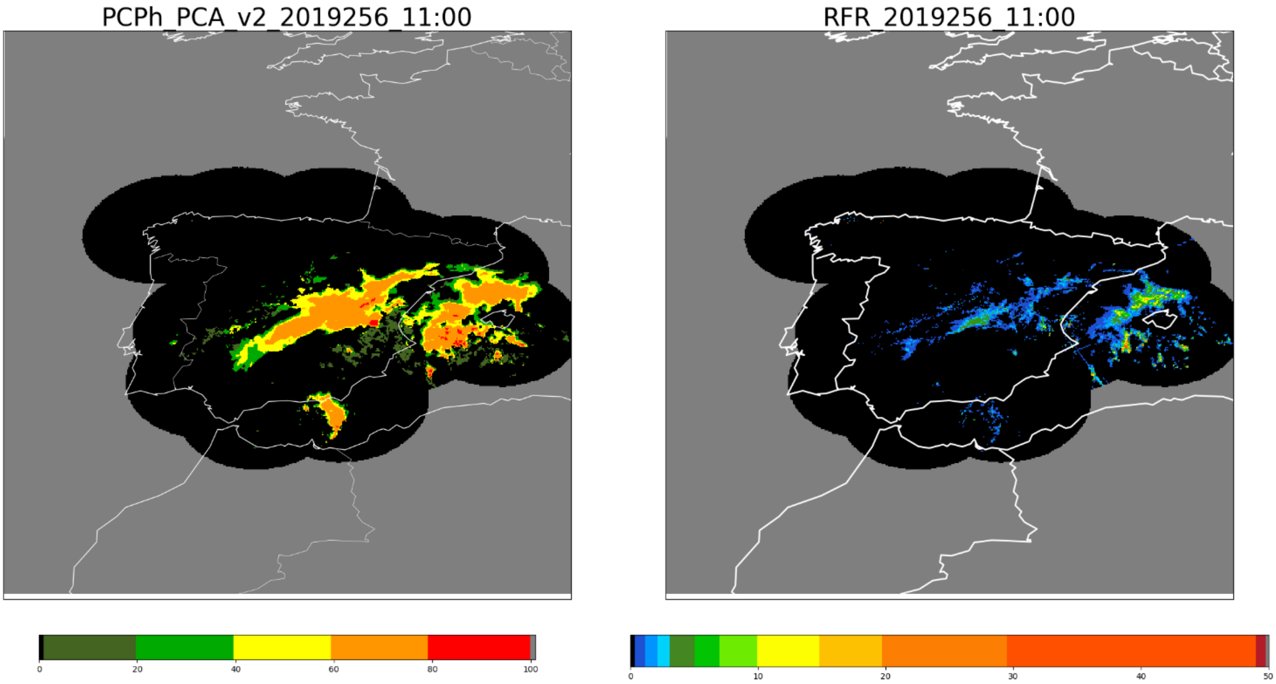



*Figure 30. PCPh Probability of precipitation (%) over Spain the 13th September 2019 at 11:00Z on the left side. Spanish composite radar (mmh$^{-1}$) on the right side.*

**NIGHT**

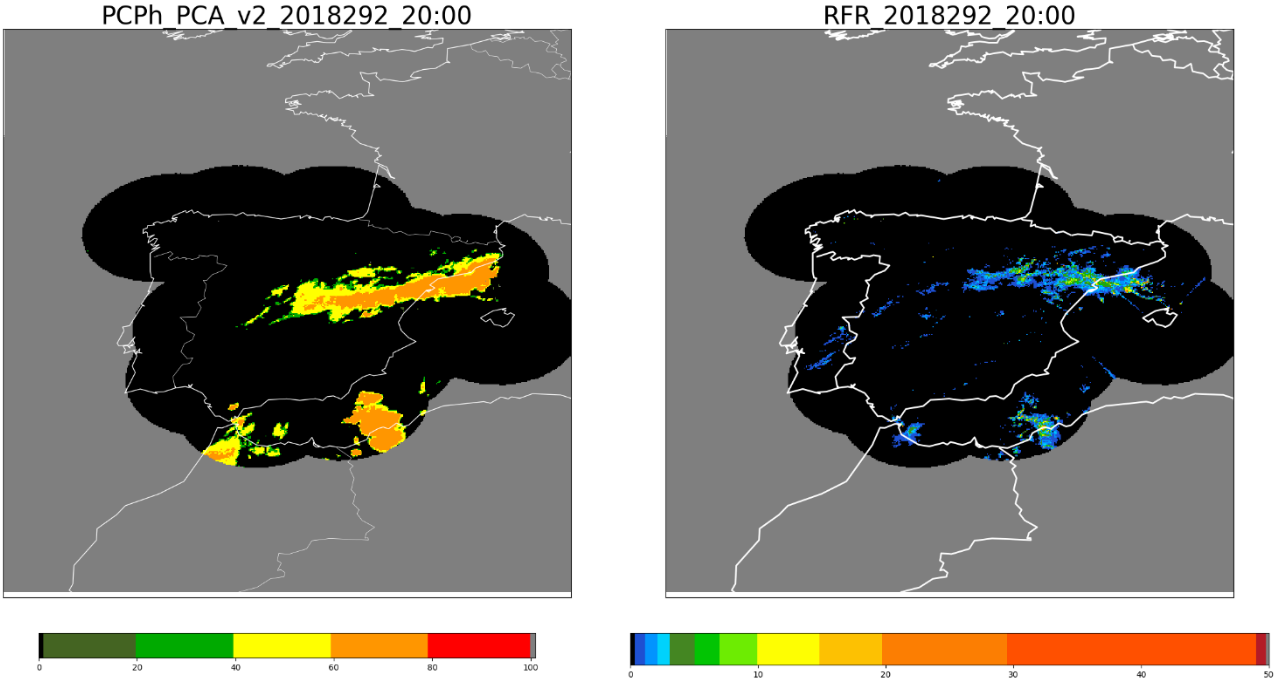



*Figure 31. Probability of precipitation (%) over Spain the 18th October 2018 at 20:00Z on the left side. Spanish composite radar (mmh$^{-1}$) on the right side.*

It can be noticed that in both day and night conditions, the probability of rain depicted in figures 30 and 31 is larger than the rainy area provided by the Spanish radar composite. CWP threshold that has been empirically established have a direct impact on the PCPh output. By choosing a higher CWP threshold (bigger than 250gm$^{-2}$) the precipitating area will have been reduced. By contrary in other situations will have not depicted probability of rain in areas where the Spanish rain radar provide with rain. A general balance must be taken.

## 5.3 Assumptions and limitations

Section 4.5 also applies to this section.

## 5.4 PCPh output

The content of the PCPh output is described in the *Data Output Format Document* [RD 9].

A summary is given below:

| **GLOBAL ATRIBUTES** | |
|---|---|
| **Product completeness** | Percentage of pixels within the region containing data |
| **Product quality** | Weighted mean of the quality of all pixels with data, using the following weights: 1: Good pixels, 0.5: Questionable quality |

<table>
<tr><th>Container</th><th>Content</th></tr>
<tr><td>pcph</td><td>NWC GEO CTMP-CRR Convective Rainfall Intensity
<table>
<tr><th colspan="4">DAY</th></tr>
<tr><td>GEO-CMIC-PHASE INPUT</td><td>GEO-CMIC PHASE INPUT CLASS</td><td>COT OR REFF FROM CMIC OR IR108, OR IR87, OR IR97, OR IR120, OR IR134, OR VIS06, OR WV62, OR WV73</td><td>PC-PH OUTPUT</td></tr>
<tr><td rowspan="2">Liquid</td><td rowspan="2">1</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Ice</td><td rowspan="2">2</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Mixed</td><td rowspan="2">3</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td>Cloud-free</td><td>4</td><td>NOT APPLICABLE</td><td>0</td></tr>
<tr><td>Undefined</td><td>5</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
<tr><td>No data or corrupted data</td><td>FillValue</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
</table>
<table>
<tr><th colspan="4">NIGHT</th></tr>
<tr><td>GEO-CMIC-PHASE INPUT</td><td>GEO-CMIC PHASE INPUT CLASS</td><td>IR108, OR IR87, OR IR97, OR IR120, OR IR134, OR WV62, OR WV73</td><td>PC-PH OUTPUT</td></tr>
<tr><td rowspan="2">Liquid</td><td rowspan="2">1</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Ice</td><td rowspan="2">2</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td rowspan="2">Mixed</td><td rowspan="2">3</td><td>NO DATA</td><td>NO DATA</td></tr>
<tr><td>DATA AVAILABLE</td><td>pcph(%) = scale_factor * counts + add_offset</td></tr>
<tr><td>Cloud-free</td><td>4</td><td>NOT APPLICABLE</td><td>0</td></tr>
<tr><td>Undefined</td><td>5</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
<tr><td>No data or corrupted data</td><td>FillValue</td><td>NOT APPLICABLE</td><td>NO DATA</td></tr>
</table>
pcph(%) = scale_factor * counts + add_offset<br>
where:<br>
scale_factor = 1.0<br>
add_offset = 0.0</td></tr>
<tr><td>pcph_status_flag</td><td>5 bits indicating<br>
Data Availability:<br>
Bit 0: $R_{eff}$ or COT not computed (out of cloud, night time or undefined phase)<br>
Bit 1: Phase not computed or undefined<br>
Bit 2: IR band missing (used in parallax correction)<br>
Applied Correction:<br>
Bit 3: Parallax correction applied<br>
Other information<br>
Bit 8: pcph_intensity was a hole because of the parallax correction, and then was filled by the median filter</td></tr>
</table>

***Geophysical Conditions***

| Field | Type | Description |
|---|---|---|
| Space | Flag | Set to 1 for space pixels |
| Illumination | Parameter | Defines the illumination condition<br><br>0: N/A (space pixel)<br>1: Night<br>2: Day<br>3: Twilight |
| Sunglint | Flag | Set to 1 if Sunglint |
| Land_Sea | Parameter | 0: N/A (space pixel)<br>1: Land<br>2: Sea<br>3: Coast |

***Processing Conditions***

| Field | Type | Description |
|---|---|---|
| Satellite_input_data | Parameter | Describes the Satellite input data status<br><br>0: N/A (space pixel)<br>1: All satellite data are available<br>2: At least one useful satellite channel is missing<br>3: At least one mandatory satellite channel is missing |
| NWP_input_data | Parameter | Describes the NWP input data status<br><br>0: N/A (space pixel or NWP data not used)<br>1: All NWP data are available<br>2: At least one useful NWP field is missing<br>3: At least one mandatory NWP field is missing |
| Product_input_data | Parameter | Describes the Product input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All input Product data are available<br>2: At least one useful input Product is missing<br>3: At least one mandatory input Product is missing |
| Auxiliary_input_data | Parameter | Describes the Auxiliary input data status<br><br>0: N/A (space pixel or Auxiliary data not used)<br>1: All Auxiliary data are available<br>2: At least one useful Auxiliary field is missing<br>3: At least one mandatory Auxiliary field is missing |

***Quality***

| Field | Type | Description |
|---|---|---|
| Nodata | Flag | Set to 1 if pixel is NODATA |
| Internal_consistency | Flag | Set to 1 if an internal consistency check has been performed. Internal consistency checks will be based in the comparison of the retrieved meteorological parameter with physical limits, climatic limits, neighbouring data, NWP data, etc. |
| Temporal_consistency | Flag | Set to 1 if a temporal consistency check has been performed Temporal consistency checks will be based in the comparison of the retrieved meteorological parameters with data obtained in previous slots. |
| Quality | Parameter | Retrieval Quality<br>0: N/A (no data)<br>1: Good<br>2: Questionable<br>3: Bad (REMOVE)<br>4: Interpolated |

# 6. CWP AND $VIS_{0.6}$ NORMALIZED SIMULATION IN NIGHT MODE TO PCPH AND CRRPH COMPUTATION

Since CWP and $VIS_{0.6}$ are not available at night time and they are compulsory inputs to compute both PCPh and CRRPh, it is necessary to simulate them to be used at night time.

The method to generate CWP and $VIS_{0.6}$ is based on a Principal Component Analysis.

As at night there are only infrared and water vapour channels, those will be the inputs two train our dataset.

It has been used $IR_{8.7}$, $IR_{9.7}$, $IR_{10.8}$, $IR_{12.0}$, $IR_{13.4}$ and $WV_{6.2}$ and $WV_{7.3}$ SEVIRI channels between 10Z and 15Z as dataset to build the Covariance matrix. It has been kept the first two principal components that explain a 99 % of the variance.

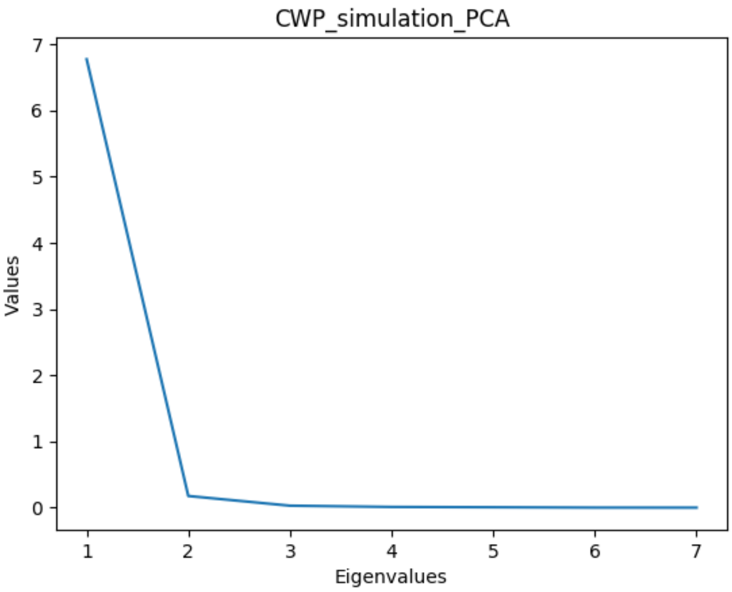



*Figure 32. Eigenvalues for CWP and $VIS_{0.6}$*

## 6.1 INPUTS

| $IR_{8.7}$ µm | $IR_{9.7}$ µm | $IR_{10.8}$ µm | $IR_{12.0}$ µm | $IR_{13.4}$ µm | $VW_{6.2}$ µm | $WV_{7.3}$ µm |
|---|---|---|---|---|---|---|

Brightness temperature of these channels have been parallax corrected.

## 6.2 NORMALIZATION

Normalizing every pixel consists of subtracting a fixed value (mean value) from every SEVIRI channel and divide by another fixed value (standard deviation)

Normalized value $_{\text{CHANNEL}}$=(Pixel value $_{\text{CHANNEL}}$ –Mean value $_{\text{CHANNEL}}$)/Standard Deviation $_{\text{CHANNEL}}$

| Channel | Mean value | Standard Deviation |
|---|---|---|
| IR108 | 242.03643436 | 21.27664634 |
| IR120 | 240.61806287 | 20.58345066 |
| IR134 | 234.23087442 | 13.51909372 |
| IR87 | 242.08655814 | 20.78070428 |
| IR97 | 235.34225756 | 10.67359157 |
| WV62 | 227.09770767 | 6.63554842 |
| WV73 | 233.72949906 | 12.15742079 |

*Table 12. Normalizing Parameters for $VIS_{0.6}$ and CWP*

## 6.3 PROJECTIONS

Every pixel should be processed as follows:

$P_1=IR_{10.8}$ normalized $*v_{11}+IR_{12.0}$ normalized $*v_{12}+IR_{13.4}*v_{13}+.......+WV_{7.3}$normalized$*v_{19}$

$P_2=IR_{10.8}$ normalized $*v_{21}+IR_{12.0}$ normalized $*v_{22}+IR_{13.4}*v_{23}+.......+WV_{7.3}$normalized$*v_{29}$

| Eigenvectors | | | |
|---|---|---|---|
| $\mathbf{v_{11}}$ | -0.3818856 | $\mathbf{v_{21}}$ | 0.22521461 |
| $\mathbf{v_{12}}$ | -0.38244994 | $\mathbf{v_{22}}$ | 0.17259767 |
| $\mathbf{v_{13}}$ | -0.38326863 | $\mathbf{v_{23}}$ | 0.01525338 |
| $\mathbf{v_{14}}$ | -0.38146901 | $\mathbf{v_{24}}$ | 0.24101417 |
| $\mathbf{v_{15}}$ | -0.37517376 | $\mathbf{v_{25}}$ | 0.38387953 |
| $\mathbf{v_{16}}$ | -0.3798613 | $\mathbf{v_{26}}$ | -0.27833609 |
| $\mathbf{v_{17}}$ | -0.36115374 | $\mathbf{v_{27}}$ | -0.79770511 |

*Table 13. Eigenvectors to compute the projections*

## 6.4 LOOK UP CALIBRATION TABLE

The calibration Look up table has the following characteristics:

- 2Dimensions with 200 bins.
- X Axis goes from -8 to 8, Y Axis goes from -3.5 to 3.5, both with 200 divisions.
- Central value of the first division is -7.96 along X axis with a 0.8 increase. Highest value is 7.96.
- Central value of the first division is -3.4825 along Y axis with a 0.035 increase. Highest value is 3.4825.

Once P1 and P2 have been calculated for every pixel, $VIS_{0.6}$ and CWP are computed as follows:

$VIS_{0.6}$=LUT(x=p1,y=p2) according to Figure 33

CWP=LUT(x=p1,y=p2) according to Figure 34

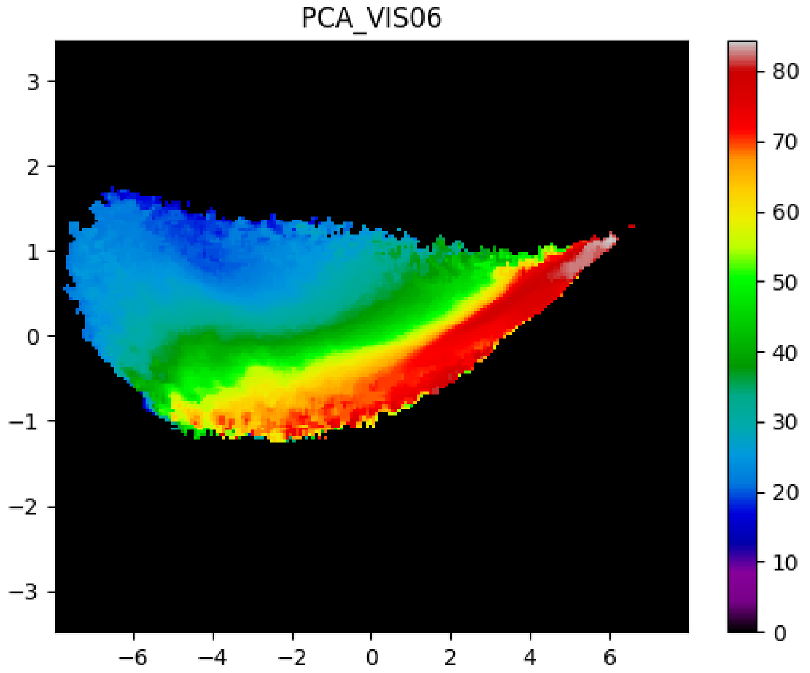



*Figure 33 . 200*200 LUT for $VIS_{0.6}$*

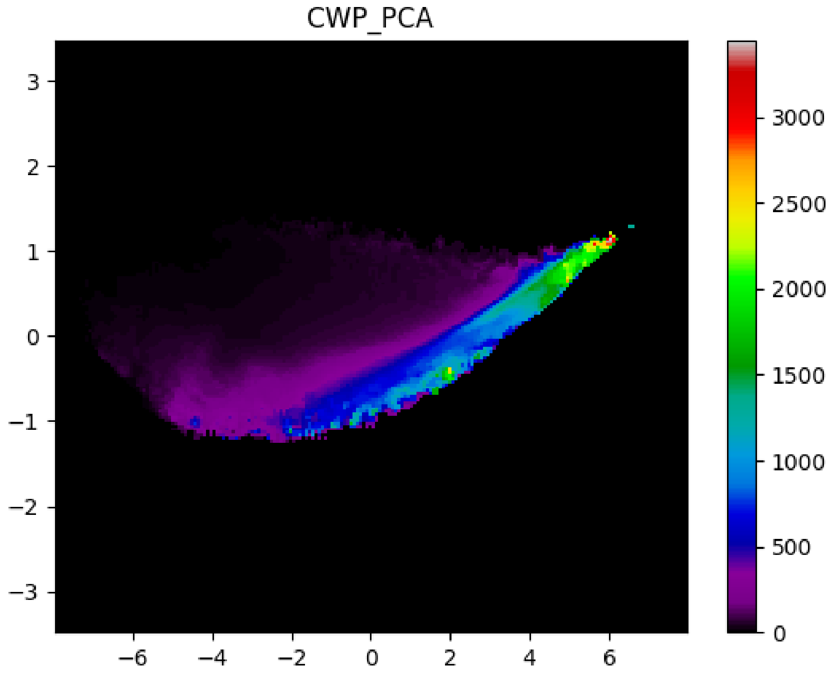



*Figure 34. 200*200 LUT for CWP*

Z axis is the colour palette and it represents reflectance in Figure 33 and Cloud Water Path ($gm^{-2}$) in Figure 34.

As far as $VIS_{0.6}$ is concerned for every pair of points (p1,p2) the mean value of the $VIS_{0.6}$ reflectance is computed for the 2015 dataset. In the case of the CWP the mean value of the 2015 dataset it has been also computed. The main reason for choosing the mean value and not a Percentile with a higher value are the False Alarms. Thinking on Hits it would be a better idea to choose a 75% Percentile or even higher because this election would have led to higher values of CRRPh rain rates and higher values of probability of rain (PCPh). CRRPh is able to detect early stages of developing thunderstorms, but it tends to slightly overestimate the precipitating area. However, when it comes to visually check False Alarms, it seems that it is less cumbersome giving a False Alarm with a lower rain intensity value than a False Alarm with high rain rates.

Both LUT have been smoothed in 3*3 boxes with a median filter to reduce some noise.

$VIS_{0.6}$ values vary from zero to 100 % of reflectance and values from CWP range from zero to values slightly higher to 7000 $gm^{-2}$. The statistical chosen has been the mean value. That is the reason why, as it can be noticed in the colour palette, $VIS_{0.6}$ simulated reflectance do not reach the 100 % of reflectance and CWP only reaches values about 4000 $gm^{-2}$.This election have an impact on the rain intensities and probabilities of rain at night, that are in general lower. However, adopting this criterion make us obtain better results on average computing a whole year in terms of depicting the precipitation area.

## 6.5 EXAMPLE OF VISUALIZATION

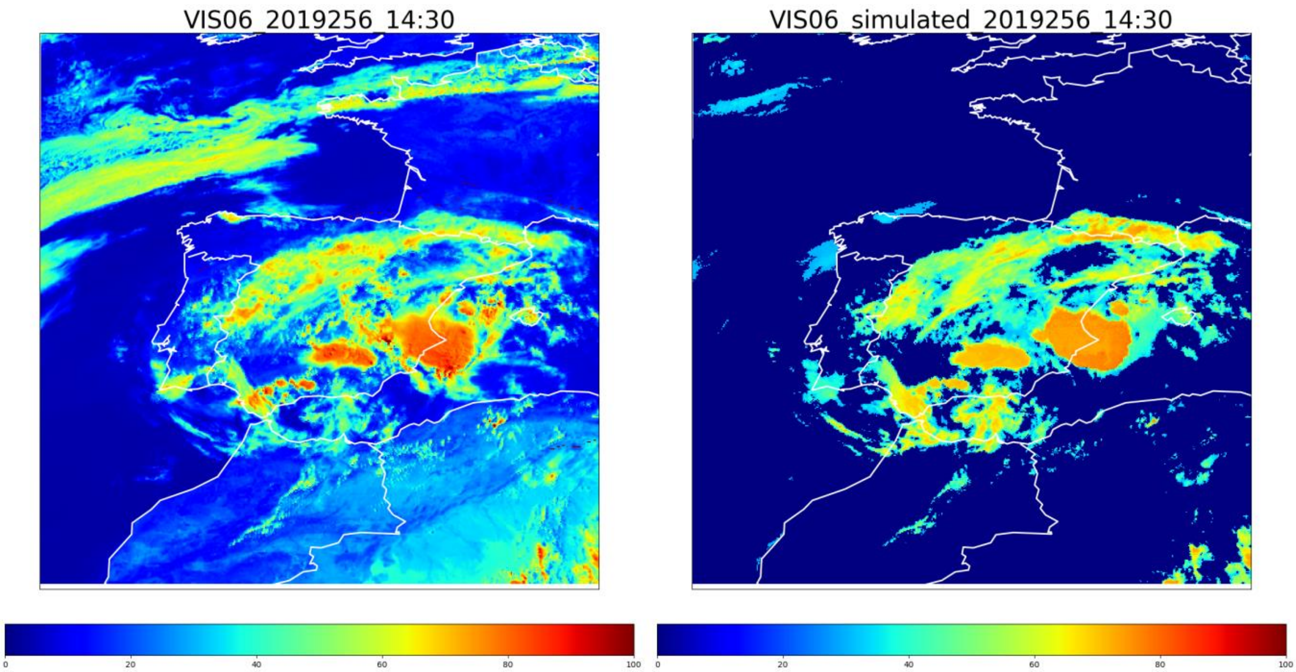



*Figure 35. $VIS_{0.6}$ simulation over Spain the 13th September 2019 at 14:30Z*

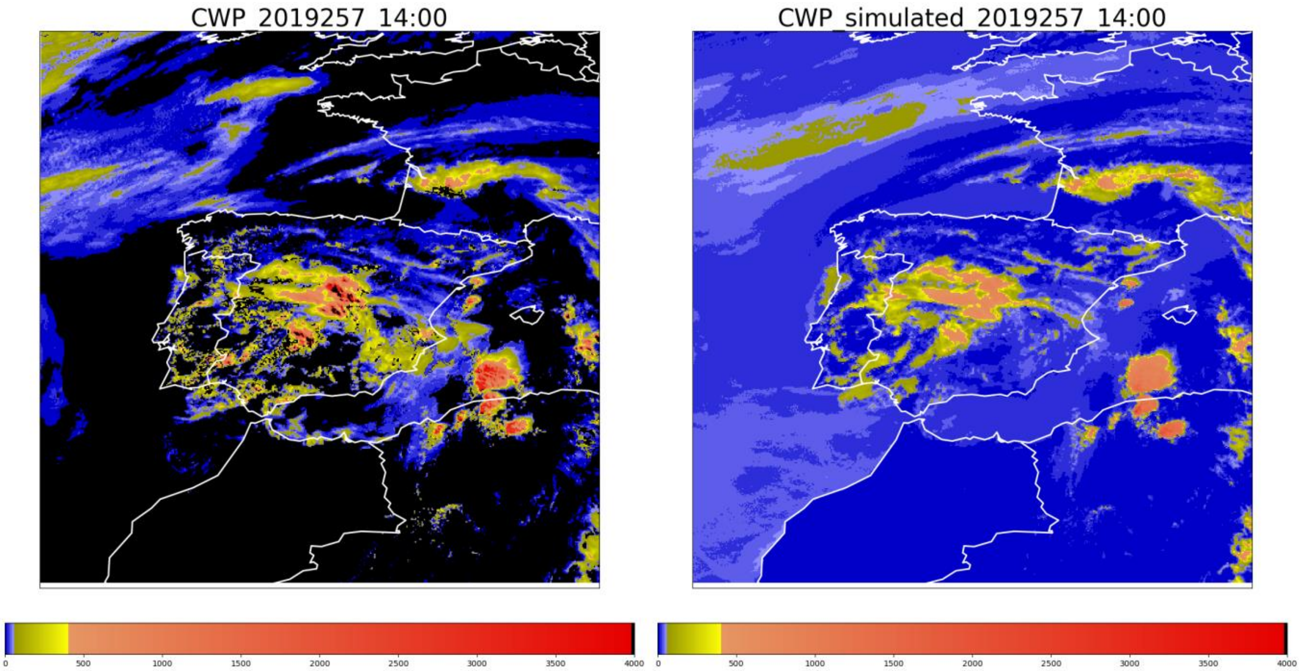



*Figure 36 Cloud Water Path ($gm^{-2}$) simulation over Spain the 14th September 2019 at 14:00Z. The colour palette is the standard one for the NWC SAF CMIC product.*

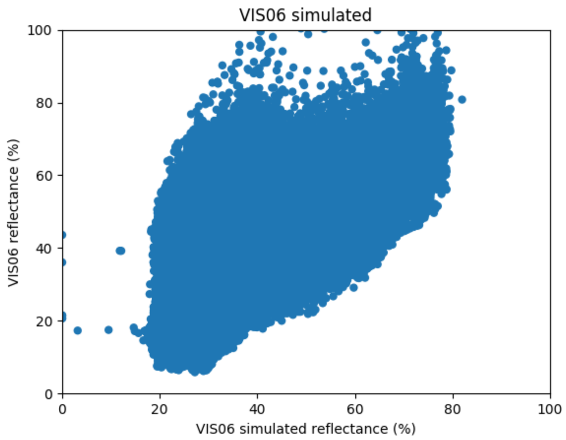



*Figure 37. VIS06 scatter plot the 13th September 2019 at 14:30Z over Spain*

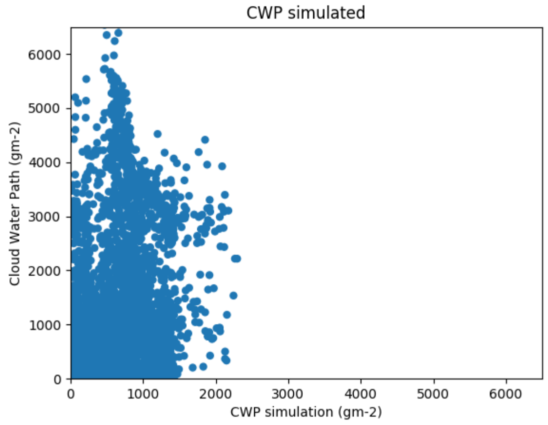



*Figure 38. CWP scatter plot the 14th September 2019 at 14:00Z over Spain*

Above it has been represented CWP and VIS06 scatter plots for the two examples of visualization for one time slot.

In order to have a more detailed idea about the pseudo channels simulation it has been computed a larger period: from May to September, every thirty minutes between 10:00Z and 15:00Z. Figures 39 and 40 are shown below.

It can be appreciated that whilst CWP simulation in general underestimates the real CWP values, VIS0.6 simulation slightly overestimates the real VIS0.6 values. It seems to be better simulated VIS0.6 values than CWP values. It is likely that a larger standard deviation compared against its central value in CWP have an impact on this simulations.

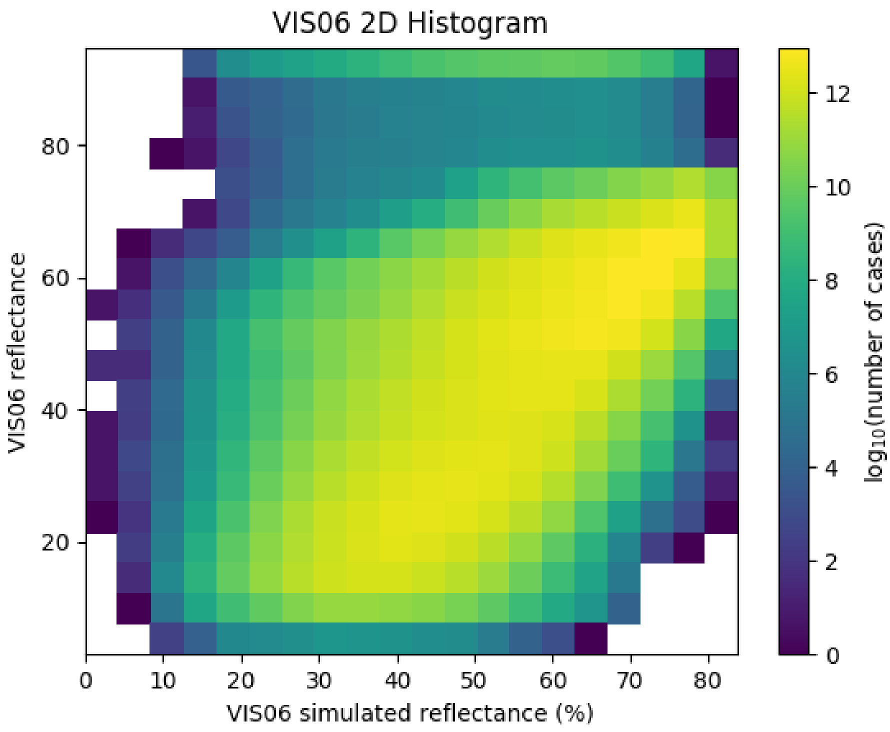


*Figure 39. VIS06 two dimensions histogram evaluated for Spain from May to September between 10:00Z and 15:00Z every thirty minutes. The colour palette is in a log scale.*

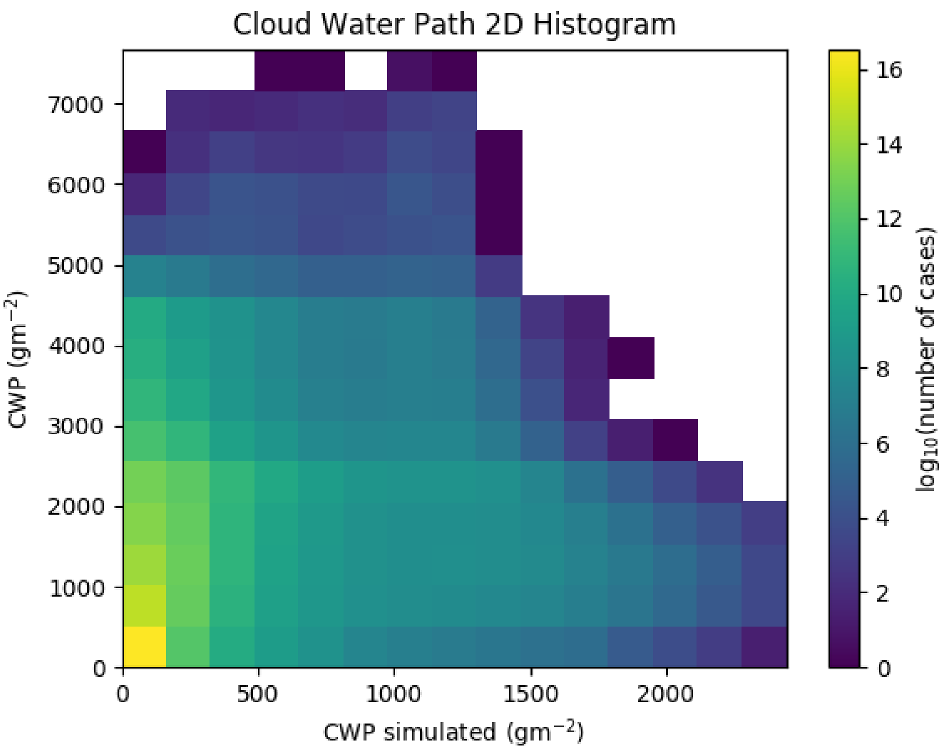



*Figure 40 Cloud Water Path two dimensions histogram evaluated for Spain from May to September between 10:00Z and 15:00Z every thirty minutes. The colour palette is in a log scale.*

In Figures 39 and 40 the colour palette appears in log-scale to provide a better visualization. In the first case values tend to concentrate in the diagonal, whilst in the second one it seems to be an underestimation of the CWP values. In figure 40, log scale facilitate the visualization of the occurrences of CWP below 1000 gm$^{-2}$, and around the minimum threshold values.

## 6.6 ASSUMPTIONS AND LIMITATIONS

As it has been explained in section 4.3 the standards related to the calibration area for PCPh and CRRPh have been different, however it has been used the same the pseudo-VIS06 and pseudo-CWP simulations at night. This has been done because it is considered that the "training" dataset is representative enough. It has been selected boxes of 25*25 pixels centred in convective pixels (RFR>10 mmh and ET>6km.) to calibrate both VIS06 and CWP at day time to be used at night time. The idea behind was to have a more homogeneous dataset, with less proportion of pixels with low rain rates.

In prior stages to have the final precipitation products, the VIS06 and the CWP calibration processes have been probed in different scenarios. Selecting boxes with the same criterion as CRRPh have been done a try. Although there were some differences, it didn´t have big impact in the validation report. Therefore, in the end it was decided to unify criteria and to use the same VIS06 and CWP for both products, PCPh and CRRPh.

The way the CWP has been created at night time it has an impact on the rain intensities. Looking back at figure 34, once (p1,p2) have been calculated for every pixel, there is a large number of CWP for every (p1,p2). It has been selected the mean value of the large data list of CWP values. By choosing a higher Percentile in CWP, it will be traduced on a higher rain rate in the CRRPh output. It must be taken into account that especially in winter at night time, there are False Alarms out of convective situations. A balance in False Alarms with low rain rates and Hits with higher rates must be taken into consideration.

Cloud water path thresholds for day and night time are used to delimit the precipitation area. The higher the threshold value of the CWP is set the smaller the precipitation area becomes.

In order to reduce some noise, and to be consistent with other LUT smoothing processes, the calibrating LUT has been smoothed in 3*3 boxes with the mean value.

Taking a look at figure 35 it can be noticed that both images are quite similar in terms of general appearance. The simulation does not reach so high values and there are some regions with low reflectance that are not well depicted. However, lower values of reflectance do not have a big impact on the precipitating area since they belong to thin clouds, and reflectance lower than 30 % are discarded.

Something similar can be appreciated in figure 36, CWP and the simulated one have a similar appearance. Let's take into account that the thresholds for CRRPh and PCPh are different.

It is not worthy to pay attention to areas with CWP lower than 250 $gm^{-2}$ since they are below the threshold to PCPh and will not have an impact in the probability of rain. It is also not use paying attention to CWP values lower than 350 $gm^{-2}$ since they are below the threshold to CRRPh and hence they do not contribute to rain rate output.

# 7. ANNEX A: PARALLAX CORRECTION

Two important factors for accurate precipitation estimations from satellite imagery are the position of the cloud tops and the influence of orographic effects on the distribution of precipitation.

The exact cloud position with respect to the ground below is needed to apply the CRR orographic correction. This is not a problem when a cloud is located directly below the satellite; however, as one looks away from the sub-satellite point, the cloud top appears to be farther away from the satellite than the cloud base. This effect increases as you get closer to the limb and as clouds get higher. Since parallax correction rectifies this effect, it is needed to be applied before orographic correction in the case of the CRR product.

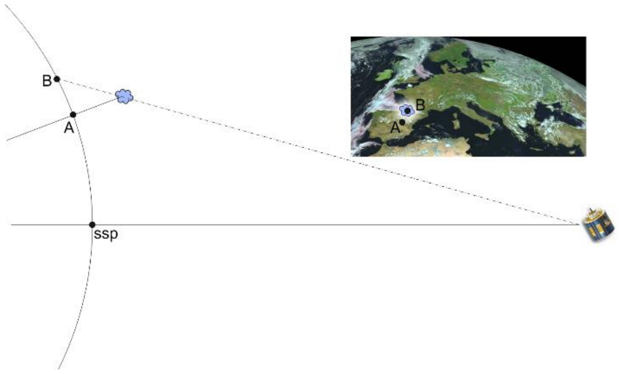



*Figure 41. Parallax geometry*

The parallax correction depends on three factors: a) the cloud height, b) the apparent position on the earth of that cloud and c) the position of the satellite.

The last two factors are known, but the first one has to be estimated. Two height estimation methods have been studied: numerical model and climatic profile obtained from the 1962 standard atmosphere model. Both of them are based on the conversion of each 10.8IR brightness temperature to height.

By default, height is estimated using NWP data. Parallax correction needs the NWP geopotential and temperature data at some levels (1000, 925, 850, 700, 500, 400, 300, 250 and 200). If NWP previous and next (according to the forecast time) models are available for the current slot time, a linear interpolation between these two models is performed.

Using 10.8IR brightness temperature, a linear interpolation is done among NWP temperatures and geopotential giving as a result the cloud height for each pixel. This height is then converted to meters.

In case of lack of NWP data or different number of pressure levels found (between temperature and geopotential) the NWP method for height calculation won't be used, and the climatic profile will be applied instead.

The used climatic data contain geopotential and temperature information related to five zones: 0º-15º, 15º-30º, 30º-45º, 45º-60º and 60º-75º. Two seasons are considered, summer and winter. A linear interpolation is used for latitude position and a cosine interpolation is used for Julian date.

Cloud height (in meters) is obtained using a bi-lineal interpolation according to the pixel temperature and considering the nearest four climatic temperature and geopotential measurements.

Parallax correction begins by converting the point and satellite locations into cartesian coordinates using the Earth centre as the origin. The Earth's surface is considered as an ellipsoid with an equatorial radius of 6378.077 Km. and a polar radius of 6356.577 Km. A virtual ellipsoid (as the earth's one) is performed using the distance from the cloud top to the earth centre. The cross point between the line joining the satellite and the apparent cloud surface position and this ellipsoid is found. The surface point connecting it with the Earth centre is then obtained, providing as result the new co-ordinate of the pixel. Finally, cartesian coordinates are converted into geographical ones.

When Parallax Correction is working, a spatial shift is applied to every pixel with precipitation according to the basic CRR value. In this re-mapping process, and only for a very small percentage of pixels, it could happen that (1) two pixels of the original image are assigned to the same pixel of the final image or (2) a pixel of the final image is not associated to any pixel of the original image (a “hole” appears in the final image). To solve these special cases, the next solutions have been implemented in the software:

- Case (1): the algorithm takes the maximum value of the rainfall rate
- Case (2): the software identifies the pixels with “hole”. A 3x3 median filter centred on that hole pixel is applied in order to assign a rainfall rate value (to compute the median, the pixels within the 3x3 box identified as holes are excluded)

The theoretical basis used in the computing of the Parallax correction in the CRR, PCPh and CRRPh products and the Parallax Correction Processor of the NWC/GEO [RD 7] is the same.

# 8. ANNEX B: LIGHTNING ALGORITHM

The lightning algorithm is based on the assumption than the higher is the spatial and temporal density of lightning occurrence, the stronger is the convective phenomenon and the higher is the probability of occurrence and the intensity of convective precipitation.

Only Cloud-to-Ground lightning flashes are used by this algorithm. To incorporate this information into the product a rain rate has been assigned to every lightning depending on:

- the time distance ($\Delta\tau$) between the lightning event and scanning time of the processing region centre.

- the location of the lightning

- the spatial density of lightning in a time interval.

In order to know the rain rate to be assigned to each lightning the process proposed in Tapia et all. (Tapia et al., 1998) has been followed in this way:

A representative set of convective storms occurred over Spain have been selected. For each of them a Rainfall-9io (RLR) has been computed. This RLR takes into account the quantity of precipitation measured as well as the number of lightning occurred during each event. The mean of the RLR obtained for the selected storms is 10.08 mm/lightning.

The procedure followed is the following:

First of all, the number of lightning occurred within an interval $\Delta t$ before the scanning time of the processing region centre, are assigned to each pixel according to its latitude and longitude. The interval $\Delta t$ is selected by the user (default value: 15 minutes).

Afterwards a rain amount is assigned to every pixel according to the number of lightning allocated to it. The variability of the spatial correlation between lightning and rainfall within the storm area suggest the use of a uniform distribution of rainfall about lightning flashes (Tapia et al., 1998). For this reason, instead of assigning the RLR just to one pixel, this quantity of precipitation is spread around the pixel in order to obtain a more homogeneous pattern of precipitation in this way:

| | | | | |
|---|---|---|---|---|
| Z4 | $\frac{Z3+Z4}{2}$ | Z3 | $\frac{Z3+Z4}{2}$ | Z4 |
| $\frac{Z3+Z4}{2}$ | $\frac{Z2+Z3}{2}$ | Z2 | $\frac{Z2+Z3}{2}$ | $\frac{Z3+Z4}{2}$ |
| Z3 | Z2 | Z1 | Z2 | Z3 |
| $\frac{Z3+Z4}{2}$ | $\frac{Z2+Z3}{2}$ | Z2 | $\frac{Z2+Z3}{2}$ | $\frac{Z3+Z4}{2}$ |
| Z4 | $\frac{Z3+Z4}{2}$ | Z3 | $\frac{Z3+Z4}{2}$ | Z4 |

Figure 42. Spreading of the RLR value in a 5 by 5 pixels box

Being Z1, Z2, Z3 and Z4 the rain rate assignments according to the RLR obtained in the calibration process. The spreading of the RLR value has been done in the following way:

Z1= 0.228 * RLR (default value: 2.30 mm)

Z2= 0.074 * RLR (default value: 0.75 mm)

Z3= 0.025 * RLR (default value: 0.25 mm)

Z4= 0.010 * RLR (default value: 0.10 mm)

Simultaneously, the time of occurrence of each lightning event is taken into account. Since the point of view of instantaneous precipitation rates, lightning closer in time to the instant of rainfall measurement are better spatially correlated to the convective nuclei at that moment. So a higher weight is given to those lightning that occurred closer in time to the scanning time of the processing region centre (CRRPh time). To do that, all rain rates already assigned are multiplied by the factor COEFF_τ being:

$$COEF\_\tau = -1*10^{-7}(\Delta\tau)^4 - 3*10^{-3}(\Delta\tau)^2 + 1$$

Where Δτ is the interval of time between the time of occurrence of the lightning and the CRRPh time:

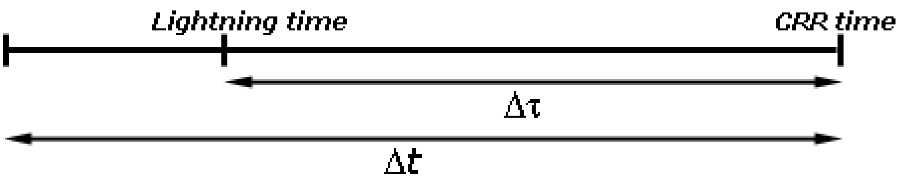



Figure 43. Diagram that shows the relationship between Δτ and Δt

Based on the fact than the higher is the spatial density of lightning occurrence the higher is the probability of the occurrence of greater intensities of precipitations, the density of lightning around each pixel is taken into account in the last step. To do that, rain rate corresponding to each pixel is multiplied by COEFF_N with:

$$COEFF\_N = a*\left(1-b^N\right)$$

Where N is the number of lightning occurred in a 11x11 pixels box centred on every pixel within the Δt interval. a and b are the parameters of the equation (default values: a=0,45; b=0,7).

It is emphasized that once the precipitation pattern has been computed, it is compared to the CRRPh precipitation pattern in order to obtain the final product. This final product contains the highest rain rate of the two.

Instructions on how to tune lightning algorithm can be found in the User Manual for the Precipitation Product Processors of the NWC/GEO [RD 5].

# 9. ANNEX C: HOURLY ACCUMULATIONS

At the end of the process the final values of the rainfall rates in mm/h are used in order to obtain hourly accumulations. A trapezoidal integration (Sánchez-Sesma and Sosa, 2004) is performed in order to compute the hourly accumulations.

Normal mode:

Six scenes are used in this process: the instantaneous scene corresponding to the time of the hourly accumulation and the five previous instantaneous scenes. The rain rate in mm/h output is the one used to make the computing.

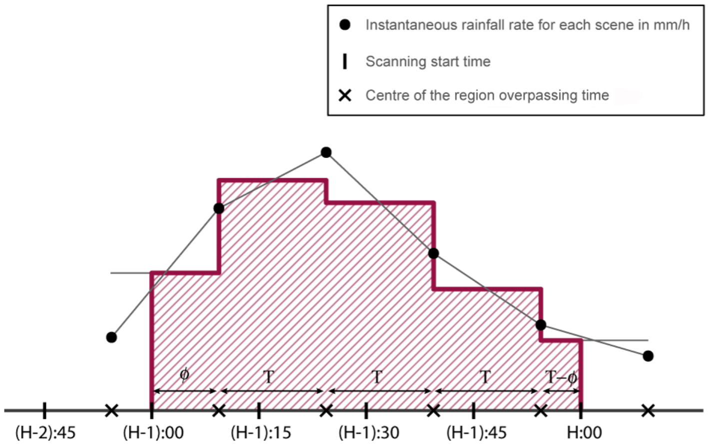



Figure 44. Trapezoidal integration

The nominal time of a scene corresponds to the moment when the satellite starts the scanning. Some minutes are needed to overpass the centre of the region where the product is being running. In order to avoid the time window effect, the following equation has been used to compute the hourly accumulations:

$$A_6 = \frac{I_1 + I_2}{2}\phi + \frac{I_2}{2}T + I_3 T + I_4 T + \frac{I_5}{2}T + \frac{I_5 + I_6}{2}(T - \phi)$$

Where:

- $A_i$: hourly accumulation, in mm, corresponding to the time i.
- $T$: time interval between scenes in hours (T= 0.25)
- $\Phi$: part of $T$ that corresponds to the time that takes the satellite to reach the centre of the region.
- $I_i$: Instantaneous rainfall rate for each scene in mm/h

The hourly accumulation won't be computed when there is a lack of more than two scenes or two consecutive ones in the complete interval.

Rapid Scan mode:

Fourteen scenes are used in this case: the instantaneous scene corresponding to the time of the hourly accumulation and the thirteen previous instantaneous scenes.

The equation that is used in the trapezoidal integration for the Rapid Scan mode is:

$$A_{14} = \frac{I_1 + I_2}{2}\phi + \frac{I_2}{2}T + \left(\sum_{i=3}^{12} I_i\right)T + \frac{I_{13}}{2}T + \frac{I_{13} + I_{14}}{2}(T - \phi)$$

Where:

- $A_i$: hourly accumulation, in mm, corresponding to the time i.
- $T$: time interval between scenes in hours (T= 1/12)
- $\Phi$: part of $T$ that corresponds to the time that takes the satellite to reach the centre of the region.
- $I_i$: Instantaneous rainfall rate for each scene in mm/h

The hourly accumulation won't be computed when there is a lack of more than six scenes or four consecutive ones in the complete interval.