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

NWC/CDOP3/GEO/MF-PI/SCI/ATBD/Convection, Issue 1, Rev 1.0

22nd February 2024

Applicable to

GEO-CI 2.23 (NWC-089)

GEO-RDT 5.23 (NWC-090)

Prepared by METEO-FRANCE Toulouse (MFT) / Direction des Opérations – Prévision Immédiate

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	METEO-FRANCE MFT (M. Claudon and R. Houel)	Signed, M. Claudon Signed, R. Houel	22nd February 2024
Reviewed by	METEO-FRANCE MFT (J.-M. Moisselin)	Signed, J.-M. Moisselin	22nd February 2024
Authorised by	P. Ripodas NWCSAF Project Manager		

DOCUMENT CHANGE RECORD

Version	*Date*	*Changes*
1.0	1st September 2021	First edition of the document for v2021 Some information concerning MTG are in the document even if MTG is not committed for this version
1.0.1	28th February 2022	Minor changes + changes in paragraph “improvement from previous version” to stick with other documents
1.1.0	22nd February 2024	v2021.3 release: - bug correction and handling of long convective trajectories - update of the product_quality NetCDF attribute

Table of contents

List of Tables and Figures

1 INTRODUCTION

1.1 Scope of the document

The ATBD document provides the scientific description of the convection products algorithm. It points out assumptions done on algorithms and limitations of products. This document summarizes product validation result and describes the outputs.

1.2 Scope of other documents

The UM (User Manual) provides all useful information to users.

The VR (Validation Report) depicts the accuracy of each product.

The Interface Control Documents ICD/1 (Interface Control Document n°1) describes the External and Internal Interfaces of the NWC/GEO software.

The Interface Control Documents ICD/2 (Interface Control Document n°2) describes the input and output data formats of the NWC/GEO software.

1.3 Software Version Identification

This document describes the products obtained from the GEO-CI 2.23 (Product Id NWC-089) and from the GEO-RDT-CW 5.23 (Product Id NWC-090) implemented in the release of the NWC/GEO software package.

1.4 Improvement from previous version

Improvements from previous releases are:

- GEO-RDT-CW and GEO-CI (in decreasing order of importance)
    - In v2021.1: support to GOES-18 satellite
    - In v2021.3: update of the product_quality NetCDF attribute: value is penalized if NWP data are missing
    - In v2021.1: addition of the land/sea mask, of the sun and satellite angles in RDT-CW attributes
    - In v2021.1 and v2021.2 : code optimization and bug corrections
- GEO-RDT-CW (in decreasing order of importance)
    - In v2021.2: switch MSG3 / MSG4: update of RDT discrimination files.
    - In v2021.1: Improvement of the cell detection toward a limitation of number of cells in very specific cases
    - In v2021.3: Bug correction and handling of long mature trajectories (toward a faster declassification)
    - In v2021.3: distinction of the polarity of intra-cloud lightning (provided by some ground-based lightning network)
    - In v2021.3 optional choice of the type of cells to output in RDT NetCDF file.
    - In v2021.3 Bug correction in the overshooting tops algorithm

- GEO-CI
  - In v2021.1: addition of a filter to remove pixels belonging to a large cell (over 10 000 km²)
  - In v2021.1 the cloud type a previous slots is taken into account when computing parameter's trends.

1.5 Glossary, Acronyms and Abbreviations

See [R.D. 1] for a complete list of acronym for the NWC SAF project.

1.6 References

1.6.1 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: http://nwc-saf.eumetsat.int

Table 1: applicable documents

Ref	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	11/4/2016
[AD.2.]	Project Plan for the NWCSAF CDOP3 phase	NWC/CDOP3/SAF/AEMET/MGT/PP	1.3	12/7/2019
[AD.3.]	Configuration Management Plan for the NWCSAF	NWC/CDOP3/SAF/AEMET/MGT/CMP	1.0	21/2/2018
[AD.4.]	NWCSAF Product Requirement Document	NWC/CDOP3/SAF/AEMET/MGT/PRD	1.5	1/12/2021
[AD.5.]	System and Components Requirements Document	NWC/CDOP2/GEO/AEMET/SW/SCRD	1.3	13/11/2020
[AD.6.]	Interface Control Document for Internal and External Interfaces of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/ICD/1	1.1	1/10/2019
[AD.7.]	Interface Control Document for the NWCLIB of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/ICD/2	1.1	1/10/2019
[AD.8.]	Data Output Format for the NWC/GEO	NWC/CDOP3/GEO/AEMET/SW/DOF	2.0.3	22/02/2024
[AD.9.]	Component Design Document for the NWCLIB of NWC/GEO	NWC/CDOP2/GEO/AEMET/SW/ACDD/NWCLIB	2.0	27/2/2017
[AD.10.]	Component Design Document for the Convection Product Processors of the NWC/GEO	NWC/CDOP2/GEO/MFT/SW/ACDD/Convection	2.1	21/1/2019
[AD.11.]	User Manual for the Tools of the NWC/GEO	NWC/CDOP3/GEO/AEMET/SCI/UM/tools	1.0	21/1/2019

Ref	Title	Code	Vers	Date
[AD.12.]	MTG-SAF System Budget Analysis and SAF Level 2 Products Generation and Dissemination Baseline for MTG - [L2SAF]	EUM/PPS/DOC/09/0032	v1H e-signed	

1.6.2 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://nwc-saf.eumetsat.int

Table 2: reference documents

Ref	Title	Code	Vers	Date
[R.D. 1]	The Nowcasting SAF Glossary	NWC/CDOP3/SAF/AEMET/MGT/GLO	1.0	20/10/2020
[R.D. 2]	Best Practice Document, 2013, for EUMETSAT Convection Working Group, Eds J.Mecikalski, K. Bedka and M. König »	Available on CWG Website		2013
[R.D. 3]	Karagiannidis, A., 2016, Final Report on Visiting Scientist Activity for the validation and improvement of the Convection Initiation (CI) product of NWC SAF v2016 and v2018, Visiting Scientist Activity followed in Nowcasting Department of Météo France, Toulouse, France Period June-December 2016 »	available on NWCSAF Website		2016
[R.D. 4]	Scientific Report on verification of RDT forecast	NWC/CDOP3/GEO/MFT/SCI/RP/01	1.1	2018
[R.D. 5]	Schultz,C.J., W.A. Petersen, and L.D. Carey, 2009: Preliminary developmeent and evaluation of lightning jump algorithms for the realtime detection of severe weather. J.Appl. Meteor. Climatol., 48, 2543-2563			2009
[R.D. 6]	de Laat, A., Defer, E., Delanoë, J., Dezitter, F., Gounou, A., Grandin, A., Guignard, A., Meirink, J. F., Moisselin, J.-M., and Parol, F., 2017, Analysis of geostationary satellite-derived cloud parameters associated with environments with high ice water content, Atmos. Meas. Tech., 10, 1359-1371, https://doi.org/10.5194/amt-10-1359-2017, 2017			2017

2 DESCRIPTION OF CI (CONVECTION INITIATION) PRODUCT

2.1 PRODUCT OVERVIEW

2.1.1 Goal of the product

CI provides the probability for a cloudy pixel to become a thunderstorm during the following minutes included in a given period (30, 60 or 90 minutes). The product aims to detect the first steps of initiation of convection, when the first convective signs occur after the formation of clouds, with a modification of environmental conditions.

2.1.2 Product description

Probability of the formation of a thunderstorm depends on evolution of local condition and on advection of clouds. Since the information is often too scarce for a full object-approach, CI algorithm is a mix of pixel-approach and object-approach. CI final product is at pixel scale.

CI is defined for three time-steps (0-30', 0-60' and 0-90') and for four classes of probability (0-25%, 25-50 %, 25-50 %, 75-100%). For example if a given pixel at a given time T has a probability interval of 50-75% for the range 0-60', it means that the pixel has a probability between 50% and 75% to become a thunderstorm between T and T+60 minutes.

2.1.3 Terminology

For a given slot a time T

- Pixel ***eligible-CI***: a pixel passing a first filter verifying if basic conditions for convection are satisfied regarding cloudy environment, instability indices, etc. The objective is to avoid non-initiation phases and very cold pixels.
- Pixel ***pre-CI***: a pixel eligible-CI passing a second filter verifying if convection has chance to start regarding BTD or BT values.
- Pixel ***CI***: a pixel that is likely to become a thunderstorm. Pixel CI-30 is a pixel that will become a thunderstorm within the interval [T, T+30']. CI-60 and CI-90 are defined in a similar way.

2.2 ALGORITHM DESCRIPTION

2.2.1 Theoretical Description

2.2.1.1 Overview

CI is a mix of object and pixel analysis, of physical and statistical approach. The methodology is

1. to identify areas of interest, which are areas of ***eligible-CI*** pixels
2. to determine a guess of ***2D movement field*** to be representative of cloudy pixel movement
3. to undertake, over areas of interest, ***cloud cell detection and tracking*** in order to
   - *correct, update and complete* the 2D movement field

- increase the number of slots from which some pixels are tracked

4. to calculate satellite characteristics of ***eligible-CI*** pixels, including historic of the pixel thanks to 2D movement field (***static and dynamic characteristics***)
5. to determine ***pre-CI*** pixels using relevant thresholds of parameters
6. finally to evaluate convection through probability assessment, and localize corresponding ***CI*** pixels

Relevant parameters, thresholds and some part of the algorithm are inspired from « Best Practice Document » [R.D. 2], especially SATCAST methodology for the definition of pre-CI pixels. Major improvements of versions following the first v2016 release have been defined during the Visiting Scientist activity in the period June-December 2016 [R.D. 3]. These improvements concern more specifically the tuning (thresholding, split of the algorithm in day and night parts), the use of CMIC. Learning data relies on case studies and features representative of convection triggering. It is dedicated to tune interest fields' thresholds and coefficients to evaluate a diagnosis of convection initiation.

2.2.1.2 Description of the Algorithm

2.2.1.2.1 Area of interest and eligible-CI

This step requires in optimum configuration

- NWP data, to eliminate stable areas and focus on more unstable pixels,
- Cloudy filtering, to filter non-cloudy pixels with NWCSAF cloud products, and focus on very low, low and medium categories,
- 10.8µm BT (or 10.3µm BT for GOES-16), in order to ignore cold cloudy pixels

Thus, large areas are ignored in the following processes, which focus on a restricted set of pixels to be analysed.

2.2.1.2.1.1 NWP convective mask

Convection products, and especially Convection Initiation, take benefit from NWP guidance before attempting analysis.

NWP data are used to produce a convective mask through availability or computation of several convective indexes: K index, Showalter index and Lifted index. The union of these indexes allow to identify stable areas where possibility of convection will be very low.

Indices used for convection mask can reflect unstable, unclear or stable meteorological situations. Values of the mask are 0 if all indexes are stable, 2 if at least one index is unstable, 1 in other cases. Threshold values for convective indexes are the following:

Table 3: use of instability indices in CI

	stable	unstable
Lifted Index	LIX > 0	LIX < -3
Showalter Index	SHW >3	SHW < -3
K index	K<20	K>30

Depending on regions or on user's needs, those values can be modified through specified arguments in configuration file (*.cfm, see User Manual), leading to enlarge or reduce the area of interest. Those arguments are the following: LICONV, LINOCONV, KICONV, KINOCONV, SHWCONV, SHWNOCONV.

- Full stable case: pixel value of NWP_Mask = 0 (if all convective indexes are associated to stability "stable" or at least one is associated to stability "stable" and other missing)
- Unstable case: pixel value of NWP_Mask = 2 (if at least one of the convective indexes is associated to "instability")
- Unclear case: pixel value of NWP_Mask = 1 (for other cases than above)

Regions with null (0) values, i.e. stable areas, will be ignored in further analysis.

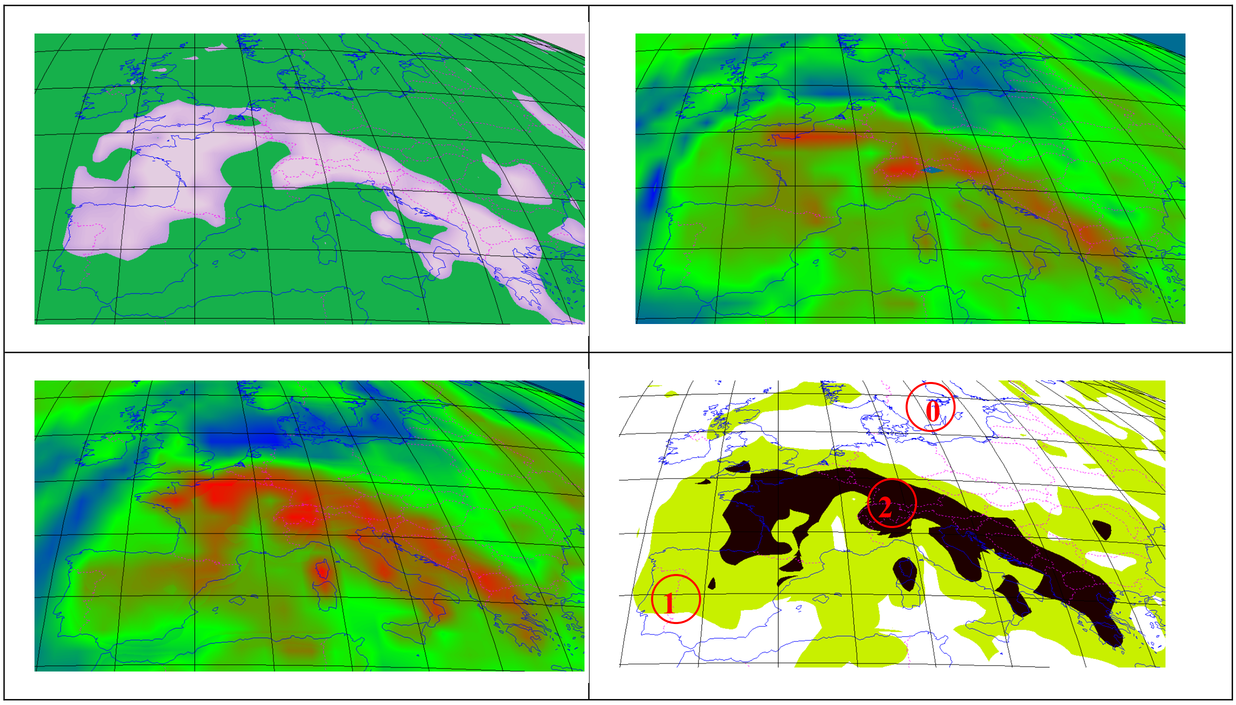


Figure 1: 25 May 2009, 12h15. Convective mask (bottom right), as a union of K index (top left), Showalter index (top right) and Lifted index (bottom left), from NWP data. Regions 1 and 2 are regions of interest for analysis

2.2.1.2.1.2 Cloud type as filtering mask

Cloud Type product is used as input to GEO-CI with two objectives:

- Use CT product as Cloud mask to exclude non cloudy pixels from further 2D analysis.
- Restrict the panel of pixels of interest to those concerned with "Very Low" to "Medium" cloud type, and ignore in this version "Fractional" cloud type or those which could be superimposed with thin cirrus.

An illustration of the interest to take into account a Cloudy filtering to complete IR10.8 BT limitation is presented in following figure. Some false alarms are removed using this step.

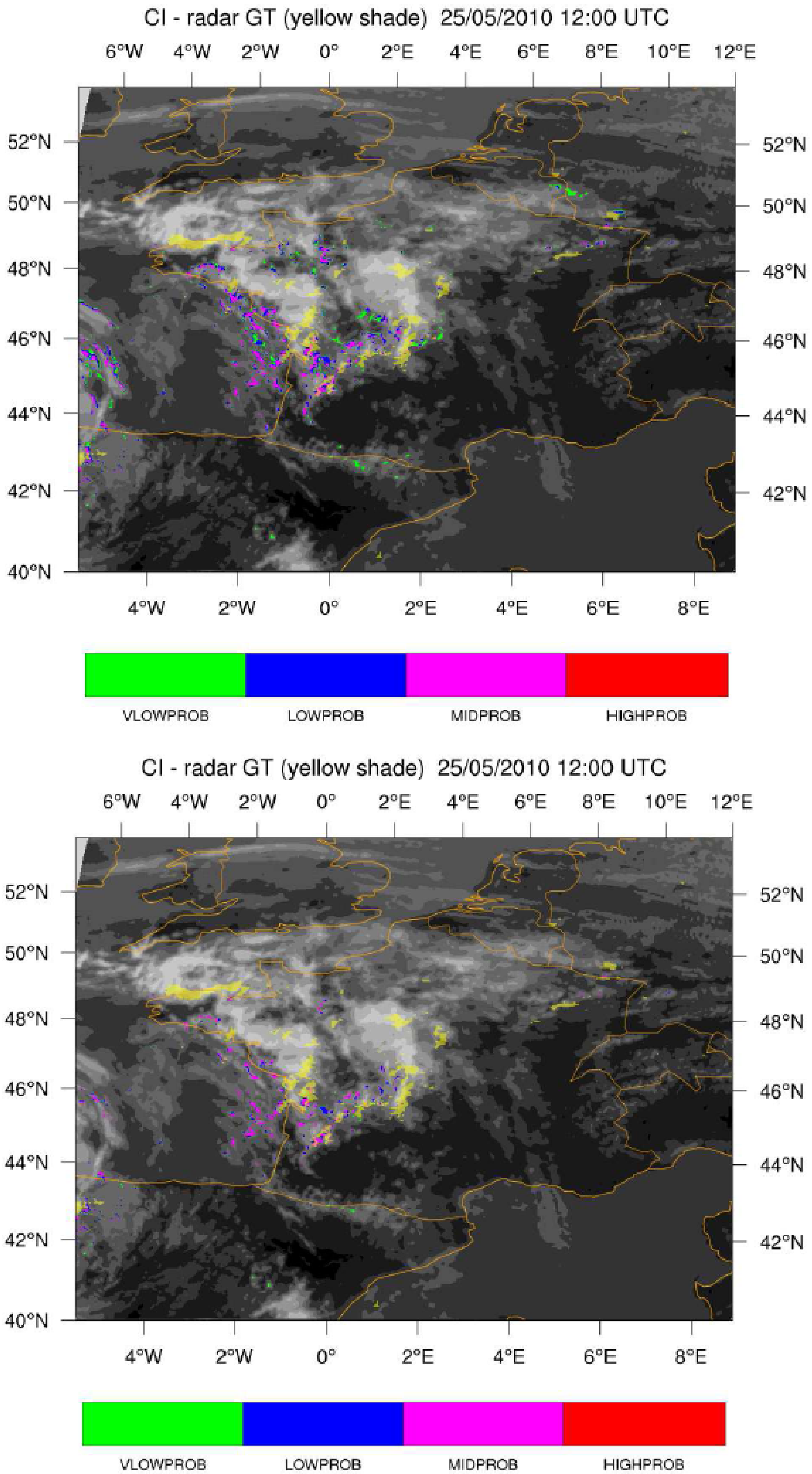


Figure 2: CI (colour bar) and radar Ground Truth (yellow shade) considering all cloud types (top map) and only very low, low and medium clouds (bottom map)

2.2.1.2.2 2D movement field

A 2D movement field is estimated in optimum configuration with last available HRW wind observations. Those observations are filtered: “Wind temperature and pressure level” processing indicator of HRW is used to limit the number of observations taken into account to lowest levels, and associate the right IR channel to 10.8 µm BT of the pixel. One HRW retrieval by pixel is kept. HRW observations are then remapped on the satellite grid-field region, through a weighted interpolation taking into account a radius of influence and a cut-off distance. An additional blending with low level (850hPa) NWP wind field can be optionally processed where no information is available, but priority is given to HRW wind observations.

This blended field is at first used as guess movement in the next process (tracking step in the object analysis process) for the movement initialization in “cold start” (first run) cases, or for orphan cells in the recovery analysis process (see chapters “The detection of cloud systems” and “The tracking of cloud systems”).

In its final state, once updated by cloud cell detection and tracking (see next step), it will be considered as a pixel tracker for trends calculations.

2.2.1.2.3 Object image-analysis: Cloud cell detection and tracking

An object analysis process is undertaken like in RDT-CW software. The objectives of this step are

- To take benefit from techniques allowing to catch cloud cells movement
- To access cloud cells’ parameters variations along its trajectory

More details about tracking can be found in chapters “The detection of cloud systems” and “The tracking of cloud systems” of this document (RDT-CW description).

CI specificities rely on:

- Limits for adaptative thresholding:
    - The warmest limit is set to +10°C in order to catch cloud cells as early as possible.
    - The coldest limit has been reset to a cold value of -75°C in order to better identify and track cloud systems. The filtering of highest / coldest pixels and cells is undertaken later in the code
- Minimum vertical extension of objects, which has been set to 3° instead of 6° in RDT-CW product to focus on lower extended cloud systems

This step takes benefit from movement guess field as *input* to increase cell’s speed reliability, and on the other hand delivers as *output* an updated movement field with the analysed objects’ speeds. All pixels belonging to a tracked and quite warm cloud system are assigned the corresponding movement speed instead of previous pixel’s movement values.

This final blended movement field is a key point for further relevant trends calculations (see next step).

2.2.1.2.4 Pixel image-analysis: BTD and trends

Brightness Temperature values and Differences are processed for each *eligible-CI* pixel from various available channels, for current data and data from previous slots.

BTDs taken into account are

- WV6.2-WV7.3,
- WV6.2-IR10.8,

- IR10.8-IR8.7,
- IR12.0-IR10.8,
- IR13.4-IR10.8,

Then, BT (IR10.8) and BTD trends are calculated for each eligible-CI pixel using the speed and direction of updated 2D movement field as guidance for identifying pairs of current and corresponding pixels in previous images. More precisely, short-time (time gap of 15 minutes or so) and long-time (time gap of 30 minutes or so) trends are computed for each pixel and are adapted to satellite scan frequency or data availability, according to the Table 4:

Table 4: time gaps between current slot and considered slot to compute short-time and long-time trends

Satellite or mission	Scan frequency or availability of data	Short-time trend – time gap	Long-time trend – time gap
MSG-0° / MSG-IODC	15 minutes	15 minutes	30 minutes
MSG-RSS	5 minutes	15 minutes	30 minutes
GOES-W / GOES-E	10 minutes	10 minutes	30 minutes
HIMAWARI	20 minutes	20 minutes	40 minutes

Pixel positions at previous slots are retrieved from the 2D movement field through retro-advection process.

Trend's value is then computed within a 3x3 pixel box of influence in the previous image. A median value is considered instead of an extrema value, in order to reduce false alarms.

When tracking of aggregated pixels (belonging to a tracked cloud system as object) is available, corresponding trends are used for some parameters instead of single pixel-trends, and should be able to provide trends over longer depth.

2.2.1.2.5 Additional pixel filtering

2.2.1.2.5.1 Day-time microphysics filtering

During day-time when microphysics parameters COT (Cloud optical thickness), LWP (liquid water path) and IWP (ice water path) are available from CMIC product, they are used as an additional filter to focus on relevant areas. The table below illustrates the ranges of values of microphysics parameters used for this purpose.

Table 5: thresholds for microphysics parameters.

Parameter name	***Relevant value***	***Meaning***
COT	13,42 (unitless)	Cloud optical thickness
LWP	0,22 kg/m^2	Liquid water path
IWP	0,17 kg/m2	Ice water path

A cloudy pixel is eligible for convection initiation, only when at least one of the three microphysics parameters has a value over the threshold. Those thresholds have been tuned subjectively regarding a subset of relevant meteorological situations, taking into account the radar ground truth (Figure 3).

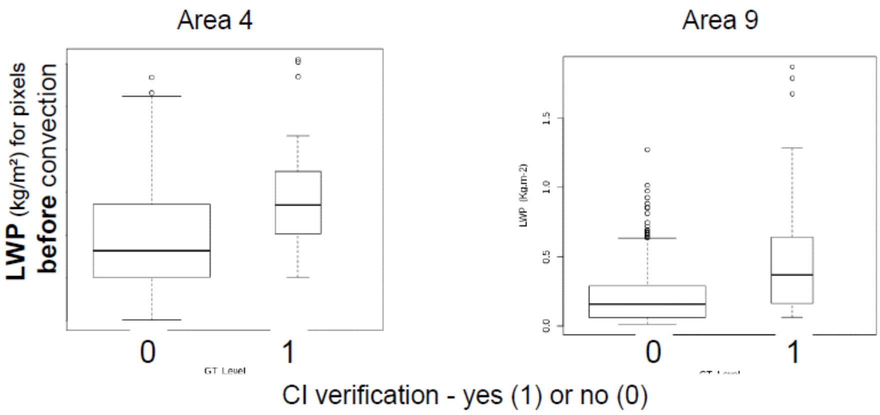


Figure 3: Box-plots for two regions and test cases used to define threshold for LWP. From [R.D. 3]

2.2.1.2.5.2 Including cell filtering

In order to reduce obvious false alarms sometimes appearing on the edge of cloud systems, the code has been adapted to identify pixels outside a main (tracked) cloud cell but belonging to an including cell.

An including cell corresponds to a cell defined at the first level below (warmer) the level of a main (tracked) cloud cell. It is supposed to include more than one main (tracked) cell. This step allows to ignore most of connected pixels belonging to a cloudy system but outside of a cloud cell object.

2.2.1.2.5.3 Large warm cell filtering

In order to reduce obvious false alarms sometimes appearing on large warm flat cloud systems, pixels belonging to large warm cloud systems are removed from further analysis based on a maximum surface (10 000 km²).

2.2.1.2.6 Definition of Pixel pre-CI

Each *eligible-CI* has then a list of BT and BTD values and trends. According to previous studies about convection initiation, parameters are grouped as:

- Representative for Cloud-top Glaciation
  - IR10.8 Brightness temperature
  - IR10.8-IR8.7 BTD
- Representative for Cloud depth and vertical extension
  - WV6.2-IR10.8 BTD
  - IR13.4-IR10.8 BTD
  - IR12.0-IR10.8 BTD
- Representative for Cloud growth (updraft)
  - All BTDs trends
  - IR10.8 BT trends

Pre-CI pixels have at least one significant relevant value (see tables below).

In the last step, some parameters of all pre-CI pixels will be analysed for a CI-diagnosis.

2.2.1.2.7 Diagnosis

2.2.1.2.7.1 Interest fields

Each pre-CI pixel is associated with a list of values from parameters groups mentioned above. The initial interest fields' threshold ranges, as illustrated in table below, are derived from previous studies on CI, and applied to a day-time algorithm, when CMIC product is available and CMIC conditions are fulfilled on microphysics parameters values.

Table 6: Initial interest fields thresholds for pre-CI filter and CI-diagnosis, when CMIC product is available

Parameter name	***Relevant value***	***Meaning***	***filter***	***CI-diagnosis***
BT IR10.8	> -25°]-20°C , 0°C]	Brightness temperature (***glaciation***)	Eligible-CI and pre-CI	X
BTFZG	[-30',0']	Time since freezing point (10.8 µm BT) (***glaciation***)	Pre-CI	X
BTD4	]-10° , 0°C[	IR10.8-IR8.7 (***glaciation***)	Pre-CI	X
BTD	]-35° , -10°C[	WV6.2-IR10.8 (***height***)	Pre-CI	X
BTD6	]-25° , -5°C[	IR13.4-IR10.8 (***height***)	Pre-CI	X
BTD5	]-3° , 0°C[	IR12.0-IR10.8 (***height***)	Pre-CI	X
WBTD	]-25° , -3°C[	WV6.2-WV7.3 (***height***)	Pre-CI	X
BT' 2.2µm		Corrected Brightness temperature		
TxBT 15'	] -4°/15' , -50°/15'[	Temperature change rate (***growth***)	Pre-CI	X
TxBT 30'	] -4°/15' , -50°/15'[	Temperature change rate (***growth***)	Pre-CI	X
TxBTD 15'	> 3°/15'	BTD 15 Trend (***growth***)	Pre-CI	X
TxBTD 30'		BTD 30 Trend (***growth***)		
TxBTD4 15'	]0°/15' , 10°/15'[	BTD 15 Trend (***growth***)	Pre-CI	
TxBTD4 30'		BTD 30 Trend (***growth***)		
TxBTD5 15'	]0°/15' , 10°/15'[	BTD 15 Trend (***growth***)	Pre-CI	
TxBTD5 30'		BTD 30 Trend (***growth***)		
TxBTD6 15'	> 3°/15'	BTD 15 Trend (***growth***)	Pre-CI	
TxBTD6 30'		BTD 30 Trend (***growth***)		

One can note that most of 30min trends are not kept in consideration. Focus remains mainly on 15min trends.

During night-time, when microphysics parameters are not available, stricter thresholds are used for non-trend (static) parameters in order to limit the number of false alarms. Table below illustrates those thresholds.

Table 7: Night-time Interest fields thresholds for static parameters for pre-CI filter and CI-diagnosis

Parameter name	***Relevant value***	***Meaning***	***filter***	***CI-diagnosis***
BT IR10.8	> -25° > -20°C	Brightness temperature (***glaciation***)	Eligible-CI and pre-CI	X
BTFZG	[-30',0']	Time since freezing point (10.8 µm BT) (***glaciation***)	Pre-CI	X
BTD4	]**-2,1°** , 0°C[	IR10.8-IR8.7 (***glaciation***)	Pre-CI	X
BTD	]**-34°** , **-12°**C[	WV6.2-IR10.8 (***height***)	Pre-CI	X
BTD6	]**-17°** , **-6°**C[	IR13.4-IR10.8 (***height***)	Pre-CI	X
BTD5	]**-2,2°** , 0°C[	IR12.0-IR10.8 (***height***)	Pre-CI	X
WBTD	]**-17°** , **-7°**C[	WV6.2-WV7.3 (***height***)	Pre-CI	X
BT 2.2µm		Brightness temperature		

2.2.1.2.7.2 Subjective Tuning of Interest Fields' thresholds

Interest fields' thresholds values have been assessed (day-time with CMIC) and tuned (night-time or without CMIC) regarding a limited relevant set of situations, focusing on regions where a CI product is expected to bring valuable information. Subjective analysis appeared to be the most appropriate approach to assess the objectives of CI product. This tuning approach is more detailed in the next chapter.

2.2.1.2.7.3 Probability assessment - Empirical rules

The CI output is then estimated with empirical rule defined by count of relevant criteria. The principle is to sum up the number of relevant parameters (i.e. above relevant threshold given in the tables above) by group, giving greater importance to *growth* family parameters, then *glaciation* parameters, and finally vertical extension (height) group.

Table 8: Empirical rules for CI-diagnosis. HIGHPROB means between 75 and 100%, MODPROB between 50 and 75%, LOWPROB between 25 and 50%, VLOWPROB between 0 and 25%.

Number of Growth relevant parameters (over 4)	***Nb of Glaciation relevant parameters (over 3)***	***Nb of Height relevant parameters (over 4)***	***Result***
≥ 3	≥ 3	≥ 4	HIGHPROB
		≥ 3	MODPROB
	≥ 2	≥ 4	LOWPROB
≥ 2	≥ 3	≥ 4	MODPROB
		≥ 3	LOWPROB
	≥ 2	≥ 4	VLOWPROB
>=1	≥ 3	≥ 4	VLOWPROB
Other cases			0

Relevant parameters have been adapted to the various forecast depth, once the movement field has allowed to reanalyze the corresponding Ground truth after an advection scheme.

2.2.1.2.8 Forward scheme

Once a pixel is diagnosed as CI, the diagnosis is spread along the trajectory defined accordingly the pixel tracker. The forward scheme was not activated in first release (v2016) of the product. Figure below illustrates the forward scheme.

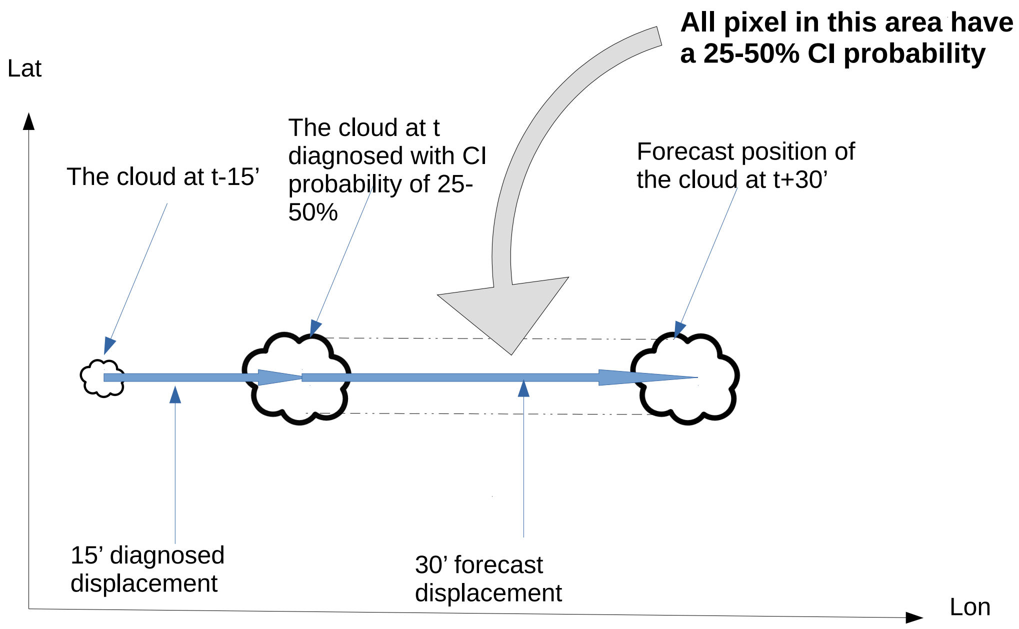


Figure 4: Forward scheme of CI

2.2.2 Practical Considerations

2.2.2.1 Calibration and tuning

Tuning CI probability assessment is undertaken using various ground truths, among which mainly radar data (for convective signature, Radar Ground Truth - RGT), but also RDT-CW (for convective cells identification, RDT Ground Truth - RDTGT).

CI probabilities concern a given period [0-xx mn]. Ground truth has to be compliant with this notion: not only convective signatures and cloud cells' positions at given slot have to be taken into account, but also all pixels along the path of convection over the period:

- When applied to radar data, all pixels over 30 dBZ during the given period are accumulated, and smoothed thanks to a median filter in order to produce a mask of convective activity. Meteorological situations are chosen to focus on convective events, and NWP convective masks ensure to avoid high reflectivity not related to convection.
- When applied to RDT, it has been decided to consider the envelope (convex hull) of all cloud cell contours during the period, to produce a mask of convective activity.

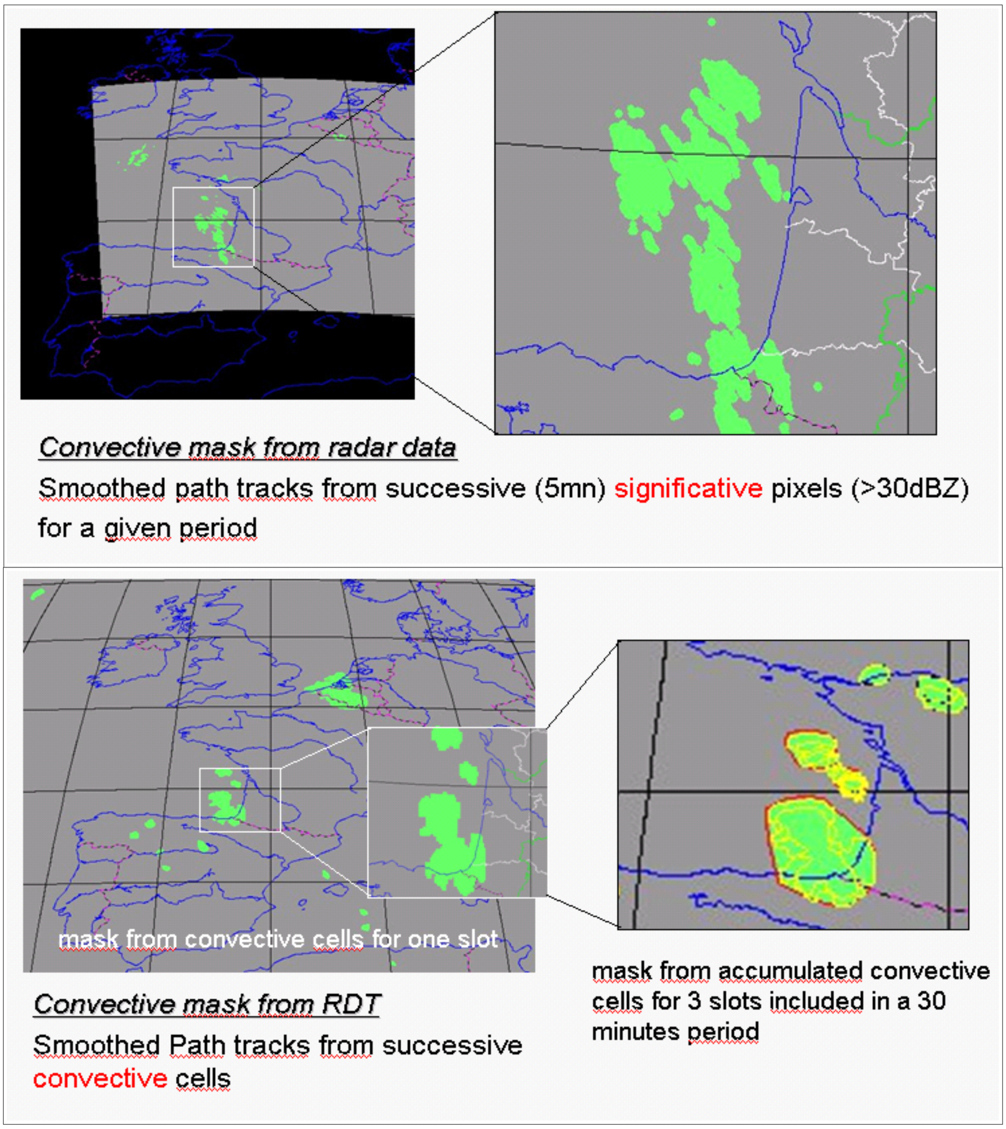


Figure 5: Ground truth: radar (top) and RDT (bottom) path over [0-30min] period for CI

Thus, different Ground truth masks have to be considered, corresponding to the three periods of convection initiation probabilities [0-30mn], [0-60mn], [0-90mn]. For each mask, Pixel GT-CI-xx are set to 1, the other are set to 0.

Complementary additional Ground Truth may be used with lightning data (enlarged cumulated strokes during a period). Since this kind of data remains punctual and sparsed, it is used within a Complete Ground Truth (CGT) with all sources of data. It is foreseen in the optimal tuning configuration to consider ground truth convective activity from:

- Radar (>30dBZ) convective masks
- RDT convective masks
- RDT + radar (>30dBZ) + lightning strokes convective masks

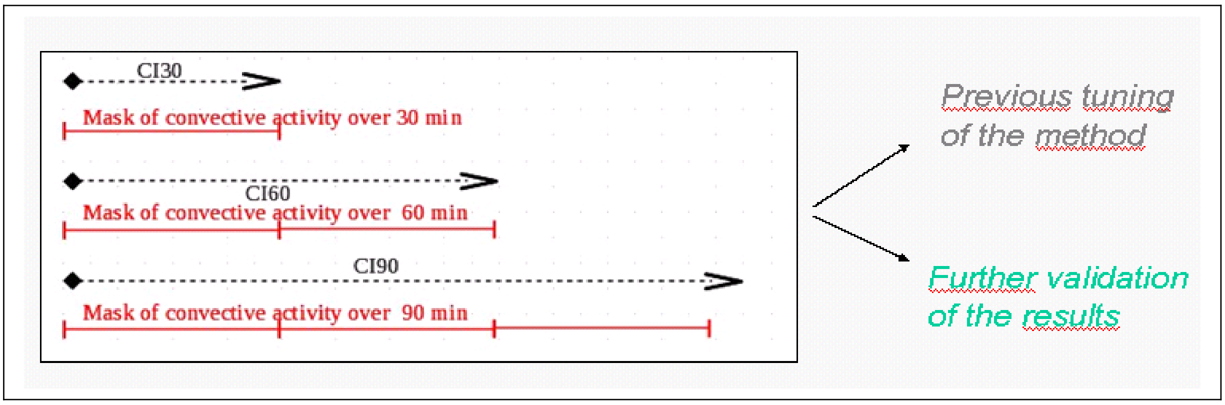


Figure 6: pre-CI pixels vs different Ground Truth masks

Interest field values of ***pre-CI*** pixels of a given slot are then analysed regarding those ground truth masks. With an approach based on a set of appropriate dates / situations and features, relevant thresholds of interest fields are identified.

As shown in the Figure below for the [0,30min] period, the spatial distribution of CI information (pre CI pixels and CI diagnosed pixels) often reveals scattered pixels or groups of pixels when compared to the signature of convection with selected ground truth (radar data in that case). Most of the positive GT regions are found in areas of cold cloud masses. These masses probably include deep convective clouds.

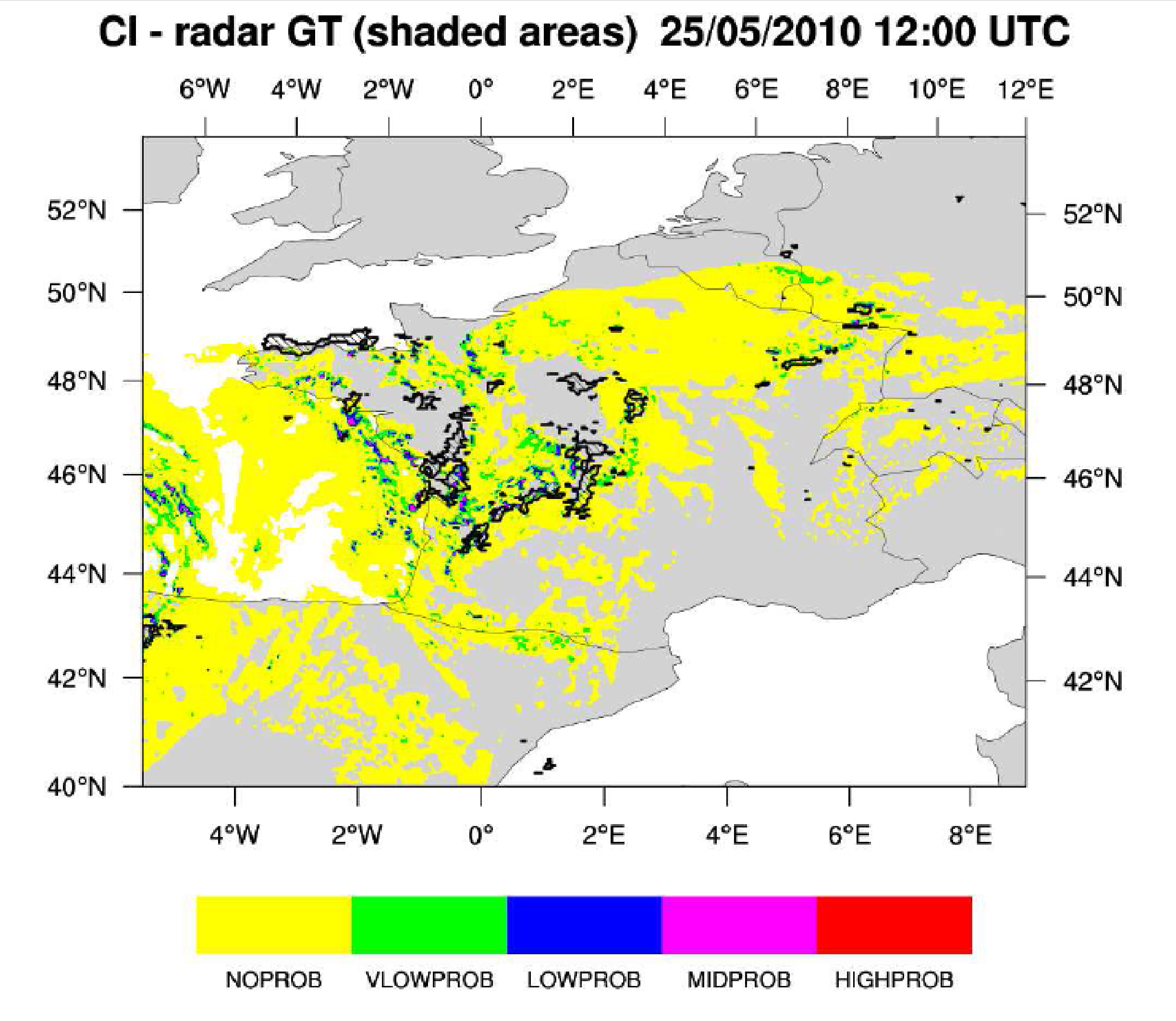


Figure 7: [0,30min] CI diagnosis vs corresponding 30min Radar Ground Truth

Such different kinds of information will obviously lead to difficulties to assess objectively the CI product, since low scores will be obtained with numbers of no detections and false alarms. Radar signature appears not only where convection initiation is expected, but also on already developed convection.

The convection signatures with RDTGT shows much more extended information, a fact rather expected due to the nature of RDT product (most of the time already developed convection) and the algorithm used for the computation of the mask. It is important to note that in most cases the RDTGT includes the RGT.

This is the reason why it was decided to focus on cases studies over limited interesting areas, where we expect the use of a CI product has sense. Moreover, the Radar Ground Truth seems the most appropriate to analyse regarding the scattered information that results from pre-CI and CI diagnosis steps.

An example of interest field analysis vs ground truth is illustrated below for a relevant region of a given meteorological situation. Such an extended approach allowed to assess or tune more precisely the thresholds for CI diagnosis.

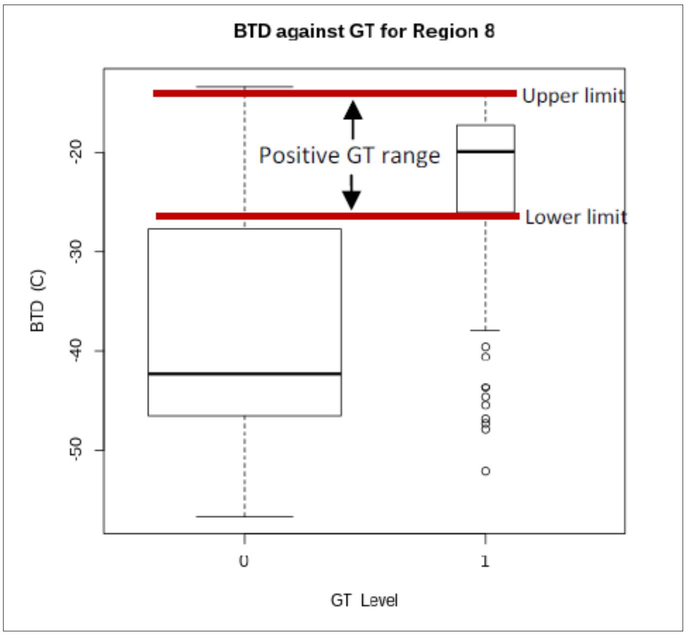


Figure 8: Analysis of an interest field parameter (here BTD) from pre-CI pixels on a selected region vs Radar Ground Truth. Ranges of values are assessed or tuned through this approach.

The same approach, over the different periods, taking into account the movement field applied to the pre-CI pixels, is then applied.

Finally, a dozen of events did constitute the data set for tuning, selected from 2010 cases study where some useful additional v2016 products had been made available. It leads to assess initial IF (interest fields) values listed in Table 7, and to tune day/night approach through Table 6 & 8 values.

Note: a final step will check that the probability value of CI for the range [0-60'] is above the probability of CI for the range [0-30'] and below the probability of CI for the range [0-90'].

2.2.2.2 Quality Control and Diagnostics

GEO-CI doesn't process real time quality control on tracking or diagnosis result. Sanity checks are in place to make, for example, the speeds realistic. But it is not used to create a quality control diagnosis.

2.2.2.3 Exception Handling

In case of missing satellite images, some error messages inform the user and CI fully recovers its quality few images later. Nevertheless, it takes into account the flag quality of CT optional input product in the CI pixel-level flag quality (ci_quality container), and fills a ci_status_flag container taking into account input data processed and used.

Moreover, the CI software produces some error messages in exception cases.

2.2.2.4 Outputs

The content of the output in NetCDF format is described in the Data Output Format document ([AD.8]). The product is an image-like product, whose target structure content three main containers dedicated to the three specified periods [0-30mn], [0-60mn] and [0-90mn]. All ci_prob containers have same structure.

Table 9: Output of CI product

Container	*Content*
ci_prob30	*NWC GEO CI Probability next 30 minutes* *Class 0: no probability to become thunderstorm* *Class 1: 0-25% probability to become thunderstorm in the next 30minutes* *Class 2: 25-50% probability to become thunderstorm in the next 30minutes* *Class 3: 50-75% probability to become thunderstorm in the next 30minutes* *Class 4: 75-100% probability to become thunderstorm in the next 30minutes* *FillValue: No data or corrupted data*
ci_prob60	*NWC GEO CI Probability next 60 minutes* *Same classes and meaning than for ci_prob30, but referred to the next 60 minutes*
ci_prob90	*NWC GEO CI Probability next 90 minutes* *Same classes and meaning than for ci_prob30, but referred to the next 90 minutes*
ci_quality	*Value 8: good* *Value 16: questionable* *Value 24: bad*
ci_status_flag	*6 bits indicating (if set to 1)* *Bit 0: High_resolution_satellite_data_used* *Bit 1: Visible_data_used* *Bit 2: IR3.9µm_data_used* *Bit 3: Cloud_type_data_used* *Bit 4: Cloud_Microphysic_data_used* *Bit 5: NWP_data_used*

2.2.2.5 Assumptions and Limitations

This version of the product offers a global approach with most relevant parameters which have been previously identified by others studies for this topic, completed by microphysics parameters during day-time and a specific subjective tuning for night-time. The objective was not to clone the previous approaches, but to offer a different use of the output for nowcasting purposes, with the wish to take into account uncertainty of the forecast through probability information. Moreover, the mix between a pixel-approach and object approach aims to catch as much as possible a reliable movement field to complete longest forecast ranges.

The tuning of CI relies on the availability of ground truth. With radar data (and for complete Ground Truth with lightning data), tuning is limited to areas covered by the corresponding networks.

3 DESCRIPTION OF RDT-CW PRODUCT

3.1 Product Overview

3.1.1 Name of product

RDT-CW is a name gathering RDT concept, as object analysis and convective objects identification, and CW for Convection Warning as a very short term forecast of the former. The aim is to distinguish three products linked to convection: RDT-CW (Convection Warning), RDT_CTRAJ (description of convective trajectories), CI (convection Initiation). In this document name RDT is sometimes still used to describe versions prior to CDOP3.

3.1.2 Goal of the RDT-CW product

The RDT-CW product has been developed by Météo-France in the framework of the EUMETSAT SAF in support to Nowcasting. Using mainly geostationary satellite data, it provides information on clouds related to significant convective systems, from meso scale (200 to 2000 km) down to smaller scales (tenth of km). It is provided to users in the form of list of numerical data stored in an output file (no image file). The objectives of RDT-CW are:

- The identification, monitoring and tracking of intense convective system clouds
- The detection of rapidly developing convective cells, where IR sensor allows for
- The forecast of the convective cells

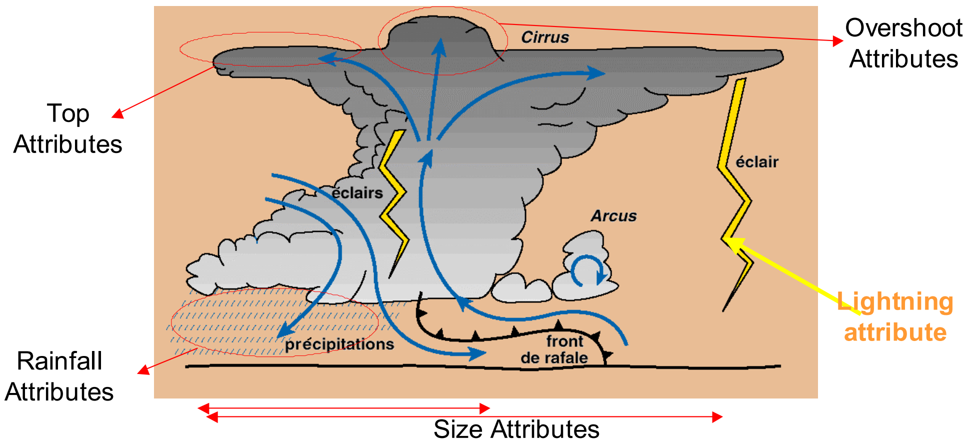


Figure 9: RDT-CW scheme

The object-oriented approach underlying the RDT-CW product allows to add value to the satellite image by characterizing convective, spatially consistent, entities through various parameters of interest for the forecaster such as motion vector, cooling and expansion rate, cloud top height,…, and their time series. It supports easy and meaningful downstream data fusion (surface observations, NWP fields, radar data, etc.).

Thereby, RDT-CW is a tool for meteorological forecasters but can also be used by research teams and end-users like aeronautical users.

3.1.3 The four steps of the algorithm

The RDT-CW algorithm could be divided into four parts:

- The detection of cloud systems
- The tracking of cloud systems
- The discrimination of convective cloud objects
- The forecast of convective cloud objects

Detection, tracking and discrimination can be grouped together in "analysis part".

3.2 RDT-CW DESCRIPTION

3.2.1 Theoretical description

3.2.1.1 Physics of the problem

RDT-CW is a mix of physical and statistical approach. The methodology is first to identify and track cloud system, then to define satellite characteristics of these cloud systems during different phases (triggering, development and mature). Learning databases are then built on the most significant parts of the trajectories of the cloud systems, for a pre-conditional tuning.

3.2.1.2 Description of the Algorithm

3.2.1.2.1 The detection of cloud systems

3.2.1.2.1.1 Main principle

Prior to any analysis, RDT-CW undertakes a cloud filtering using CT product (see 2.2.1.2.1.2). All non cloudy pixels are ignored through a dynamic cloud filtering. Since Cloud type information is further used as attribute to cloud cell objects, CT is rather used than CMA product.

The goal of the detection algorithm is to define "cells" which represent the cloud systems as seen in the infrared 10.8 µm channel. Once the "cells" are detected, a number of morphological (area, aspect ratio…) and radiative features (average and minimum brightness temperature, etc.) of the "cells" are computed in order to characterize the corresponding cloud systems. More precisely, "cells" are connected zones (8-connectivity) of pixels i) having a brightness temperature lower than a given temperature threshold T_{th} (which is not the same for all the "cells" detected in a given image) and ii) being larger than a given area threshold A_{min} (which is the same for all the detected "cells").

The use of a detection algorithm based on a fixed temperature thresholding is problematic. Indeed, the choice of a rather low temperature threshold leads to a late first detection of convective systems and the use of rather high temperature threshold leads to a merging of different convective systems into one single "cell" when these systems are embedded in a warm layer of clouds.

The RDT-CW detection method is based upon an adaptive temperature thresholding of infrared images. Thus, each cloud system is represented by one or several cells defined by its own, cell-specific, temperature threshold, ranged between a warm threshold T_{warm} and a cold threshold T_{cold}. More precisely, possible temperature thresholds are: T_{warm}, T_{warm}-ΔT, T_{warm}-2ΔT,... T_{cold} where ΔT is the temperature step of possible temperature thresholds.

RDT-CW cells point out the bottom of cloud towers included inside cloud system. The temperature threshold used to define the bottom of an RDT-CW object is the warmest one which allows to distinguish it from others nearby temperature extremes.

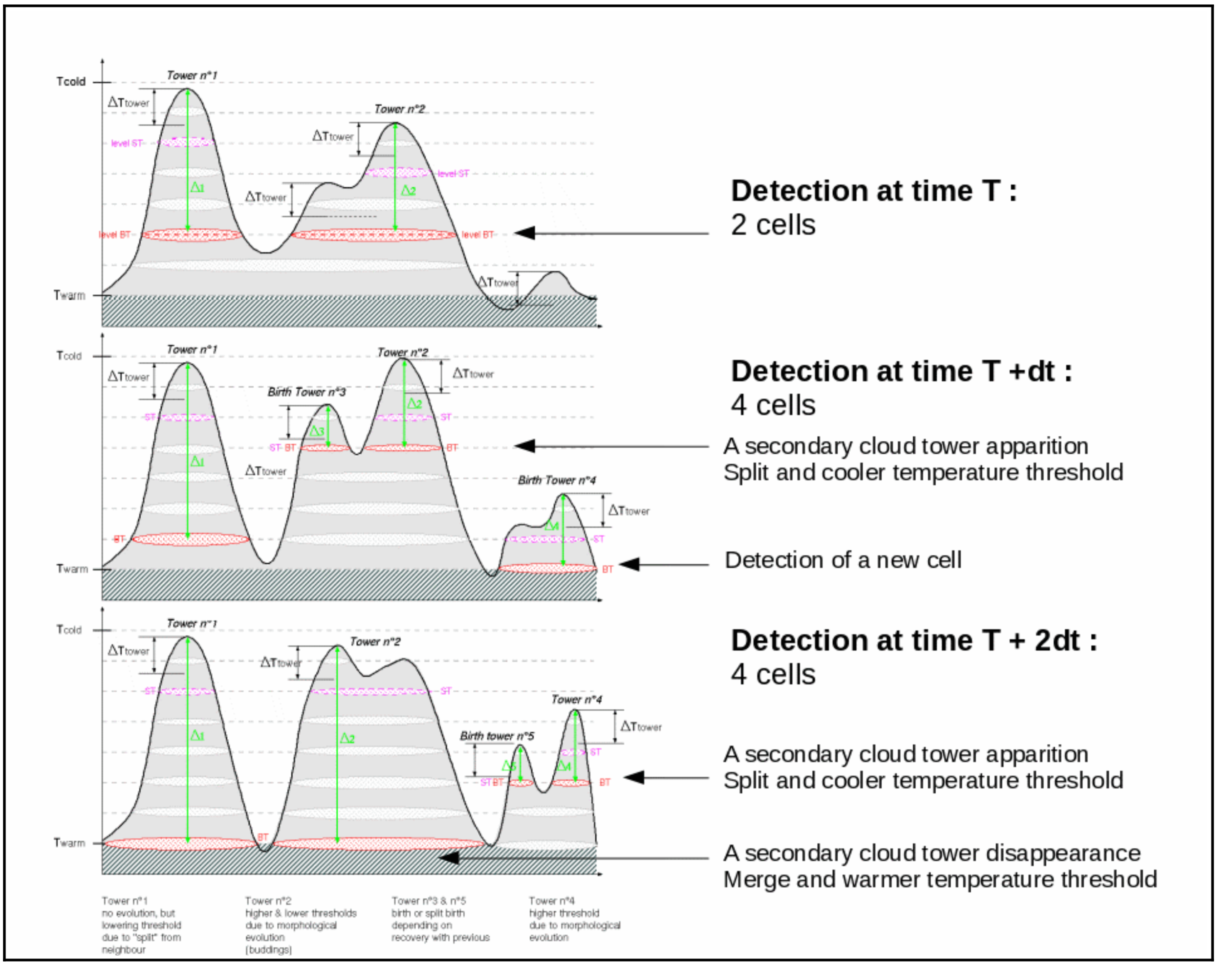


Figure 10: the principle of the detection algorithm for 3 successive slots with morphological evolution

3.2.1.2.1.2 BT limitation or correction for some special cases

3.2.1.2.1.2.1 Justification

In order to increase the relevancy of the cloud contour, modifications have been introduced. These modifications concern the choice of the value for BT.

- On one hand, before becoming mature, some cloud cells could sometimes exhibit too large BT contours regarding the top extent of growing tower. The need of a contour more focused on growing tower was satisfied by comparing areas of Base and Top of towers.
- On the other hand, aeronautical users have recently specified the need of cells' contour including as much as possible whole mature cloud systems, especially in tropical regions. This appeared essential for a correct planning of Cb avoidance. This specification has been satisfied in v2018 by introducing a configurable and an absolute temperature threshold for the BT level.

3.2.1.2.1.2.2 Outline representativeness of growing towers

Corrections are applied when both following conditions are satisfied

- The vertical morphology of the cloud presents particular shapes, with cloud systems defined by only one tower.
- The representative threshold, automatically set to the warmest value, leads to a bottom contour much larger than the horizontal extension of the top tower itself in its coldest part.

In case of a single budding rising up from a cloud layer, the algorithm described in previous paragraph identifies the bottom of the layer rather than bottom of the tower. The modification consists in detecting, between BT and ST levels, the level of maximum vertical rate of cloud cell area. The algorithm considers the vertical area ratio between two successive temperature levels S_{n+1}/S_n. The goal is to catch a more "realistic" bottom of tower (red level in Figure 11 left). A value of 0.7 for surface ratio between bottom and top levels has been tuned and used to represent a more realistic bottom of tower, particularly for convective systems.

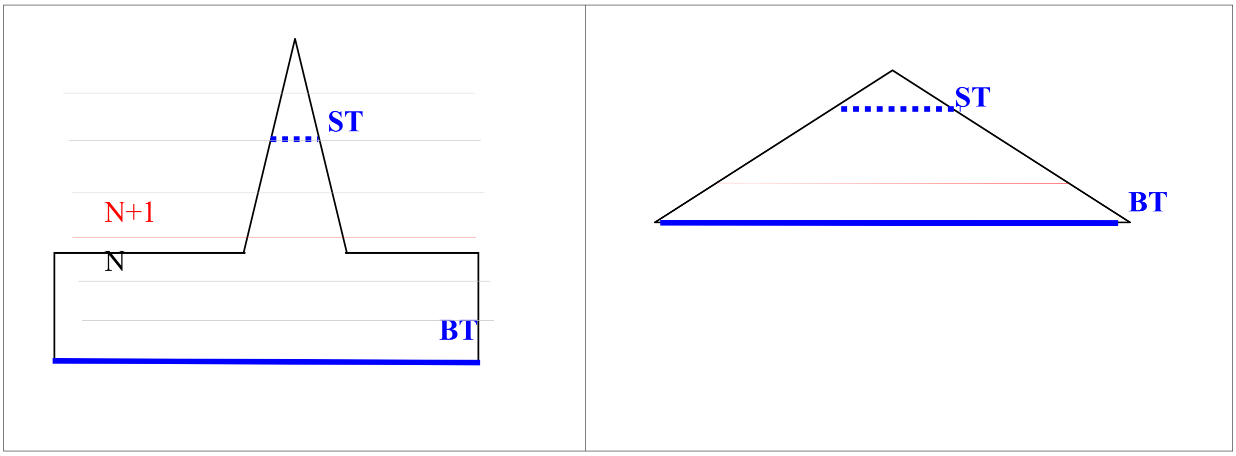


Figure 11: Left: Single tower from cloud layer. Right: Pyramidal single tower. Bottom (BT) and Top (ST) levels of Tower are represented

In case of a "flat pyramidal" shape of single-tower cloud system, here again algorithm described in previous paragraph catches the bottom of the whole system, resulting in some cases to an excessive difference between top and bottom horizontal extensions.

An upper threshold of the ratio between BT and ST areas is necessary. A ratio of "3" gave the best subjective correction of BT level.

3.2.1.2.1.2.3 Contour level limitation of mature cloud systems

The principle is first to take benefit from NWP data and tropopause temperature field (provided or re-computed) to introduce an absolute temperature threshold limitation (not configurable) for BT level: BT temperature can not be colder than the threshold. The advantage is a dynamic and coherent limitation of cloud contours from a meteorological point of view.

Additional configuration is proposed for tropical regions or in warm air masses with high tropopause levels. Indeed in these configurations, as tropopause can be very high the previous limit will not be active. An optional, additional and configurable limitation has been introduced to keep hand on BT contour of mature cloud system. Several thresholds have been tested in different regions and periods, from -40° to -70°C. Developers of the product initially recommended a value of -60° or warmer, but it really depends on situations and regions. Figure 12 below illustrates such an implementation.

Aeronautical users' feedback lead to adapt again this threshold to a colder value of -70°, for a better adaptation to huge and cold MCS over Africa. This value of -70° will be the default initialisation for BT limitation.

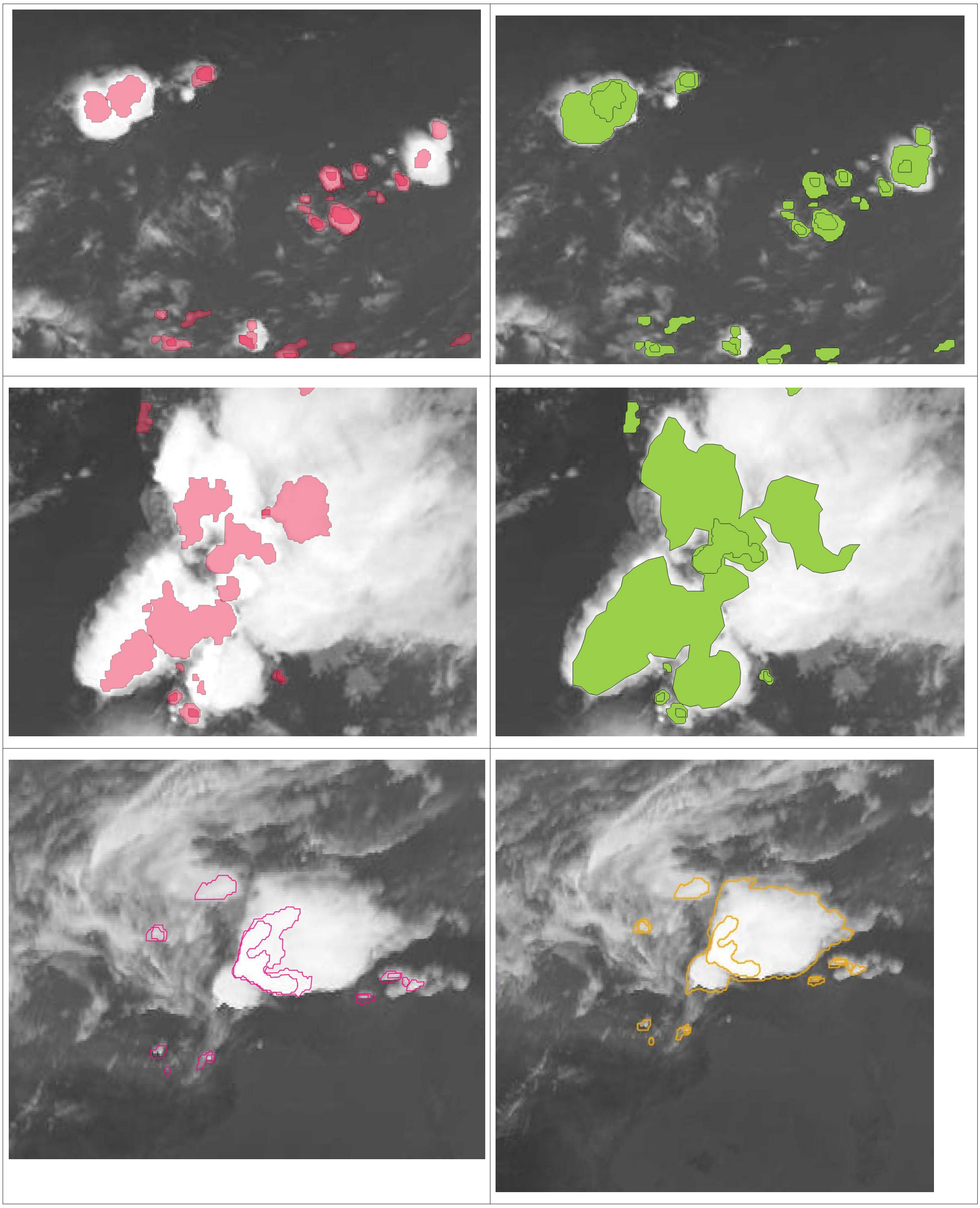


Figure 12: BT limitation of mature cloud systems. Without BT contour limitation (left), with Tropopause T° and -60° threshold limitation (right)

It is to note that BT limitation can induce cloud system with several towers inside. In that case, only the coldest will contribute to ST level, ignoring other buddings.

This approach is implemented in the algorithm, leading to:

- Some visual improvements of RDT-CW cloud contours
- An increase of the homogeneity of the convective population of mature systems

3.2.1.2.2 The tracking of cloud systems

Once the detection of cloud systems is performed, the tracking module of the RDT-CW software is applied on the detected "cells" and allows building trajectories of cloud systems from a sequence of infrared images. The tracking algorithm is based on the geographical overlapping of "cells" between two successive infrared images (Figure 13). It also handles splits and merges of cloud systems.

One main input of the tracking of the cloud system is the previous moving speed estimation of cloud cells. Main results are the current moving speed estimation, and temporal links with previous cells.

For that reason RDT-CW software now pre-calculates a movement guess field to consolidate this approach.

A 2D movement field is estimated in optimum configuration with last available HRW wind observations. Those observations are filtered (selection of IR channels observations, using the corresponding valid pixel's brightness temperature to keep one obs per pixel), and are remapped on the satellite grid-field region, through a weighted interpolation taking into account a radius of influence and a cut-off distance.

An additional blending with low level (700hPa) NWP wind field can be optionally processed where no information is available, but priority is given to HRW wind observations. This "guess" field allows:

- Initializing moving speed with "cold start" cases (first run), which will be useful for the next runs
- Initializing moving speed for cells with no recovery and no neighbouring cells, used for a retro-advection and checking recovery again

The main difficulty is the tracking of small cloud systems (typically less than 5 pixels). In order to improve the tracking of such small cloud systems, the RDT-CW tracking algorithm takes into account an estimated velocity of "cells" to compute the overlapping between "cells".

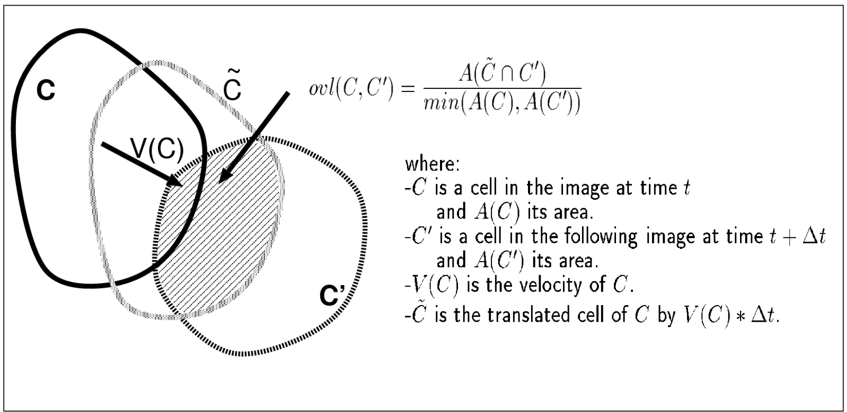


Figure 13: Definition of the overlapping between two cells

The search for an overlapping between a cloud system C' detected in the image at time t+Δt and a cloud systems C detected in the previous image at time t is described in Figure 14 and hereafter.

First, "cells" in the image at time *t* are advected using their estimated velocity. If at least one of these advected cells overlaps sufficiently with cloud system C' then a link is created between C' and this (these) cell(s).

Then, moving speed of cloud system C' is estimated from gravity centres displacement of all linked cloud cells. When no overlapping has been found for cloud system C', then its velocity is evaluated from cross-correlation technique (more costly), neighbouring speed (less costly), or pre-calculated movement guess field.

The cloud system C' is then backward-advected from this estimated velocity. If at least one of the cells detected at time t overlaps sufficiently with the backward-advected cloud system C' then a link is created between C' and this (these) cell(s).

If no overlapping is found, then the backward-advected cloud system C' is enlarged and a last search for overlapping between this enlarged backward-advected cloud system C' and cells detected at time t is done. If at least one of the cells detected at time t overlaps sufficiently with the enlarged backward-advected cloud system C' then a link is created between C' and this (these) cell(s).

If no overlapping is found then cloud system C' is identified as the beginning of a new trajectory.

Figure 15 illustrates how the steps 2 and 3 of the tracking algorithm could improve the tracking of small cloud systems. In the diagrams of this figure, the "cells" of a given cloud system in two consecutive images are showed: C is its "cell" in the image at time *t* and C' is its "cell" in the image at time *t+Δt*.

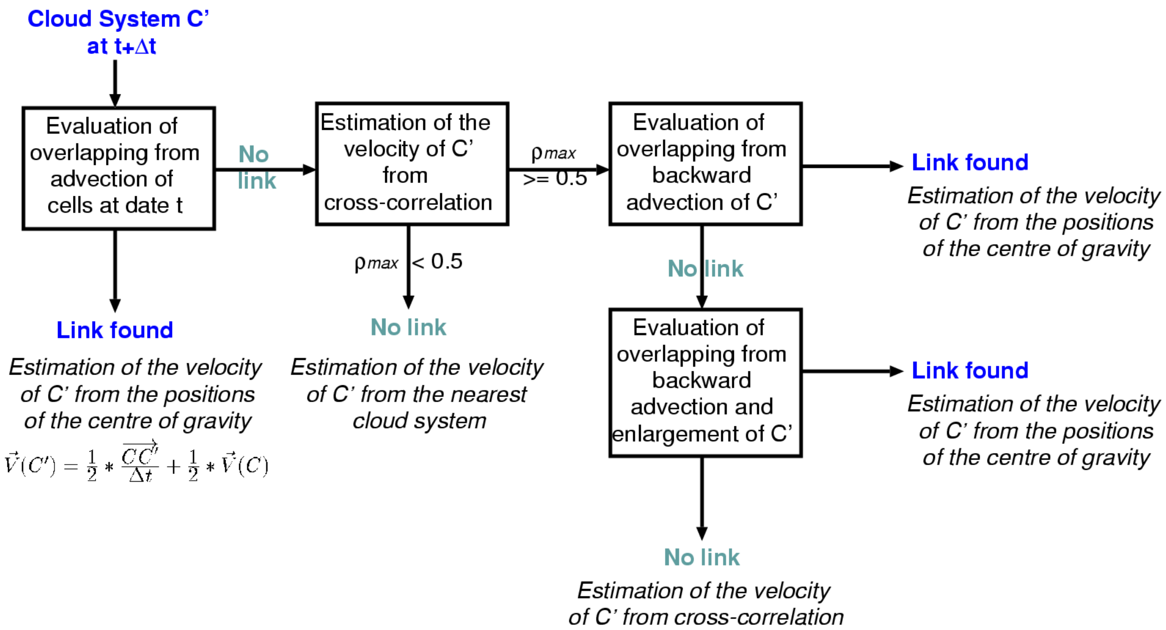


Figure 14: Main steps of the tracking algorithm. In the second step, ρ_{max} *stands for maximum correlation coefficient from cross correlation process. This step can be replaced by neighbouring speed or pre-calculated movement field.*

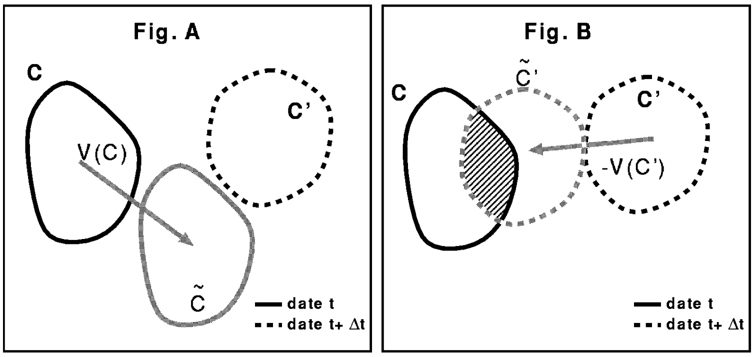


Figure 15: Principle of the tracking algorithm (steps 2 and 3). Plain line for reference date t (obs and translated), dashed for reference date t+Δt (obs and retro-advected)

In Figure 15 A, C̃ is the translated cell of C by a distance and direction that depend both on the motion vector estimated in previous tracking phase and Δt. In this case, the quality of the velocity was too low and leads to no overlapping between C̃ and C'. So, after step 1 of the tracking method, no link is created between "cells" C̃ and C' and so, if steps 2 and 3 were not in the tracking algorithm, the tracking of this cloud system would have failed.

The following analysis is then done: C' is a "cell" in the image at time *t+Δt* which overlaps with no "cell" of the previous image, consequently its velocity is evaluated using a cross-correlation technique.

Figure 15 B displays the cell C̃' which is the translated cell of C' by a distance and direction that depend both on the motion vector estimated thanks to cross-correlation and Δt: an overlapping exists between C and C̃'. Thus, the tracking algorithm creates a link between "cells" C' and C: the tracking is successful.

Step 4 of the tracking algorithm is an improvement for the tracking of very small cloud systems (less than 5 pixels). The enlargement of a cloud system consists of adding "pseudo-cloudy pixels" (see Figure 16) to the detected cells all along its edge in order to increase, artificially, the size of the cell and then to ease the occurrence of overlapping between consecutive cells corresponding to the same cloud system.

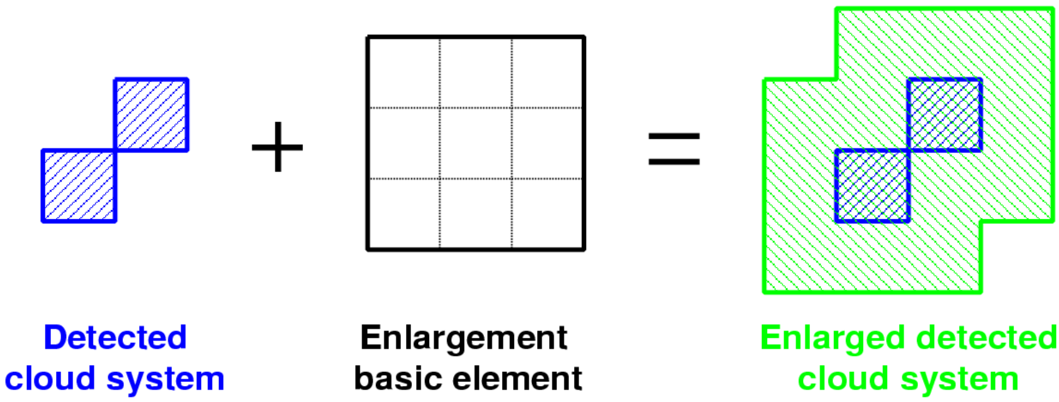


Figure 16: Principle of the enlargement of cloud systems (step 4)

3.2.1.2.2.1 Motion vector estimation from weighted gravity centres

As mentioned above, once links between cells in successive images have been identified, displacement is estimated from the successive positions of corresponding gravity centres. More precisely, displacement is processed taking into account all current and previous linked "group" of cells, weighted by the size of each cell. Thus, an estimation of "family" cell movement is done.

Previous chapter has pointed the particularity of cloud cell identification with temperature threshold. This threshold defines the 2D geometric extension of cloud cell objects. Two kinds of gravity centres are calculated from this 2D representation during detection step:

Geometric gravity centre: mean of latitude and longitude of each pixel constituting the 2D horizontal extension at the temperature threshold defining the cloud cell

Weighted gravity centre: position of each pixel is weighted by its level (temperature), giving higher weight to higher levels (lower temperatures). The resultant localization is more representative of the 3D cloud morphology.

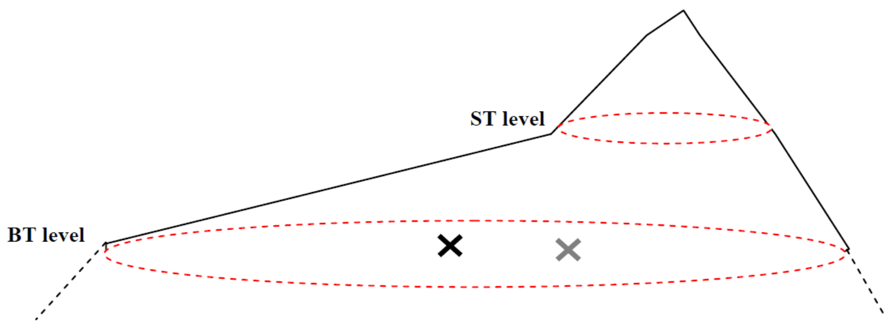


Figure 17: Positions of geometric and weighted gravity centres for a particular shape of cloud system

Motion vectors are now estimated by default using displacement of weighted gravity centres positions in successive images. This approach tends to lower the impact of temperature threshold changes of a cloud cell from one image to the next one. The figure below illustrates the kind of overestimation which can result of an estimation based on geometric gravity centres displacement in case of changes of cell's temperature threshold.

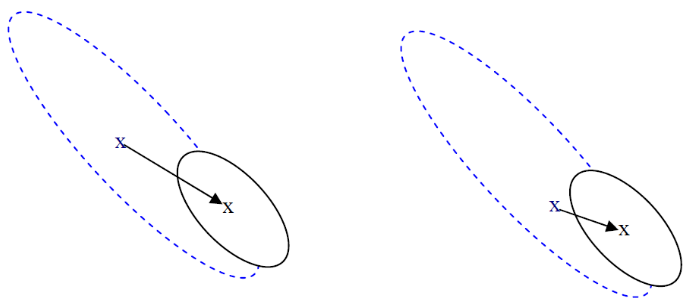


Figure 18: Synthetic example of speed estimation from geometric gravity centres (left) and weighted gravity centres (right). Previous cell dashed blue, current cell plain dark.

In a further step, a temporal coherence of moving speed is undertaken, to eliminate erratic estimation due to split/merge cases or due to different temperature thresholds and thus too large distance between gravity centres.

Improvements are illustrated in Figure 19. Some problems are identified in v2013: one excessive speed (30m/s) southern Italy (North Est part of the image), one excessive speed and opposite direction South of Sicilia (South West part of the image). The problems are corrected in later release. That improvement leads to suppress erratic values of speed or direction of motion due to merges or splits along cloud system trajectory/life. They also allow increasing horizontal coherence with neighbouring cloud systems.

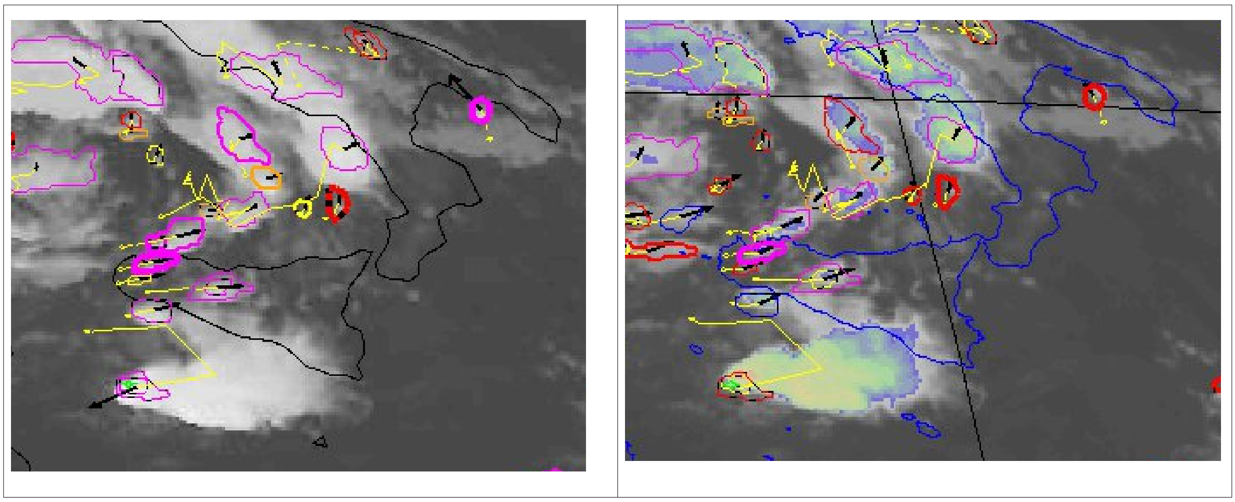


Figure 19: 11th August 2015 07h00 UTC slot. Left: RDT without temporal coherence. Right: RDT operated with NWP and HRW, and temporal coherence of cell's motion

3.2.1.2.3 The discrimination scheme

3.2.1.2.3.1 Main principles

The discrimination scheme relies on statistical models, which have to be previously tuned. Those models are provided with various parameters and trends associated to a cloud system. Electric data have been used as ground truth during this previous tuning, allowing to recognize mature thunderstorms, and identify their life cycle.

On statistical approach, the two populations, convective and no convective, are unbalanced. We can notice a ratio of one convective for more than one hundred non convective over Europe.

That is why, like for Convection Initiation product, NWP data are used to process a "NWP convective mask" as a first step prior to discrimination step, and ignore cloud systems located in not relevant regions (see chapter 2.2.1.2.1.1 , and the possibility to configure the threshold values of convective indexes to enlarge or reduce the area of interest). Thus, discrimination scheme avoids eventual false alarms, especially in winter or intermediate seasons. With this release RDT-CW rather focuses on unstable regions defined by NWP mask, and ignore stable and unclear areas (NWPCONV 1 argument in model configuration file). This approach considerably lowers false alarms during intermediate and winter seasons over mid-latitudes, without any noticeable impact for other periods or regions. Note: CI product takes into account those unclear areas defined by NWP mask, thus RDT-CW complements the analysis but focuses on unstable areas.

Moreover, this approach has been applied during the tuning of discrimination scheme, to exclude areas and cloud systems without interest from a convective point of view, reducing the unbalance mentioned above, and improving statistical approach.

A convective object has not homogeneous characteristics during its life time. Thus, it is necessary to define several statistical bodies in order to take care of various stages of convective phenomena: triggering, development, mature and decaying phases.

That is why a vertical stratification is undertaken, with the definition of categories, linked to the minimum temperature of the cloud system. Thus, a cloud system is considered only if it reveals an ascending trajectory, from lowest to highest categories.

- Mature: top temperature < -40°C since at least 30min

- Mature transition: crossing top temperature –40°C
- Cold transition: crossing top temperature –35°C
- Warm transition: crossing top temperature –25°C
- Warm2 transition: crossing top temperature –15°C
- Warm: top temperature > -15° preceding Warm2 crossing

At last, the ground truth used, electrical activity based on cloud to ground or intra-cloud occurrence, doesn't allow to diagnose the time of convection triggering or to depict the decaying period.

Therefore, the discrimination scheme is a mix between statistical decisions and empirical rules. The statistical decisions are processed for a short historical period preceding current slot, and depending on several periods of interest. Statistical decisions are only applied on no convective object to check their convective status. Once a positive decision has been set, the cloud system will inherit from this characteristic for the following slots, as long as the stage is growing. Empirical rules are then defined to declassify convective object (convection decaying or false alarm diagnosis). They are based on cooling parameters for triggering and development phase and based on cooling and global convection index for mature phase.

Figure below illustrates those main principles of discrimination scheme.

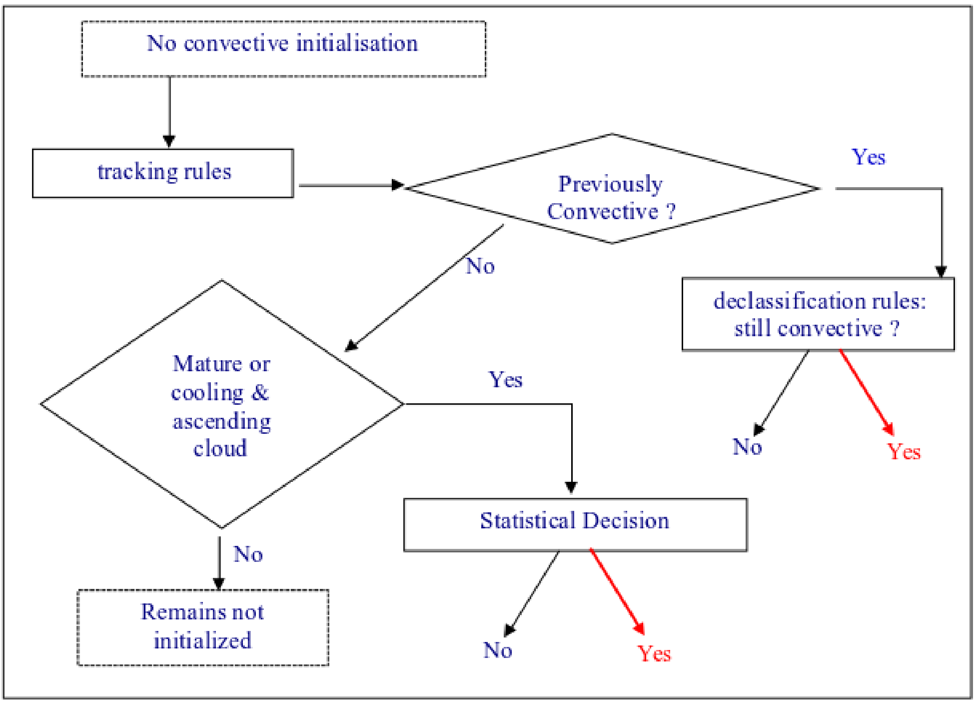


Figure 20: The progression of discrimination scheme

The initial statistical models, previously tuned with MSG over France, remain available in this version of RDT-CW, as generic (GEN) discrimination, since elaborated with a high-quality lightning network, and fully validated over Europe.

v2018 version has been enriched with an additional discrimination, whose statistical models have been slightly modified, and calibrated for several geostationary satellites and corresponding update rates, using a global lightning network as ground truth. This new approach, labelled calibrated (CAL) discrimination, will be the default one for satellites which have been calibrated with.

The methodology and the statistical approaches are presented hereafter.

3.2.1.2.3.2 The generic (GEN) scheme

3.2.1.2.3.2.1 The generic (GEN) discrimination

The statistical decision relies on an ensemble of logistic regressions, here named statistical models, tuned and applied for various steps of cloud development, and various configurations. The objective is to distinguish the convective cells from the other one's.

The discriminating parameters associated to a cloud object are processed from several MSG channels (IR 10.8µm, IR8.7µm, IR12µm, WV 6.2µm and WV 7.3µm).

Moreover, the cloud tracking allows to estimate rates and extremes on various past period.

Discriminating parameters are averages, standard deviations and extreme values over time-depths during the recent history:

- Main IR (10.8 µm or equivalent) channel:
  - Values: the minimum value (of all the pixel at each step), maximum value of the following parameter defined at each step "average minus minimum"
  - Trends: various values are taken into account: instantaneous (between the current slot and the previous one), average and standard deviation over the time-depth
  - Morphology: maximum value of following parameters defined at each step: average and extreme spatial gradients at the outline of the cell, volume, surfaces, ratio aspect
- Other channels
  - BT: extreme values over the cell are estimated and various trends are processed, and then included in predictor set
  - BTD representative values are estimated and various trends are processed, and then included in predictor set

Statistical models are defined for each vertical category. Those based on temperature threshold crossing are named transition models. The models defined on mature population are named mature models, and those defined on warm population warm models.

Each model is defined for a recent past history, from 15 up to 60 minutes, depending on the availability.

In order to provide a classification for several configurations, applicable to various geostationary cases, models are defined for following configuration cases, or modes:

- "Classic with NWP" configuration (6.2µm+7.3µm +8.7µm +10.8µm+12.0µm + NWP data)
- "Classic" configuration (6.2µm+7.3µm +8.7µm +10.8µm+12.0µm)
- "Limited with NWP" configuration IR10.8µm+WV6.2µm + NWP data
- "Limited" configuration IR10.8µm+WV6.2µm
- "Mono channel with NWP" configuration IR10.8µm+ NWP data
- "Mono channel" configuration IR10.8µm

It is to note that even if the user's configuration file does not correspond to the real time availability of data, RDT-CW is able to adapt and detect automatically the best usable configuration among the ones listed above. For that reason, each mode has benefit from a specific tuning.

To summarize, the scheme is based on discriminations defined on crossing times (vertical transition categories) and on specific mature and warm states. Each discrimination makes use of available historic data (from 15 minutes up to 60 minutes). Thus, the discrimination scheme rests on several tens of models (logistic regressions). Each model has to be specifically tuned on learning databases. For that reason, the description of discrimination tuning in this document will give only a quick overview of the results obtained.

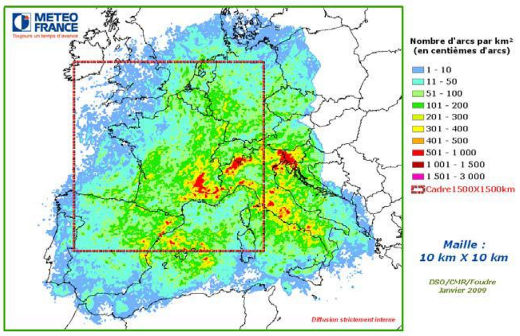
Preliminary studies had been led to assess convective discriminating skill of linear and no linear models to be incorporated into the discrimination scheme. The best results were obtained with random forest method (with 600 trees) and simple Logistic Regression method. The logistic regression had been implemented into operational version. This model is simple and fast, and provides easily a way to estimate the error of diagnosis.

3.2.1.2.3.2.2 The generic (GEN) statistical model tuning

During the discrimination tuning, 15 to 60 minutes sections centred on transition time are extracted from cloud system trajectories, and models (logistic regressions) are defined on various depth, respecting the way those models are planned to be used in real time: the choice of a model corresponds to a choice of depth, based on age in the category, age of first detection, and past historic in the warmer category. Warm category time of reference for tuning is defined relative to Warm2 transition time. Mature category tuning relies on an analysis over a longer historic.

The data used for tuning were June-August 2008 and June-September 2009, for both MSG Full Disk (MSG02 at that time) and MSG Rapid Scan (MSG01 at that time), and corresponding NWP data from Météo-France ARPEGE model, for 12h and 18h ranges (as for real-time use).

The domain used for the tuning has been a little bit widened, to take into account statistics of lightning data. Data has been provided by Météo-France Observation Department and concerns Météorage and partners network (Euclid). An accuracy of 2-4 km of detection has been taken into account.



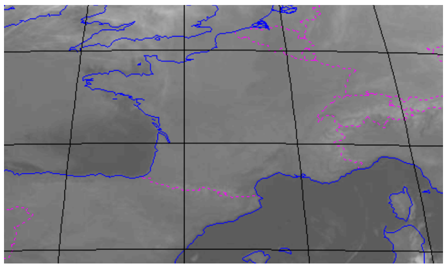


Figure 21: Left: Extension and quality (density) of Meteorage+partners network. Meteorage network coverage appears in red. Right: Domain used for GEN discrimination tuning

With numerous 2008 and 2009 data, the discrimination tuning statistical method is more stable, with an increase of the stability (robustness) of statistical models.

The use of NWP data to exclude cloud systems in stable areas has allowed to reduce in the database the imbalance between electric and non electric systems. For that reason, the method has evolved in respect to “data mining” techniques: a large learning data set without modifications of the initial proportion of population.

The ground truth used rests on a moderate lightning data activity, even for mature and transition mature categories. But the proximity to lightnings has been taken into account to built a non polluted non-convective population, still decreasing the imbalance (non convective when 50 pixels far from flashes, i.e. about 150-200 km).

Cross validation method has been implemented to reduce the dependency to the learning data set. For each statistical model, the whole data base has been taken into account for a first tuning (except 4 weeks for a further independent and coherent validation) in order to obtain a selection of relevant parameters (predictors). The coefficients of these parameters have been then “adjusted” through the processing of fifty learning-validation steps, where learning and validation dataset where randomly chosen (with respect to a proportion 80%-20%). Thus, linear model will be less dependent on learning data set.

Finally, a validation step is undertaken on an independent data set: 4 weeks distributed among 2008 and 2009, 20080713-20, 20080901-08, 20090617-24 and 20090821-28.

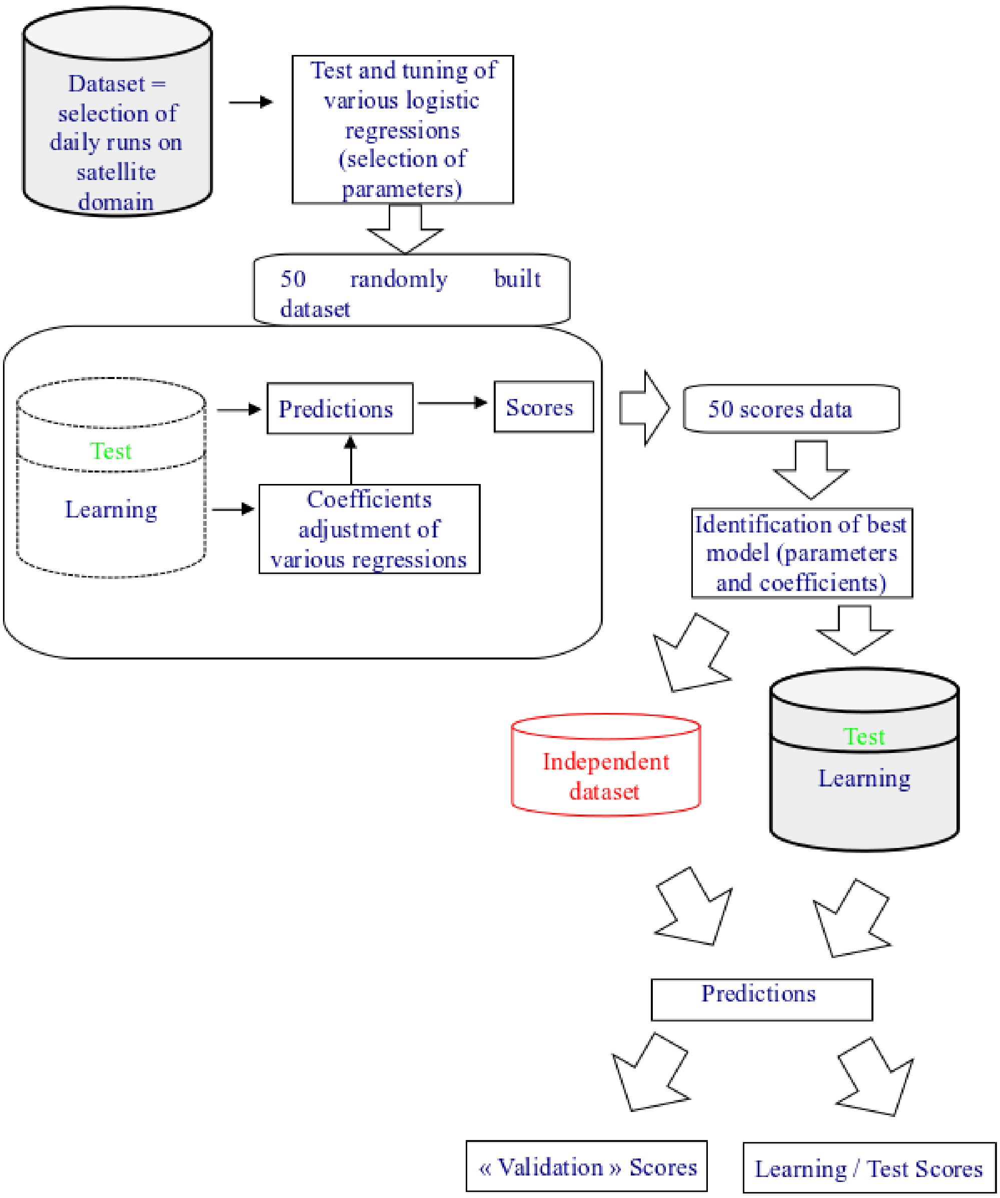


Figure 22: Discrimination tuning methodology.

The discrimination skill is depicted from threat score-false alarm distributions. This depiction allows to point out the inflexion point where the false alarm increases more than no detection decreasing, with respect to a maximum acceptable false alarm ratio (varying from 5% to 15% depending on distributions).

The threat score and false alarm distributions are displayed in graphs like figure below, where learning data set appears in black (80% of learning data set), random test data set in green (remaining 20%), and validation data set in red. Minimum value of (TS-FAR) is marked as a cross.

The automatic choice of decision threshold is made from red distribution, taking into account this marked point and a maximum value of FAR.

It has to be noted that tuning scores and graph will be dependent on several parameters:

- The use or not of a NWP convective mask as a first filter of clouds
- The tuning methodology, with the parameters used for cross-validation process (different manner to consider learning dataset, test dataset and validation dataset).
- The period of tuning: the larger it is, the more reliable tuning is
- The ground truth and flashes proximity, to define non convective population
- The area of tuning

The comparison with successive versions has to be undertaken on a subjective basis, using case studies and analysing real time situations, for a larger domain than the tuning-domain.

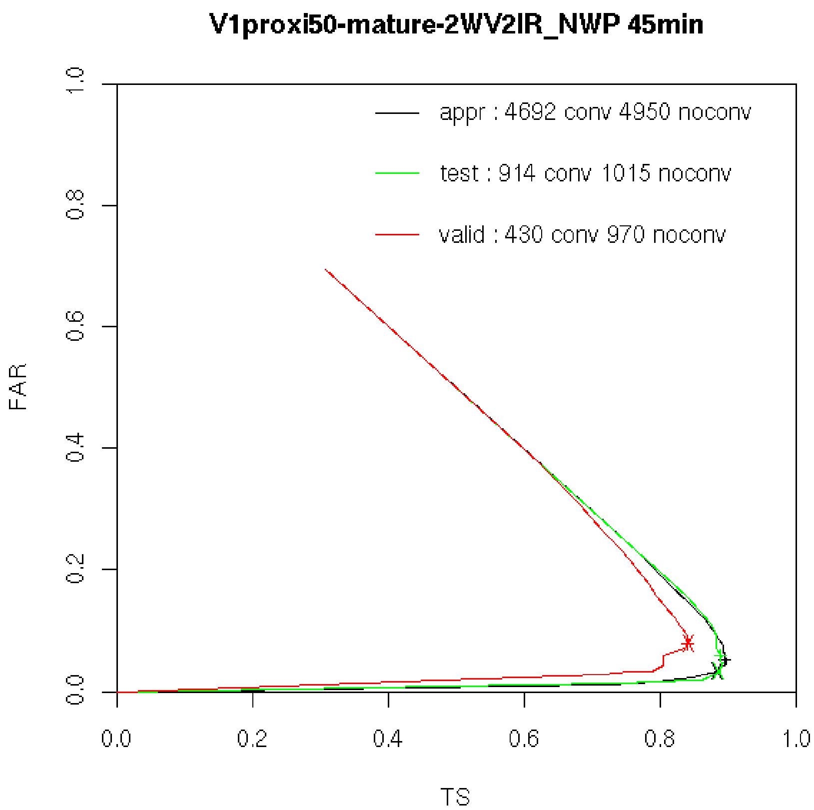


Figure 23: MSG v2011 tuning. TS/FAR curves for mature discrimination, full configuration 6.2µm+7.3µm +8.7µm +10.8µm +12.0µm +NWP, 45-min depth, for a moderate ground truth with proximity to flashes taken into account. Learning database (black), random test database (green), validation database (red)

The results issued from discrimination tuning allow to rank the configurations of RDT-CW upon scores, whatever the categories.

The automatic choice of configuration mode in the discrimination step, depending on available data (additional channels or not, NWP data or not), will respect this ranking.

Considering categories, the ranking upon scores is similar to the vertical ranking, as expected: cold categories offer better tuning and scores than warmest ones.

Considering past historic depth, larger depth most often get better scores than shorter ones, except for warmest categories, where significant signal is found and exploited even in shorter depth, with systems in ascending phase remaining few time in these categories.

Synthesizing all the results leads to invalidate some models, when they present higher false alarms.

All invalidated models are listed in a specific files in $SAFNWC/import/Aux_data/RDT-CW/files_for_discri/ConvCoeffRegr*mask, sorted by category/configuration/depth. This file is read as guidance in real time at the discrimination step.

3.2.1.2.3.3 The calibrated (CAL) scheme

3.2.1.2.3.3.1 The calibrated (CAL) discrimination

Calibrated discrimination has been set up to adapt to a given geostationary satellite and the corresponding satellite scan (update rate).

The approach is more or less the same as GEN discrimination, following those guidelines for defining statistical models:

- Same categories for vertical stratification of cloud systems
- Several modes (satellite configuration) taken into account
- Analysis of recent history with various depth

Modes and depths have been refined as detailed below.

Four modes are taken into account, among which 2 may be used during daytime:

- "IRWV" restricted mode, using one WV channel (equivalent to WV63) and main IR108 (or equivalent)
- "IRxWVx" default mode, using a panel of 5 IR channels (2 WV channels, main IR108 or equivalent, plus IR87 and IR120 channels)

  "main IR108 or equivalent": Channel IR104 can also be taken into account instead of IR112 for the GOES-16 / GOES-17 or HIMAWARI-8 / HIMAWARI-9 configurations.
- "IRxWVxJ" daytime mode, adding VIS06, NIR16 and IR38 channels to "IRxWVx" mode
- "IRxWVxJ2" daytime extra mode, adding NIR22 to "IRxWVxJ" mode (for satellites with available NIR22 channel like MTG, Himawari-8 or GOES16)

Like for GEN discrimination scheme, RDT-CW is able to adapt and detect automatically the best usable configuration among the ones listed above. The default "IRxWVx" mode will automatically switch to "IRxWVxJ" or "IRxWVxJ2" during daytime.

Five depth periods will be considered,:

- 0: for triggering cloud systems without history (first detection or split ...)
- C: short range history limited to [0-15min] or [0-20min] depending on satellite scan
- M: medium range history [0-30min] or [0-40 min] depending on satellite scan
- L: long range history [0-60min]
- VL: very long range history [0-80min] or [0-90min] depending on satellite update rate

"VL" depth will be taken into account for mature categories only, where "0" and "C" depth will only concern transition categories (i.e. neither mature nor warm categories).

The scheme is based on discriminations defined on crossing times (vertical transition categories) and on specific mature and warm states. Each discrimination makes use of available historic data (from 0 up to 90min), whatever the update rate. Thus, the discrimination scheme relies on several tens of statistical models (logistic regressions) for a given geostationary satellite. Each model has to be specifically tuned on learning databases. Here again, the description of discrimination tuning will give only a quick overview of the results obtained.

3.2.1.2.3.3.2 The calibrated (CAL) statistical model tuning

The same methodology than for GEN discrimination tuning (see Figure 22) are used, to define the statistical models. The characteristics of this tuning are the following:

- Various satellites and update rates concerned: tuning for MSG4 (Full disk mission) and MSG1 (IODC mission) in nominal mode (15 min update rate), MSG3 (rapid-scan mission), and GOES16 (GOES-E mission) for nominal 10 minutes scan rate. GOES16 at initial 15 minutes scan rate up to April 2019 was taken into account for specific tuning in former release.. Full 10 min update rate being not available in Météo-France for Himawari-8 and GOES17 it has been decided to benefit from GOES16 tuning for those two satellites, since they share similar characteristics.
- Ground truth used: lightning data have been used when detection efficiency and delay of availability of corresponding network/sensor were good enough in Météo-France operational process. METEORAGE lightning network was used for runs over Europe domain with MSG4 and MSG3-RSS. GLD360 lightning network was used for runs over a South Indian Ocean domain with MSG1-IODC. The tuning of RDT operated with 10min-GOES16 ABI data has been undertaken with GOES16-GLM flash data, converted in a ground network format compliant with [AD.6]. The quality index of those data is used to eliminate artefacts that had been previously encountered with this sensor. The tuning of RDT operated with 15 min GOES16 ABI data was undertaken with WWLLN data, due to GLM artefacts encountered during this period.
- Length of dataset: for each specified geostationary satellite, several months of operational v2018 RDT runs have been taken into account over a domain covered by lightning data.
- Previous processing of Cloud products (CMA, CT, CTTH, CMIC) on the specified domains, necessary for a full benefit to RDT-CW
- Depth ranges which can be adapted to various satellite updated rates, and a maximum depth for mature category extending to 90min
- Limited number modes specified, but extended to visible and near-infra-red channels for daily conditions
- A slightly limited set of predictors to limit redundancies and correlations

The threat score /false alarm distributions are displayed in graphs like figure below, where learning data set appears in black (80% of initial learning data set), random test data set in green (remaining 20%), and validation data set in red (independent dataset previously randomly extracted from initial dataset: 2/3rd for learning dataset, 1/3rd for validation).

For each dataset, the minimum value of (TS-FAR) is marked as a cross. Thus, for each statistical model, the automatic choice of decision threshold is mainly made from validation dataset scores, taking into account this marked point and a maximum admissible value of FAR. When scores are too low (for example in case of strong unbalance of populations or insufficient dataset length), the corresponding model is invalidated and listed in a specific auxiliary file.

It is to note that tuning scores and graph are dependent on several parameters:

- The tuning methodology, with the parameters used for cross-validation process as there are different options for defining learning dataset, test dataset and validation dataset.
- The period of tuning: the larger it is, the more reliable tuning is
- The ground truth and flashes proximity, to define non convective population
- The area of tuning (continental vs oceanic)

The comparison with previous versions has been undertaken on some case studies.

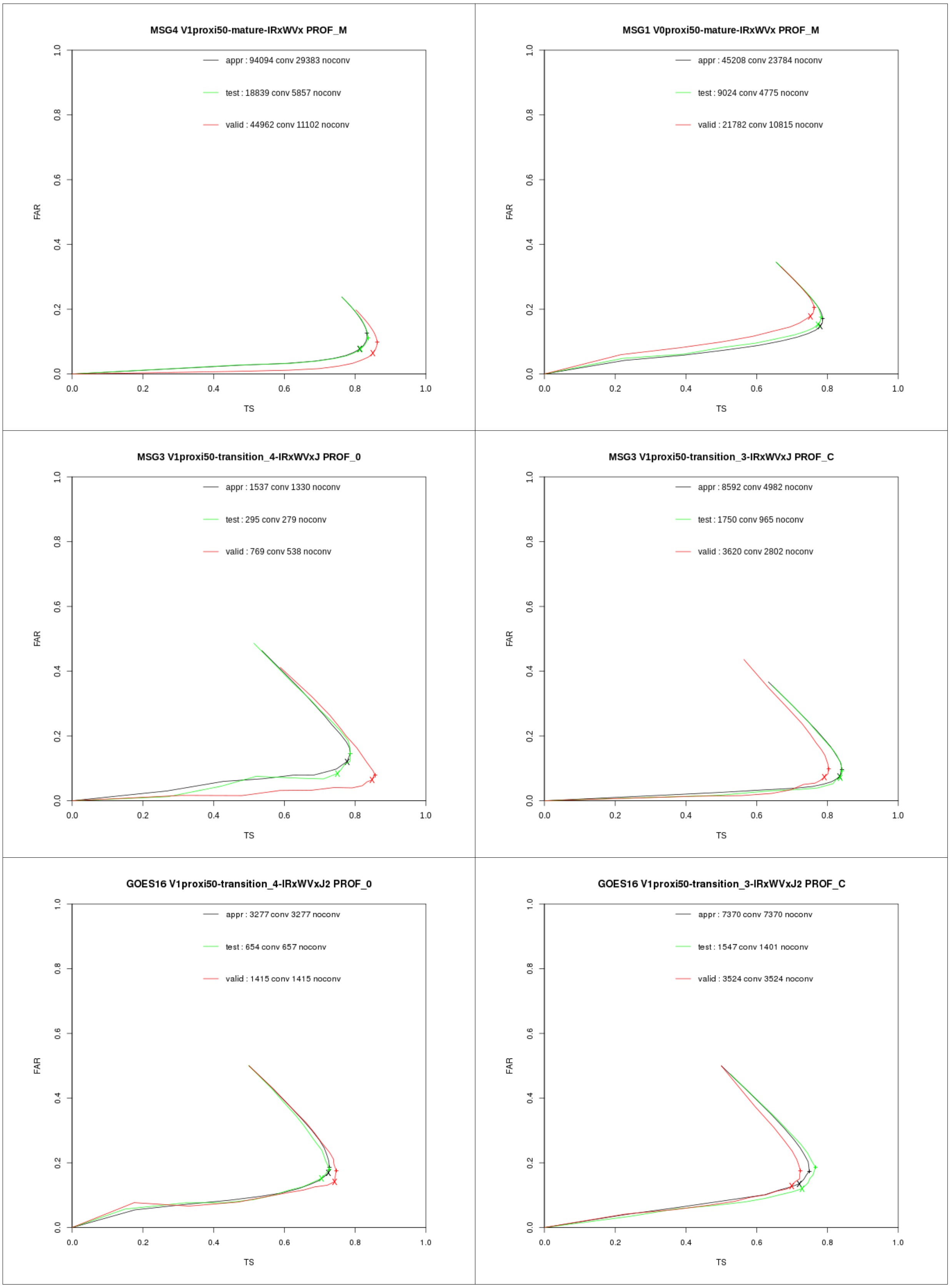


Figure 24: Examples of tuning results for 15 min-MSG4 and MSG1-IO mature category for 30 min-medium depth (top), 5min-MSG3-RSS for static and short depth warmest categories (middle) and 10min-GOES16 for static and short depth warmest categories (bottom). Standard mode for MSG4 and MSG1-OI, day-time modes for MSG3-RS and GOES16. Learning database (black), random test database (green), validation database (red)

The length of dataset for tuning (several months) has allowed to validate many statistical models, even in warmest categories. Day-time tuning brings a light improvement in many cases, especially for warmest categories and shorter depths (even for cells detected for the first time). But the tuning scores emphasize also the dependency on the used ground truth, i.e. lightning data. Since continental thunderstorms are much more electrically active than oceanic ones, tuning is less appropriate on oceanic regions.

Finally, CAL discrimination tuning leads for a matrix of coefficients specific to a given satellite with a given update rate. This matrix lists regressions' coefficients for each category / mode / depth range, corresponding to maximum 6 x 4 x 5 = 120 models, among which not all are active (for example, short depth range never applied to mature category, or "IRxWVxJ2" mode never available with MSG because SEVIRI has no 2.2 channel, or invalidated models).

For the implementation and use in GEO-RDT-CW code, the syntax of this matrix is ConvCoeffRegr.*satellitename.updaterate<.channelnumber>* where *updaterate* is expressed in minutes, and channel 15 (IR10.3) specified when different from default 16 (IR10.8). Thus, satellite name and update rate extracted from satellite configuration file will allow to drive the coefficients' matrix to use. The invalidated statistical models are listed in another auxiliary file, with the syntax ConvCoeffRegr.*satellitename.updaterate<.channelnumber>*_mask.

<u>Concerning update rate,</u> for historical reasons, some satellites have been available at a different scan rate depending on periods. In that case, discrimination files are made available for both values:

> *ConvCoeffRegr.GOES16.**10*** and *ConvCoeffRegr.GOES16.**15*** have been setup on different periods

<u>Concerning main channel,</u>

GOES16/17 and HIMAWARI-8/9 radiometers provide two distinct window channels, IR10.3 and IR11.2. Convection products consider channel #16 as main default IR channel (IR10.8 for MSG series), but now allow using channel #15 instead (argument CANAL_UTIL 15 in configuration file, see UM). In that case, discrimination tuning is undertaken with this configuration:

> *ConvCoeffRegr.GOES16.10.**15** is* now the default discrimination file for GOES16

As mentioned earlier, it has been decided to apply GOES16 statistical models for GOES17 (same generation of satellite). Météo-France presently generates products from GOES17 whith a sub-optimal 30 min update rate. A specific tuning will be attempted with GLM-GOES17 lightning data when nominal 10 min update rate will be taken into account in Météo-France (out of periods with ABI radiometer cooling problems).

> *ConvCoeffRegr.GOES17.**10.15*** is an adaptation of GOES16 models.

The same problematic applies to Himawari-8, with sub-nominal 20min update rate at Météo-France. GOES16 statistical models will also be used, as long as nominal 10min update rate is not taken into account. For the same reason, this approach is applied to Himawari-9:

> *ConvCoeffRegr.HIMA08.**10.15*** is an adaptation of GOES16 models.

> *ConvCoeffRegr.HIMA09.**10.15*** is a duplication of HIMA08 models.

3.2.1.2.3.4 <u>The inheritance and tracking rules</u>

The previous paragraphs depict convective and no convective decision depending on object attributes. This paragraph depicts empirical rules defined on convective management associated to tracking algorithm.

At first, a new detection is always classified like no convective (more precisely "not defined").

Then, from the tracking algorithm, current and previous objects may be temporally linked. A main link (named “father” link) is then identified (often based on the higher surface at a common temperature) to represent the cloud system evolution and process discriminating parameters. But in the discrimination scheme, a convective link is also searched among all temporal links, to allow inheritance of convective diagnosis from one step to the other. In most cases, main father and convective father are the same.

If an object has a convective father, the inheritance makes this object also convective. In the case of decreasing temperature category change (continuous growing), the convective time is initialised to zero for the new class. In the other cases, the convective time and category are incremented.

Then, declassification conditions may be checked.

3.2.1.2.3.5 The declassification rules

The declassification is only applied on convective object in order to diagnose the false alarms or decaying phase. The declassification rules have been empirically defined on case studies due to lack of ground truth to tune statistical models on these problematic.

The convective classification is supposed to be valid at least 60min for cloud systems, from the moment they do not present a strong warming. Due to a higher probability of false alarm, this validity is only 30 min for the warmest category for cooling systems only.

Beyond this validity time, a convective object is declassified if it stays into the same category or goes into warmer category. This schedule allows improving the stability of diagnosis, and focusing on cloud systems in ascending phase.

For the highest (coldest) mature category, declassification rules are slightly different. Moreover, some aeronautical users have expressed the need to delay the declassification of mature convective systems, to maintain attention on those systems at the beginning of the decaying stage. As for warmer layers cases, the convective classification is assumed 60 minutes at least. After this period, the convective object is declassified if one the following conditions is satisfied:

- Cloud system decreases into a warmer category (i.e. minimum temperature becomes warmer than -40°C)
- The BTD WV6.2-IR10.8 becomes lower than –1°C and its trend is negative over the last 60 minutes

If the severity index of the father-cell is positive, the cell is not declassified.

NWP instability indices are used to declassify some long mature cells without sign of activity and evolving in stable areas.

3.2.1.2.3.6 The use of overshooting tops information for convection diagnosis

If the algorithm detects one or more overshooting tops in a given cell, it is likely linked to intense updraft, corresponding to an active convective system (see below Overshooting Top Detection attribute). This information concerning the presence of an overshooting top is used to set convection diagnosis to “yes”. There’s only one possibility of change of convection diagnosis: from “no” to “yes”. No option is offered to skip this possibility.

A cloud cell not classified as convective by the discrimination scheme will be convective if the cell is associated with one or several overshooting tops.

3.2.1.2.3.7 The conditional use of lightning data for convection diagnosis

Lightning data, if available in real time, may greatly contribute to the diagnosis of convection. The object approach allows the data fusion with auxiliary data. The LGH and/or SLGH arguments of configuration file defines the way to use lightning data, respectively from ground sensor and/or satellite sensor. Set to N>0, lightning data is used to eventually change the convective diagnosis of an object. Using this option, there's only one possibility of change of convection diagnosis: from "no" to "yes".

An object not classified as convective by the discrimination scheme can be convective if more than N flashes strokes are paired with the cloud system.

3.2.1.2.3.8 The conditional minimum temperature of tropical systems for convection diagnosis

Particular approach can be applied in tropical regions. When detected and tracked by RDT-CW software, coldest cloud systems may easily be considered as convective beyond a given temperature. The TROPICALDISCRI argument of configuration file defines the minimum temperature threshold for this approach. Using this option, there's only one possibility of change of convection diagnosis: from "no" to "yes".

An object not classified as convective by the discrimination scheme can be convective if located between 30°N and 30°S and minimum temperature attribute below specified threshold, and if corresponding Cloud type corresponds to Very High Opaque or High Opaque.

3.2.1.2.4 The forecast scheme

3.2.1.2.4.1 Main principles

Convective cells have their own dynamics and can have a trajectory that does not always follow the environmental displacement fields. The object mode of RDT-CW analyses the motion of each cell, and compute the speed. The forecast scheme uses this speed estimate to forecast the successive position of each cell.

The Lagrangian method is proved to be quite efficient up to 1 hour range. Thus the RDT forecast is proposed up to this limit.

The advection scheme takes benefit from an improving quality of cell's motion (see previous chapters). It is illustrated in following figure where smoothing and dilatation options have been activated.

Note: Even if those latter options may obviously insert a touch of uncertainty, they also can sometimes lead to excessive dilated contours when large cells are associated with high expansion rates. That's why RDT-CW default configuration file keeps smoothing and dilatation options non active. Those options have to be used preferentially in case study mode.

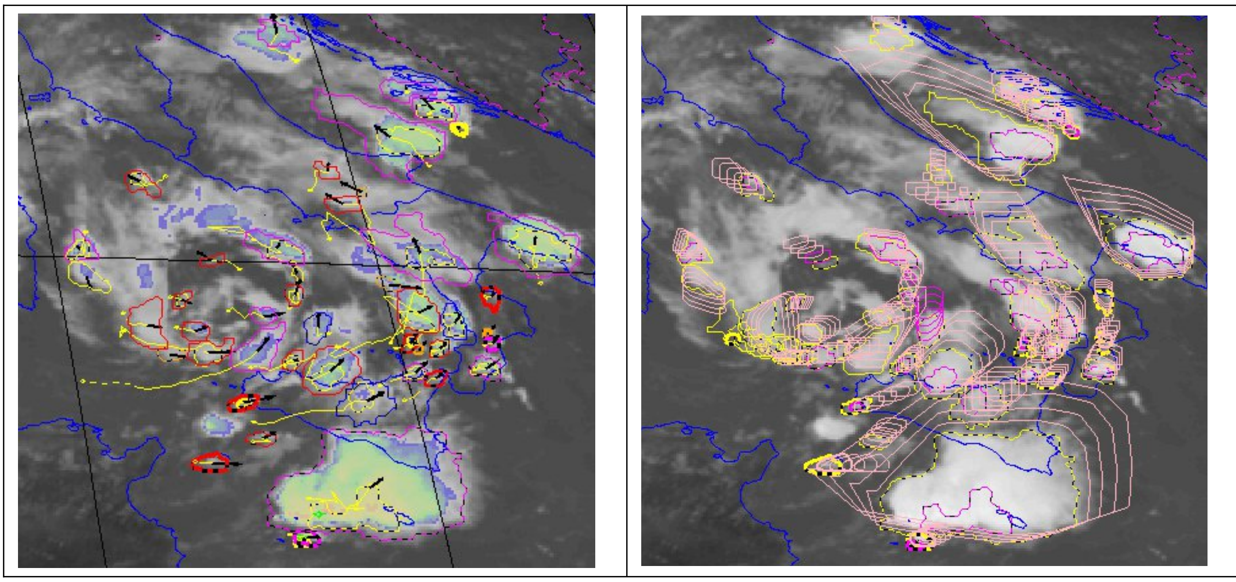


Figure 25: 11th August 2015 09h00 UTC slot – illustration of RDT advection scheme. Left: analysed RDT-CW cells with motion vectors. Right: RDT-CW +15, +30, +45 and +60 smoothed and dilated forecast cells

3.2.1.2.4.2 Specific limitations

Please note that

- If there is not speed estimate for a given cell at a given slot, there should be no forecast of this cell (case of a new cell). Thanks to initialization through guess of movement field, this case should be rare.
- EXIM product is not used as we put priority on cell by cell motion, using cell's motion vector estimate
- Overlap of forecast cells is not managed: motion vectors for neighbouring cloud cells may vary in direction and/or intensity, in particular when split/merge processes occur. In those cases, Lagrangian advection may lead to overlapping of forecast cells. Ideally, one should merge forecast cells, or produce probabilistic forecasts.
- Most of attributes are unchanged for forecast cells. Nevertheless, during first lead ranges up to 30min, trends are taken into account for some attributes. Activity/severity of the cell may be incremented when high cooling or expansion rates, or high lightning trends.
- Default configuration leads to no change in cell geometry. Smoothing advected contours and dilatation/contraction considering expansion rate and lead range may be activated through SMOOTHPTS and DILAT arguments.

3.2.1.2.5 Processing additional and synthesis attributes

A lot of cloud cell's attributes are directly determined during the detection step: extreme and average temperatures, gradients, BTDs, morphology parameters (areas, ellipses axis).

In the version v2018 the set of attributes of the second level has been enlarged.

Other attributes are provided using external sources, like lightning data, other products, NWP data, or are determined at the end of the process just before output encoding, when the maximum of information is available.

3.2.1.2.5.1 Lightning

Lightning data of an identified lightning network is associated with cloud cells when it is possible. User has to provide data compliant to ICD/1 documentation [AD.6] and to update model configuration file. NWCLIB manages to take into account the temporal tolerance for this pairing, as provided through "LGHDTANT" and "LGHDTPOST" arguments of the model configuration file, and considers the region over which RDT-CW is processed.

Then, each pixel of a window centred over a lightning is explored in order to associate the impact with the collocated cell or the closest cell. In this latter case the minimum distance is evaluated against the spatial tolerance as provided in the model configuration file ("LGHTLR" argument). Pairing is confirmed if this distance is lower than required tolerance. If not, a pairing is checked with including level of this closest cell, providing a more simple but additional information. Including level is defined using the outlines 1°C (default configuration) warmer. In general case, this outline includes several towers (see paragraph 3.2.1.2.1.1).

Since v2018 version, the pairing of lightning data is not only undertaken with cells at main "Base of Tower" level, but also at second level ("Top of Tower") when cloud cell morphology necessitates to go toward a more 3D-like cell's definition.

To summarize, pairing with lightning data may concern:

- Base level extension of current cell (exact pairing)
- Proximity to Base contour of current cell (pairing with tolerance)
- Top level extension of current cell (internal pairing)
- Including level of current cell (total lightnings pairing)

Lightning data issued from satellite sensors of different nature than ground sensors (optical sensors like MTG-LI or GOES-GLM for example), can be managed additionally in different containers.

When provided by ground sensors, lightning data are ***counted*** for each cloud cell (when available, negative, positive and intra-cloud lightning data are separately counted). Moreover, the total lightning activity paired with BT level is checked from previous cloud cell, to evaluate a ***lightning trend*** (number/sec) useful for activity or severity assessment.

When provided by a satellite sensor, flash data are ***counted*** for each cloud cell. The total lightning activity paired with BT level is checked from previous cloud cell, to evaluate a ***satellite lightning trend*** (number/sec) useful for activity or severity assessment.

3.2.1.2.5.2 Lightning Jump detection

Since v2018 version, a lightning trend at higher temporal resolution to approach the concept of “lightning jump” is also taken into account. It is an interesting feature derived from total lightning activity analysis (see [R.D. 5] from which the approach has been adapted). The trend is not only evaluated between each satellite slot as explained in previous paragraph but also at minute scale, if the lightning input data allows the calculation.

- Raw lighting data paired with cloud cell at "Base of Tower" level is stored in a time series array at minute scale. To analyse and diagnose a jump at lightning scale, previous studies did rely on a minimum of 12 minutes of lightning history. Thus, lightning time series of successive ancestors of a current cloud cell may be joined to complete the period and allow analysis
- Those data are then processed for a 2min-smoothed time series of flash rates (s^{-1}), and a corresponding time series of flash rates trends.
- For each significant flash rate (over a default configurable value), previous averages and standard deviations of flash rates trends are processed to determine the threshold (2 x rms) to compare with current flash rate trend. When flash rate trend is over this threshold, Lighting Jump is diagnosed as active (LJ_ON). If flash rate trend becomes below 0, previous diagnosis is set to non active (LJ_OBS) or cancelled (NO_LJ)
- Active Lightning Jump may be inherited from previous cell

Lightning Jump (LJ) attribute of Cloud cell may have following values: 0 (NO_LJ), 1 (inherited from previous cell), 2 (LJ_ON diagnosed and still active) or 3 (LJ_OBS has been observed during period).

Lightning Jump is known as a potential very short-term precursor of strong events like Hail or strong gusts. Running RDT-CW with lightning data may give the possibility to benefit from a lightning jump diagnosis, when the lightning data network provides total lightning activity at high temporal resolution.

Lightning Jump detection is processed from ground lightning data and from satellite lightning data when available. The information is made available separately in the end-product.

Lightning Jump algorithm parameters have now been set configurable through model configuration file, and default threshold for Flash Rate (FR) has been set to 20 s^{-1} in order to limit too much diagnosis.

3.2.1.2.5.3 Cloud type

Cloud type value extracted from CT product is examined for each pixel of cloud object. The ***most frequent value*** is attributed to the cloud cell when enough values are available.

The quality of CT products is directly taken into account to assess the quality of this attribute.

3.2.1.2.5.4 Cloud microphysics, Ice Content and risk of Ice Crystals Icing

Cloud micro-physic parameters and values are extracted from CMIC product, and are examined with two objectives:

- Make a 2D analysis of micro-physical values at pixel level in order to identify pixels with probability of high ice water content, and thus likely to contribute to icing at high altitude
- Give informative value of micro-physic parameters as attribute to each cloud object among which icing hazard linked to high altitude ice crystals

High altitude ice crystals are crystals of the size of a grain of flour that may damage the engine of aircrafts and interfere with measurements. A 2D risk index for high altitude icing has been developed thanks to the outcome of the large-scale integrated FP7 HAIC project (High Altitude Ice Crystals). This synthetic index is linked to significant values of high ice water content.

When “phase” value corresponds to ice and Cloud top height is above 5000 m, values of Cloud Optical Thickness (day only) and total Cloud water Path are taken into account to increment an intensity index when above corresponding thresholds (COT above 40, CWP above 0.2 kg/m2). At this stage, only 0 (no severity) and 2 (moderate severity) are encoded, algorithm adapted from KNMI ([R.D. 6]), see following figure for illustration.

The Cloud cell's attributes are following

- The ***most frequent value*** for the “phase” parameter is considered. The value is set to “ice” if the proportion of ice is above 60%. The value is set to “water” if the proportion of liquid water is above 60%. The value is set to “mixed” in other cases.
- The ***maximum*** values of Cloud Optical Thickness (day only), Radius Effective, Ice Water Path and Liquid Water Path are determined over the horizontal cloud cell extension.
- The ***maximum*** value of 2D icing index field is determined over cloud cell extension, and qualifies the risk of high altitude icing (High Icing Hazard) for the whole cloud object. For this attribute, only 0 (no severity) and 2 (moderate severity) values are available. This attribute may be inherited in case of strong continuous cooling or overshooting top detection.

Among the objectives of the micro-physics attributes are

- To accumulate some cases and IWC retrievals to make easier future studies and comparisons.
- To strengthen RDT as a IWC detection tool. But users have to keep attention on the use of this information, as geostationary retrieval of high altitude IWC hazard is still very challenging.

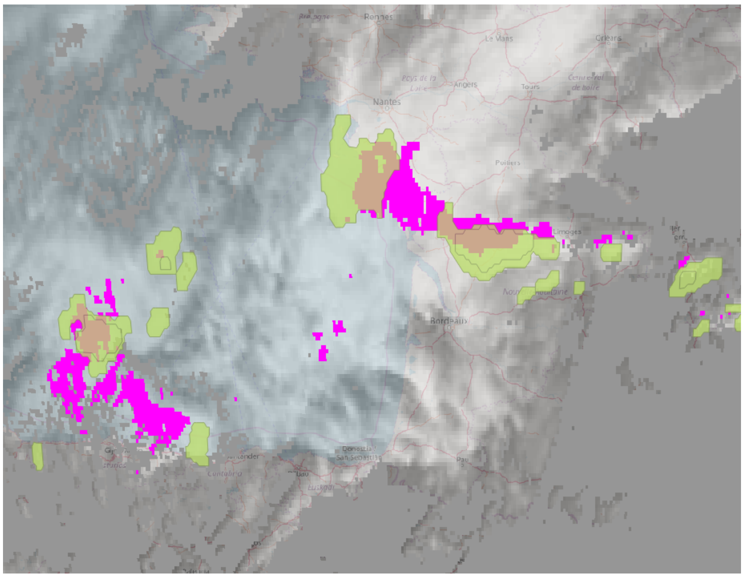


Figure 26: 13th June 2017, 15h00Z, French Biscay Bay. Illustration of 2D field of high altitude icing index derived from CMIC parameters (magenta), superimposed with RDT cells (green). Cells including highlighted pixels are associated with high altitude icing hazard attribute set to 2

3.2.1.2.5.5 Cloud Top Pressure

RDT-CW uses mainly the Cloud Top Pressure value from CTTH product. This approach can fulfil some specific user's needs.

The ***Q10*** value of pressure is considered among the pixels of the cloud-cell, to surround the vertical extension of the cloud. The quality of CTTH product is directly taken into account to assess the quality of the attribute.

3.2.1.2.5.6 Convective Rain Rate

Convective Rain Rate intensity values are issued from CRRPh products. Values are analysed for the cloud-cell horizontal extension and the maximum is associated as attribute to the RDT-CW object.

3.2.1.2.5.7 NWP model inputs

Lifted Index and Tropopause temperature and pressure are the NWP attributes of cloud cells. Those parameters can be used in the discrimination scheme

Tropopause parameters correspond to the ***median value*** of the pixels of the cloud-cell, whereas representative convection index is seen as the minimum of Lifted Index values of pixel of the cloud-cell.

3.2.1.2.5.8 Overshooting Top Detection

Overshooting top (OT) detection is facilitated by the availability of a multi-parametrical description of a cloud cell. Hereafter the description of this part of this algorithm. Relevant parameters, thresholds and some parts of the algorithm are inspired from « Best Practice Document » [R.D. 2].

A first pre-selection of "candidates" overshooting tops is undertaken during the detection step. Once all input data has been managed and analysed, a final confirmation of relevant overshooting tops is done for each cloud cell.

The first step consists in a detailed morphological analysis of the cloud top, combining static and morphological criteria and, if available and corresponding argument set in configuration file, data from visible HR channel.

The minimum temperature of the cloud cell, defining the first pixel of interest, has to be colder than a given value to be considered as OT (called "COLD_Threshold" and set to –50°C in mid-latitude regions and –70°C in tropical regions).The maximum BTD=WV6.3-IR10.8 of the cloud cell has to be above a given value (called "BTD condition" and set to 0°C).

Then, the vicinity of the pixel of interest is analysed inside the cloud cell in order to

1) confirm that the colder spot is above surrounding warmer pixels, identified as following:

- Exploration of surrounding pixels towards height directions and up to two characteristic distances. Only pixels belonging to the cell are then considered. Those distances are multiple values (twice and four times) of typical OT-size of 20km (50km in tropical latitudes) and depend on the size of the pixel.
- Warmer surrounding pixels are identified if the temperature difference with OT is above a given value (called "minimum vertical extension threshold" and set to 6°C)

2) Define the horizontal extension of the OT. Neighbour pixels belonging to OT are identified as following:

- Exploration of all cell's pixels inside a twice "OT-size"-window centred on the pixel of interest
- Temperature difference with OT lower than a given value (called "maximum threshold" and set to 3°C)

If the OT morphology is not confirmed in IR channel, and if HR resolution is available for VIS06 channel, maximum reflectance and gradient of reflectance are analysed for the cloud cell, and may confirm OT detection.

Other pixels of interest may be taken into account giving relevant BTD (WV6.2-IR108) and VIS06 (NR, or HR if available) maximum values over cloud cell extension, if the corresponding pixels are not close to a previously identified OT candidate ("distance threshold" set to 50km in mid-latitude regions and 200km in tropical regions).

This step identifies a cloud cell's list of so-called "OT candidates".

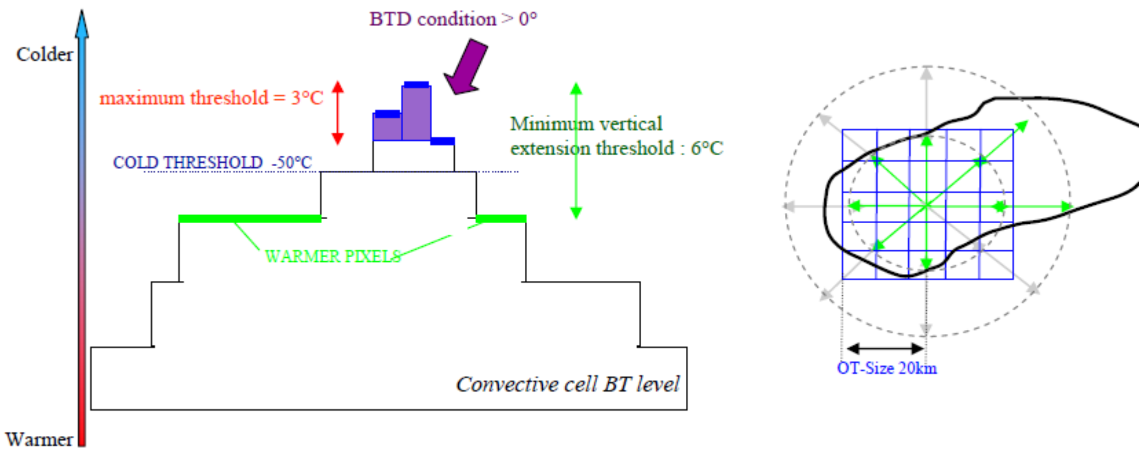


Figure 27: First step of OT detection: Values of criteria for vertical morphological analysis (left), illustration of a horizontal analysis (right). Arrows point surrounding pixels to check (green when belonging to the cloud cell), blue grid represent the research window for OT extension.

For each of those pre-identified OT, various parameters are calculated: IR minimum and gradient, various BTD values, reflectance maximum and gradient (NR and/or HR).

The second step takes benefit from tropopause information read or processed from NWP data. Tropopause data are a direct input of product or can be re-processed from other NWP parameters. The conditions to confirm a candidate as overshooting top are described below. They refer to default thresholds listed in an header file:

- If NWP tropopause available, OT colder than tropopause. This represents the main characteristic of the relevancy of OT extension. Greater impact if OT at least 5°C colder than tropopause
- BTD between WV6.2 and IR10.8 taken into account. This criteria highlights a relevant intensity of OT
- During daytime, HR or NR VIS06 high reflectance. With HR visible mode, strong reflectance gradient is also taken into account. This criteria offers an additional alternative signature of cloud morphology
- OT temperature largely colder than average temperature of cloud cell. This represents an alternative criteria for budding above surrounding
- Following additional BTDs can also be taken into account, depending on availability and configuration:
    - BTD4 between IR87 and IR108
    - BTD6 between IR134 and IR108
    - BTD7 between IR97 and IR108

It is to note that without NWP tropopause data or visible channel information, OTD of RDT-CW will be less reliable. The more channels are taken into account, the more reliable OTD is.

3.2.1.2.5.9 The synthesis attribute ”phase of development”

RDT-CW output product includes a specific attribute named “phase of life” or “phase of development”, ranging from “triggering” (or “split”) to “growing”, “mature” then “decaying” values. This attribute is an attempt to diagnose the stage of development of the tracked cloud system.

The true development stage of a convective system is difficult to diagnose: conceptual models do not always spread over the same period of time, and depend on the way the convection is organized.

This attribute is a short mix of parameters linked to

- History, through its duration and detection step (birth or triggering from a split)
- Vertical extension, through "category" of the cloud cell,
- Activity of the system, through expansion or cooling rates, lightning activity, presence of overshooting tops.

All those parameters allow to get a relevant and realistic diagnosis.

Apart from “triggering” and “split” cases, driven by the history of the system, “growing” case relies mainly on the activity of the cloud system, where “mature” corresponds to active Mature category. The table hereafter synthesize the decision tree.

Table 10: Diagnostic of phase of development

History	***Cooling***	***Vertical category (based on Tmin)***	***Expansion***	***Activity: Type, lightning, OTD***	***Encoded phase***
birth					***0 = Triggering***
Split					***4 = Triggering from a Split***
continued	*Cooling*	***Warm***			***1 = growing***
		Warm Transition			***1 = growing***
		Warm transition2			***1 = growing***
		Cold Transition			***1 = growing***
		Mature	*no expansion*		***2 = mature***
			expansion		***1 = growing***
	~same T°	***Warm***	*expansion*	*or **Conv** or Light.*	***1 = growing***
			no expansion +	*not Conv + no Light.*	*3 = decaying*
		Warm Transition	*expansion*	*or **Conv** or Light.*	***1 = growing***
			no expansion +	*not Conv + no Light.*	*3 = decaying*
		Warm transition2	*expansion*	*or **Conv** or Light.*	***1 = growing***
			no expansion +	*not Conv + no Light.*	*3 = decaying*
		Cold Transition	*expansion*	*or **Conv** or Light.*	***1 = growing***
			no expansion +	*not Conv + no Light.*	*3 = decaying*
		Mature			***2 = mature***
	Warming	***Warm***		*Lightning or (**Conv** and low warming)*	***1 = growing***
				else	*3 = decaying*
		Warm Transition		*Lightning or (**Conv** and low warming)*	***1 = growing***
				else	*3 = decaying*
		Warm transition2		*Lightning or (**Conv** and low warming)*	***1 = growing***
				else	*3 = decaying*
		Cold Transition		*Lightning or (**Conv** and low warming)*	***1 = growing***
				else	*3 = decaying*
		Mature		*Lightning or (**Conv** and low warming) or OTD*	***2 = mature***
				else	*3 = decaying*

3.2.1.2.5.10 The synthesis attribute "severity"

RDT-CW is a result of cloud cell morphological analysis, tracking, convective diagnosis and data blending with external data sources. For that reason, there is a possibility to undertake a kind of synthesis of parameters which are supposed to be linked to a strong activity or severity of a thunderstorm. Several attribute values are analysed and contribute to severity index:

- Presence of overshooting top
- Strong cooling (thresholds of temperature change -50°C/h and -100°C/h)
- High BTD trend (+50°/h)
- Strong horizontal expansion rate (threshold: +100%/15')
- High lightning activity (thresholds 20, 50 strokes or flashes paired with the cloud cell)
- Occurrence of Lightning Jump
- High values of convective rain rate (50 mm/h)
- Occurrence of high altitude icing hazard

Those elements increment level on a 4-scale index, and give a global estimation of the activity of the convective cloud system, low, moderate, high or very high.

Then characterization of most relevant phenomena is undertaken from available information: turbulence when presence of overshooting top, high ice water content when corresponding index is on, heavy rain when high values of Convective Rain Rate.

3.2.2 Practical Considerations

3.2.2.1 Quality Control and Diagnostics

RDT-CW doesn't process real time quality control on tracking or diagnosis result. Sanity checks are in place to make, for example, the speeds realistic. But it is not used to create a quality control diagnosis.

3.2.2.2 Exception Handling

In case of missing satellite images, some error messages inform the user and RDT-CW fully recovers its quality few images later. Nevertheless, RDT-CW manages the flag quality of some optional input products like CT, CTTH and CRR.

Moreover, the RDT-CW software produces some error messages in exception cases.

3.2.2.3 Output

The final product is numerical data which depict satellite characteristics and move information associated to RDT-CW cells, and provides a full spatial description (3D-like) of each cell. Numerical data is provided as a list of current cells and characteristics valid for the current slot.

The output NetCDF default format is described in the Output Format Document of NWCSAF (see [AD.8]).

The output file lists a large number of variables/attributes, taking into account horizontal, vertical and temporal description of each cloud cells. But this output may also include an image part.

- The bulletin-like part of the product relies on several dimensions: number of cells, number of contour points for horizontal description, number of levels, slices and overshooting tops for vertical description, number of trajectory points for a temporal description.
  - The *overview part* of the output lists some characteristics of cloud cell population. It may also include an additional optional *map of type and phase* of cloud cells (not the default mode). It is possible to activate the encoding of this map (user configuration with NCMAPINCLD argument of configuration file set to 1)
  - The *cell part* details the spatial and temporal description of the cloud system
    - The *main description part* lists for each cloud system identity characteristics, date type and other characteristics (type, movement, cooling rate, severity …) which concern the whole cloud system.
    - The *level and contour description part* lists for each cloud system and each "bottom" and "top" threshold level the localization parameters (contour and gravity centre), satellite characteristics, morphological characteristics and data fusion parameters (lightning, other products, etc.).
    - The *vertical surface description part* lists for each cloud system pairs of threshold brightness temperature / surface allowing vertical morphological description.
    - The *historical description part* lists for each cloud system a limited set of characteristics of its recent past (maximum 12 time steps corresponding to satellite refresh rate): localization of gravity centre, satellite and morphology characteristics, movement and trends. This part makes the RDT-CW output independent of previous outputs, when users want to manage trajectory of the cloud system and main temporal characteristics. A more complete temporal description will imply to manage previous outputs.
  - The *overshooting top description part* lists all detected overshooting tops, their localization, characteristics and reference to the corresponding cloud system

- Forecast products
  - Forecast products are only bulletin-like product, without map container. Moreover, the set of variables/attributes is more restricted: only main and bottom level description.
  - There is one forecast product for each given lead range. The maximum lead range may be configurable with FCST argument, but cannot exceed 1 hour. Forecast products are available at a forecast step defined by (and configurable with) FCSTEP argument, usually set to satellite update rate. For example, default configuration for MSG with FCST=60 and FCSTEP=15 leads to generate 5 output files, one for the analysis, and 4 for +15, +30, +45 and +60min lead ranges.

3.2.2.4 Access to cloud system history

The recent main characteristics of a cloud system are described inside the NetCDF product, with a corresponding dimension describing trajectory steps.

Thus, user may benefit (or not) from the last steps of a current cell, without accessing previous products. Past localization parameters as well as important characteristics (brightness temperatures, surfaces, trends …) are now available in each current product.

3.2.2.5 Assumptions and Limitations

The tuning of discrimination scheme relies on the availability of an efficient lightning data network used as ground truth for constituting learning and validation data sets. It had been carried out on a summer period over a domain centred over France using French Meteorage lightning data network. It may also use data from extended network (like EUCLID data for extended validation).

Nevertheless, the discrimination scores during winter period could remain weak since the lack of convective situations impacts the quality of learning data set for tuning purposes.

3.3 CTRAJ INTERMEDIATE PRODUCT

3.3.1 Goal of the CTRAJ intermediate product

CTRAJ is a technical intermediate product. When RDT-CW represents a full spatial (3D like) description of cloud cells, CTRAJ represents a full temporal description of terminated cloud systems.

CTRAJ product represents the difference between current internal backup file and previous one, storing characteristics of achieved cloud system trajectories. Thus, a trace of past convective activity can be made available.

Depending on the use, thus on user's choice and configuration, CTRAJ file may content full or brief information of cloud systems.

The following main uses of CTRAJ intermediate product can be identified:

- Discrimination Tuning (full description – backup-like)
  Trajectories produced on a specific area with all kinds of data constitute a basis for convective discrimination tuning of RDT. This facility is needed for managing new input data or new releases of RDT.
- Monitoring (synthetic description)
  Real-time trajectory production allow to have the first step of convection validation. For example to provide forecasters or end-users some element about the quality of the product. Real time validation depends on available data. It also helps have an overview of a complete trajectory.

- Climatological studies (statistical description)
  Real-time trajectory production can allow to undertake studies to define hazardous areas with high probability of RDT-CW, to define for a given place the beginning and the end of the convective season, etc.

3.3.2 Optional production outline

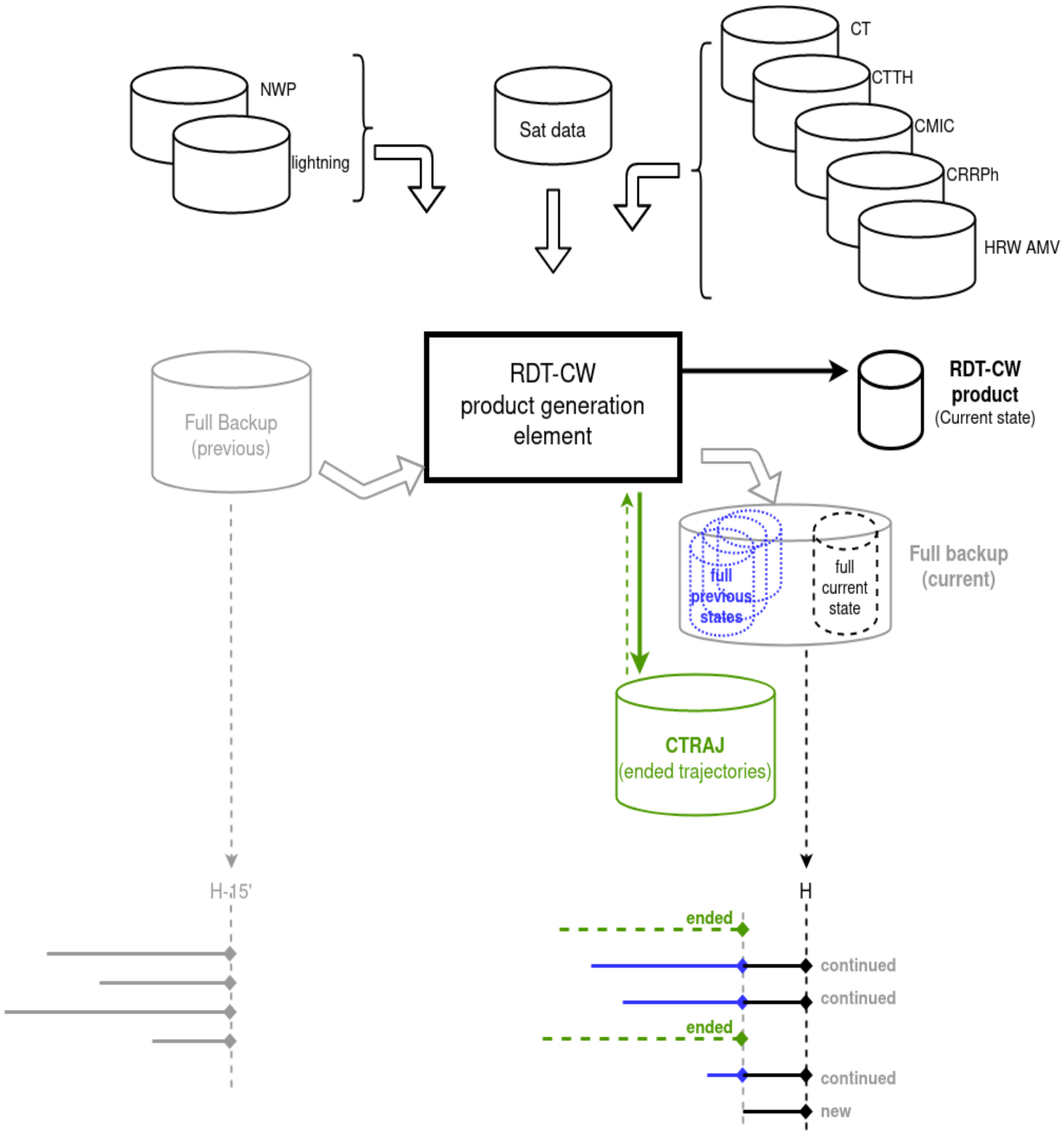


Figure 28: Production of RDT-CW and the associated intermediate product CTRAJ

The production of CTRAJ file is optional.

The advantage of this intermediate product is the possibility to define the characteristics of trajectories that will be kept and eventually to get specific attributes related to the whole trajectory: maximum of activity, area swept by the convective cell, length of trajectory, etc. Options help the user to configure the level of description of the convective cells trajectories.

For an optimal production of CTRAJ intermediate product, an optimal production of RDT-CW is mandatory, i.e. with all possible external or internal input data (NWP, lightning data, CMA, CT, CTTH, CRRPh, CMIC, HRW, etc.).

3.3.3 Practical Considerations

Output product will remain in ASCII format (at least in a first stage), by default located in SAFNWC/tmp directory like other databuf files.

It describes some of the selected characteristics of the previous cloud system trajectories that ended at the current slot, i.e. whose tracking was interrupted or finished on the current image.

Three output modes are available, upon user's request:

- Distinct output at each slot (default mode)
- Daily output updated at each slot (input/output file)
- Monthly output updated at each slot (input/output file)

For the two latter options, previous output file is updated with the latest trajectories. If there is, for example, no cloud activity during a short period, file won't be updated and possibly removed by Task Manager monitoring cleaning tasks.

This is why we offer the possibility to configure the output directory for this intermediate file with an additional argument TRAJDIR in model configuration file. In that case, CTRAJ files will be written (and eventually updated) in SAFNWC/${TRAJDIR} directory, which has to be created first. For example, following argument will redirect CTRAJ files in standard output directory for RDT products:

TRAJDIR export/RDT

3.3.4 CTRAJ content

At the end of the processing, RDT-CW may provide (depending on user's configuration) a consolidated record of all achieved trajectories during the time slot. The trajectory file is coded in ascii format. The name of the file depends on the frequency of production (driven by TRAJPROD argument of model configuration file):

- Monthly (0) S_NWC_CTRAJ_*satelliteId_regionName_YYYYMM*
- Daily (1) S_NWC_CTRAJ_*satelliteId_regionName_YYYYMMDD*
- Each slot (2 - ***default***) S_NWC_CTRAJ_*satelliteId_regionName_YYYYMMDDTHHmm*

This file is organized in rows. The first character differentiates different kinds of lines:

- A line beginning with '***T*** indicates the beginning of the description of a whole trajectory. The description begins from the first detected cell to the last one, sequentially.
- Each cell is described by a group of different lines (***I, S, S2, X, H, O, L, LI***) each documenting a particular aspect of the concerned cell
  - Lines beginning with 'I' = full identification of the cell
  - Lines beginning with 'S' = satellite morphology and radiative characteristics of tracked cell ('S2' = characteristics of corresponding tower top cell)
  - Lines beginning with 'X' = radiative and temporal characteristics of additional satellite channels, information from other PGEs or from NWP data
  - Lines beginning with 'O' = description of overshooting top detected in cell
  - Lines beginning with 'H' = Vertical description of cell
  - Lines beginning with 'L' = association with ground-based lightning activity

- Lines beginning with 'LI' = association with space-based lightning activity

More details about encoded parameters can be found in annex.

ANNEX – CTRAJ INTERMEDIATE PRODUCT CONTENT DESCRIPTION

1. Line "T" (Trajectory)

The T line describes a trajectory. The cells that compose it are outlined in the following lines.

Format:

T ch2 ch3 ch4 … ch14 ch15 ch16

ch2: The type of trajectory:
- 0: Path not convective
- 1: Trajectory convective
- 2: *internal use*

ch3: Flag test cutoff temporal trajectory
- 0: No break
- 1: The system has been detected just after a missing image
- 2:Trajectory with missing time step
- 3: The cutoff spatial index is not zero for at least one cell of the trajectory
- 4: 1 & 2
- 5: 1 & 3
- 6: 2 & 3
- 7: 1, 2 & 3
- 14: End of episode (edge of domain)

ch4: Flag test cut-off spatial trajectory
- 0: No break
- 1: Cell touching the edge of the domain
- 2: Pixel (s) of incalculable value to the cell
- 3: 1 & 2

ch5: The duration in minutes
ch6: The number of the first cell of the trajectory
ch7: The start date (Format YYYYMMDDhhmm)
ch8: The end date (Format YYYYMMDDhhmm)
ch9: Characteristic of trajectory start
- Normal: 'n'
- After a split 's'
- After a merge: 'g'
- After a complex case (undefined: merge + split) 'c'
- Error code: '.'

ch10: Number of cells integrated over the trajectory
ch11: Number of cells expelled during the course (cell split)
ch12: Characteristic of closure
- Normal: 'n'
- Termination in split case: 's'
- Termination in merge case: 'g'
- Termination in complex case: 'c'
- Error code: '.'

ch13: *internal use*
ch14: *internal use*
ch15: *internal use*
ch16: The probability of error of the method of discrimination (in%)

2. Line "I" (Identity)

The "I" line identifies a cell.

Format:

I ch2 ch3 ch4 ch5 ch6 ch7 ….. ch20 ch21 ch22

ch2: The ***date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: The cell ***identifier*** (Format YYYYMMDDhhmm_ <temperature of threshold> _ <latitude gravity center> _ <longitude gravity center >)
ch5: convective ***type*** (0 to 15)

- 0: diagnosis not convective (static model results)
- 1: diagnosis convective (static model results)
- 2: *internal use*
- 3: cell previously convective and declassified
- 4 & 5: convective nature of inherited primary link (5 if under development stage)
- 6 & 7: convective nature inherited a secondary link (7 if under development stage)
- 8 & 9 & 10: undefined
- 11 & 12 & 13 & 14: forced diagnosis (LGH, OTD, CRR, TROP)
- 15: inherited forced diagnosis

ch6: the likelihood of ***error*** diagnostics above (%) (-9999 if no diagnostic)
ch7: ***class/category*** of cell

- 0: mature Tmin <= -40 °C
- 1: mature transition Tmin <= -40 °C
- 2: Cold transition Tmin <= -35 °C
- 3: Warm transition Tmin <= -25 °C
- 4: Warm transition Tmin <= -15 °C
- 5: Hot cell (no active discrimination): Threshold temperature > -5 °C

ch8: the cell ***class/category of convective*** diagnosis (-9999 otherwise)
ch9: ***age*** (minutes) in the class
ch10: ***convective age*** (minutes) convective (-99999 otherwise)
ch11: ***history*** (minutes) available with 5 channels
ch12: ***history*** (minutes) available with 3 channels IR10.8, WV6.2, IR12.0
ch13: ***history*** (minutes) available with 3 channels IR10.8, WV6.2, WV7.3
ch14: ***history*** (minutes) available with 2 channels IR10.8, WV6.2
ch15: ***history*** (minutes) available with channel IR10.8 only
ch16: ***history*** (minutes) with the maximum available channel IR10.8 without time interruption
ch17: ***history*** (minutes) with available NWP data
ch18: ***significant*** characteristic of the cell (0 or 1)
ch19: ***severity*** index of the cell (0=NIL, 1=light,2=moderate,3=high,4=very high, 5=unknown)
ch20: type of hazard of cell (0=NIL, 1=turbulence, 2=lightning, *3=icing, 4=high ice water content, 5=hail,* 6=heavy rain rate, 7=unknown)
ch21: cell ***number at birth***
ch22: ***day/night flag*** for cell (day=0, twilight=1, night=2)
ch23: **land/sea flag** for cell (space or missing=0, land=1, sea=2, coastal=3)
ch24: **azimutal solar** angle at cell gravity center
ch25: **zenithal solar** angle at cell gravity center
ch26: **phase of life** associated to cell

3. Line "S" or "S2" (Satellite)

The S line gives the satellite characteristics of the cell previously defined in line "I".

Format:

S ch2 ch3 ch4 ch5 ch6 ch7 … ch33 ch34

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: ***Threshold*** temperature at which the cell is defined
ch5: ***internal number*** of the cell into the threshold process
ch6: Position of gravity centre: ***Latitude***
ch7: Position of gravity centre: ***Longitude***
ch8: Position of weighted gravity centre: ***Latitude***
ch9: Position of weighted gravity centre: ***Longitude***
ch10: ***Average*** temperature of the cell
ch11: ***Standard deviation*** of cell temperature
ch12: ***Minimum*** temperature of the cell
ch13: ***Area*** (km2)
ch14: ***Minor axis*** of the ellipse approaching the cell (in km)
ch15: ***Major axis*** of the ellipse approaching the cell (in km)
ch16: ***Orientation*** of the ellipse approaching the cell (angle in degrees from north)
ch17: Internal use
ch18: Internal use
ch19: Internal use
ch20: Average temperature ***gradient*** on the cell
ch21: Average of ***peripheral gradient*** of temperature on the cell
ch22: ***Percentile 95%*** of peripheral gradient
ch23: ***Number of pixels*** on the periphery of the cell
ch24: X axis (horizontal) ***Speed*** of movement in the (pixels / hour)
ch25: Y axis (vertical) ***Speed*** of movement in the (pixels / hour)
ch26: ***cooling rate*** of the cell based on the minimum temperature (° C / h and positive when Tmin decreases)
ch27: ***minimum pressure*** of top of the cloud (CTTH)
ch28: height of cloud base (missing)
ch29: Pixel location of the minimum temperature: ***Latitude***
ch30: Pixel location of the minimum temperature: ***Longitude***
ch31: ***Closest lightning*** = pixel distance of the nearest impact
ch32: moving ***speed*** (m/s) of the cell
ch33: moving towards ***direction*** of the cell
ch34: Cloud Top Pressure ***trend*** (Pa/sec)

4. Line "X" (eXtra data)

The X line describes the characteristics of additional channels, additional PGEs or NWP data of the cell previously defined in line "I".

Format:

X ch2 ch3 ch4 ch5 ch6 ch7 …….ch58

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: Attribute external Classification Cloud = predominant ***cloud type*** (CT)
ch5: Cloud Field Classification = surface corresponding ***proportion***
ch6: Cloud Field Classification = cumuliform or stratiform ***type*** (CT)
ch7: Attribute = channel WV6.2 ***Minimum*** temperature
ch8: Attribute = channel WV6.2 ***trend***
ch9: Attribute = channel WV7.3 ***Minimum*** temperature
ch10: Attribute = channel WV7.3 ***trend***
ch11: Attribute = WV-IR Brightness Temperature Difference ***maximum*** value on cell
ch12: Attribute = WV-IR Brightness Temperature Difference ***75% percentile***
ch13: Attribute = WV-IR Brightness Temperature Difference ***90% percentile***
ch14: Attribute = WV-IR Brightness Temperature Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch15: Attribute = WV62-WV73 Brightness Temperature Difference ***maximum*** value
ch16: Attribute = WV62-WV73 Brightness Temperature Difference ***75% percentile***
ch17: Attribute = WV62-WV73 Brightness Temperature Difference ***90% percentile***
ch18: Attribute = WV62-WV73 Brightness Temperature Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch19: Attribute = channel VIS06 ***Maximum*** reflectance
ch20: Attribute = channel VIS06 ***trend***
ch21: Attribute = channel IR1.6 ***Maximum*** reflectance
ch22: Attribute = channel IR1.6 ***trend***
ch23: Attribute = IR1.6-VIS0.6 Reflectance Difference ***maximum*** value
ch24: Attribute = IR1.6-VIS0.6 Reflectance Difference ***percentile 75%***
ch25: Attribute = IR1.6-VIS0.6 Reflectance Difference ***percentile 90%***
ch26: Attribute = IR1.6-VIS0.6 Reflectance Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch27: Attribute = channel IR3.9 ***Minimum*** temperature
ch28: Attribute = channel IR3.9 ***trend***
ch29: Attribute = channel IR8.7 ***Minimum*** temperature
ch30: Attribute = channel IR8.7 ***trend***
ch31: Attribute = channel IR12.0 ***Minimum*** temperature
ch32: Attribute = channel IR12.***trend***
ch33: Attribute = IR3.9-IR10.8 IR Brightness Temperature Difference ***maximum*** value
ch34: Attribute = IR3.9-IR10.8 Brightness Temperature Difference ***percentile 75%***
ch35: Attribute = IR3.9-IR10.8 Brightness Temperature Difference ***percentile 90%***
ch36: Attribute = IR3.9-IR10.8Brightness Temperature Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch37: Attribute = IR8.7-IR10.8 Brightness Temperature Difference ***maximum*** value
ch38: Attribute = IR8.7-IR10.8 Brightness Temperature Difference ***percentile 75%***
ch39: Attribute = IR8.7-IR10.8 Brightness Temperature Difference ***percentile 90%***
ch40: Attribute = IR8.7-IR10.8Brightness Temperature Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch41: Attribute = IR12.0-IR10.8Brightness Temperature Difference ***maximum*** value
ch42: Attribute = IR12.0-IR10.8 Brightness Temperature Difference ***percentile 75%***
ch43: Attribute = IR12.0-IR10.8 Brightness Temperature Difference ***percentile 90%***
ch44: Attribute = IR12.0-IR10.8 Brightness Temperature Difference ***ratio***
(nb relevant pixels surrounded by relevant pixels SIG) / (nb relevant pixels)
ch45: Attribute NWP= predominant value over cell of ***convective mask***
0=NOCONV, 1=NEUTRAL, 2=CONV

ch46: Attribute NWP= convective index ***type***: 0=K index, 1=Lifted index, 2=Showalter index
ch47: Attribute NWP= convective index ***value*** of (10% or 90%percentile depending on index)
ch48: Attribute NWP= ***Tropopause Temperature*** median value over cell
ch49: Attribute NWP= ***Tropopause Pressure*** median value over cell
ch50: Attribute = Cloud ***Phase*** - water (1) or ice (2) or mixed (3) (CMIC)
ch51: Attribute = maximum ***convective rain rate*** (mm/h) (CRR)
ch52: Attribute = Ice Crystal Risk
ch53: Attribute = maximum ***radius effective*** (m) (CMIC)
ch54: Attribute = maximum ***cloud optical thickness*** (unitless) (CMIC)
ch55: Attribute = maximum ***liquid water path*** (kg/m2) (CMIC)
ch56: Attribute = maximum ***ice water path*** (kg/m2) (CMIC)
ch57: Attribute = channel *IR2.2µm maximum reflectance*
ch58: Attribute = channel *IR2.2µm reflectance* trend
ch59: Attribute = channel *HR VIS maximum reflectance*
ch60: Attribute = channel *HR VIS maximum reflectance gradient*
ch61: Attribute = channel BTD7 (IR9.7µm-IR10.8µm) maximum value
ch62: Attribute = channel *BTD6 (IR13.4µm-IR10.8µm) maximum value*

5. Line "O" (Overshooting top)

The "O" line documents overshooting top characteristics associated with the cell defined in line "I". There are as many "O" lines as overshooting tops in the cell. In practice cells with more than one overshooting top are not frequent.

Each overshooting top is described by at least one point, corresponding to the main characteristic (minimum temperature and/or maximum BTD, etc …).

O ch2 ch3 ... ch 23 ch24 ....

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: its ***number*** (internal functioning of PGE)
ch4: main ***characteristics*** of overshooting top: "IRMin", "BTDMax", "ReflMax"
ch5: ***temperature*** (°C) of overshooting top (IR10.8)
ch6: ***horizontal gradient*** (°C/km)
ch7: maximum ***temperature difference*** (°C) between OT and surrounding pixels (always >= 0)
ch8: number of warmer ***surrounding pixels***
ch9: ***BTD***=WV6.2-IR10.8 (always >= 0)
ch10: number of ***pixels*** with BTD >=0
ch11: ***WBTD***=WV6.2-WV7.3 (when available)
ch12: ***BTD3***= IR3.9-IR10.8 (when available)
ch13: ***BTD4***=IR10.8-IR8.7 (when available)
ch14: ***BTD5***=IR12.0-IR10.8 (when available)
ch15: ***BTD7***=IR9.7-IR10.8 (when available)
ch16: ***BTD6***=IR13.4-IR10.8 (when available)
ch17: ***reflectance*** (VIS) (when available)
ch18: horizontal ***gradient*** of reflectance (VIS) (when available)
ch19: ***Reflectance Difference*** IR1.6-VIS0.6 (when available)
ch20: ***reflectance*** (HR VIS) (when available)
ch21: horizontal ***gradient*** of reflectance (HR VIS) (when available)
ch22: ***pressure gap*** to tropopause (hPa)
ch23: ***temperature gap*** to tropopause (°C)
ch24: ***number of points*** belonging to overshooting top (always >=1)

ch25: ***latitude*** of main point of overshooting top
ch26: ***longitude*** of main point of overshooting top
ch27: (latitude of next point)
ch28: (longitude of next point)
.../...

6. Line "H" (morpHology)

The "H" line documents the vertical structure associated to the cell defined in line "I"

Format:

H ch2 ch3 ch4 ch5 ch6 ch7 …

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: The threshold temperature of the ***hottest temperature*** for which the surface of the system is processed
ch5: The temperature ***increment*** between two consecutive surfaces
ch6: The ***number*** N of surfaces evaluated
ch7 to ch X (X = 6 + N): The ordered ***surfaces*** as follows: S (ch4), S (ch4 + ch5), S (ch4 +2*ch5), ..., S (ch4 + (N-1)*ch5)

7. Line "L" (Lightning)

The L line documents the ground-based electrical activity of the cell (identified with the line 'I' preceding the line)

Format:

L ch2 ch3 ch4 ... ch12 ch13 ch14 ch15 ...

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: Number of negative ***intra-cloud*** lightning in the cell in the]date-date_before; date+date_after]
ch5: Number of positive ***intra-cloud*** lightning in the cell in the]date-date_before; date+date_after]
ch6: Number of ***negative impacts*** lightning in the cell in the]date-date_before; date+date_after]
ch7: Number of ***positive impacts*** lightning in the cell the]date-date_before; date+date_after]
ch8: ***Average intensity*** of ***negative*** impacts in the]date-date_before; date+date_after]
ch9: ***Standard deviation*** of intensity of ***negative*** impacts in the]date-date_before; date+date_after]
ch10: ***Average intensity*** of ***positive*** impacts in the]date-date_before; date+date_after]
ch11: ***Standard deviation*** of intensity of ***positive*** impacts in the]date-date_before; date+date_after]
ch12: ***Average intensity*** of ***negative*** **intra-cloud** lightning in the]date-date_before; date+date_after]
ch13: ***Standard deviation*** of intensity of ***negative*** **intra-cloud** lightning in the]date-date_before; date+date_after]
ch14: ***Average intensity*** of ***positive intra-cloud lightning*** in the]date-date_before; date+date_after]
ch15: ***Standard deviation*** of intensity of ***positive intra-cloud lightning*** in the]date-date_before; date+date_after]

ch16: Number of total lightning strokes paired with including cell (when no strokes paired with current cell)
ch17: total ***lightning trend*** expressed in number / second
ch18: ***maximum*** value of 1minute ***flash rate*** expressed in number / second
ch19: ***Lightning Jump*** detection
ch20: ***length*** of minute-lightning array (0 if no data)
ch21: Date of ***first impact*** of lightning associated with that cell (YYYYMMDDHHMMSS)
ch22 .... : list of ch15 elements corresponding to 1 minute - time serie of lightning paired with the cell

8. Line "LI" (space-based Lightning Imager)

The LI line documents the electrical activity issued from satellite sensor for the cell identified with the line 'I' preceding the line

Format:

LI ch2 ch3 ch4 ... ch12 ch13 ch14 ch15 ...

ch2: ***Date*** of the cell (Format YYYYMMDDhhmm)
ch3: His ***number*** (internal functioning of PGE)
ch4: Number of lightning ***flashes*** in the cell in the]date-date_before; date+date_after] period
ch5: Number of lightning ***flashes*** paired with ***including cell*** (when no flash paired with current cell)
ch6: ***lightning flash trend*** expressed in number / second
ch7: ***maximum*** value of 1minute ***flash rate*** expressed in number / second
ch8: Satellite ***lightning Jump*** detection
ch9: ***length*** of minute-lightning array (0 if no data)
ch10.... : list of ch9 elements corresponding to 1 minute - time series of lightning paired with the cell