Validation report of the Convection Product Processors of the NWC/GEO

NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection, Issue 2, Rev. 0.1

28th February 2021

Applicable to
  GEO-CI v.2.2 (NWC-089)
  GEO-RDT-CW v.5.2 (NWC-090)

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

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	Météo-France MFT (F. Autonès and M. Claudon)	Signed, F. Autonès Signed, M. Claudon	*28th February 2021*
Reviewed by	Météo-France MFT (J.-M. Moisselin)	Signed, J.-M. Moisselin	*28th February 2021*
Authorised by	P. Ripodas NWC SAF Project Manager		*28th February 2021*

DOCUMENT CHANGE RECORD

Version	Date	Changes
2.0	*10th January 2021*	*First version of the document for v2021* *Track of changes from v2018.1* *Submitted to DRR*
2.0.1	*28th February 2021*	*Document modified after DRR outcome + typos correction*

Table of Contents

List of Tables and Figures

		Validation report of the Convection Product Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
NWC SAF	METEO FRANCE		Issue: 2.0.1 Date: 28th February 2021
			File: NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
			Page: 7/75

1 INTRODUCTION

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

1.1 SCOPE OF THE DOCUMENT

This document is the convection product validation report applicable to NWC/GEO software package v2021. The accuracies of the convection products (GEO-CI, Convection Initiation and GEO-RDT-CW, Rapidly Developing Thunderstorm Convection Warning) are discussed.

1.2 SOFTWARE VERSION IDENTIFICATION

This document describes the products obtained from the GEO-CI v2.2 (Product Id NWC-089) and from GEO-RDT-CW v5.2 (Product Id NWC-090) implemented in the release 2021 of the NWC/GEO software package.

1.3 REQUIREMENTS

Skill requirements had been expressed in PRD Table for RDT and CI (see [AD.4.]).

- CI:
  - Accuracy: FAR<0.6 POD>0.4 for +30' ahead
  - Target FAR<0.5 POD>0.5 for +30' ahead
  - Optimal: FAR<0.4 POD>0.7 for +30' ahead

- RDT:
  - Accuracy
    - 1) early detection (before first lightning occurrence) 10%
    - 2) 30 minutes after first lightning occurrence 30%
    - 3) overall thunderstorm detection skill 50%
  - Target
    - 1) early detection (before first lightning occurrence) 25%
    - 2) 30 minutes after first lightning occurrence 50%
    - 3) overall thunderstorm detection skill 70%
  - Optimal
    - 1) early detection (before first lightning occurrence) 50%
    - 2) 30 minutes after first lightning occurrence 75%
    - 3) overall thunderstorm detection skill 90%

1.4 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

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

1.5 REFERENCES

1.5.1 Applicable documents

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

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

Current documentation can be found at the NWC SAF Helpdesk web: http://nwc-saf.eumetsat.int.

Table 1: List of 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*	*1.6*	*1/12/2021*
[AD.3.]	*Configuration Management Plan for the NWCSAF*	*NWC/CDOP3/SAF/AEMET/MGT/ CMP*	*1.1*	*1/4/2021*
[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/CDOP3/GEO/AEMET/SW/ SCRD*	*1.0*	*1/9/2021*
[AD.6.]	*Interface Control Document for Internal and External Interfaces of the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ ICD/1*	*2.0*	*1/9/2021*
[AD.7.]	*Interface Control Document for the NWCLIB of the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ ICD/2*	*2.0*	*1/9/2021*
[AD.8.]	*Data Output Format for the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ DOF*	*1.1*	*1/10/2019*
[AD.9.]	*Component Design Document for the NWCLIB of the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ ACDD/NWCLIB*	*2.0.1*	
[AD.10.]	*Component Design Document for the Convection Product Processors of the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ ACDD*	*2.0*	*1/9/2021*
[AD.11.]	*Algorithm Theoretical Basis Document for the Convection Product Processors of the NWC/GEO*	*NWC/CDOP3/GEO/MFT/SCI/ ATBD/Convection*	*1.0*	*1/9/2021*
[AD.12.]	*User Manual for the tools of the NWC/GEO*	*NWC/CDOP3/GEO/AEMET/SW/ UM/Tools*	*1.0*	*21/1/2019*
[AD.13.]	*NWC SAF CDOP-3 Project Plan Master Schedule*	*NWC/CDOP3/SAF/AEMET/MGT/ PP/MasterSchedule*	*1.1*	*28/2/2018*

1.5.2 Reference documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable ones, and can be looked up to enhance the

information included in this document if it is desired. They are referenced in this document in the form [RD.X]

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

Current documentation can be found at the NWC SAF Helpdesk web: http://nwc-saf.eumetsat.int.

Table 2: List of Referenced Documents

Ref	Title	Code	Vers	Date
[RD.1.]	The Nowcasting SAF glossary	NWC/CDOP3/SAF/ AEMET/MGT/GLO	1.0	20/10/2020
[RD.2.]	Best Practice Document, 2013, for EUMETSAT Convection Working Group, Eds J.Mecikalski, K. Bedka and M. König »	Available on CWG Website		2013
[RD.3.]	Scientific Report on verification of RDT forecast	NWC/CDOP3/GEO/MFT/ SCI/RP/01	1.2	5/9/2018
[RD.4.]	Lenk, S., Senf, F., Deneke, H., Final Report on the Associated Scientist Activity for the Validation of the Convection Initiation (CI) product of NWCSAF v2018	NWC/CDOP3/SAF/MF-PI/SCI/RP/ CI_Improv_Tropos available on NWCSAF Website	1.0	November 2018
[RD.5.]	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
[RD.6.]	Validation report of the Convection Product Processors of the NWC/GEO	NWC/CDOP3/GEO/MF-PI/SCI/VR	1.0	15th October 2016
[RD.7.]	Scientific Report on Extended Validation of RDT	SAF/NWC/CDOP/MFT/ SCI/RP/1		7th December 2012
[RD.8.]	Haberlie, A. M., Ashley, W., S., Pingel, T. J., 2014, The effect of urbanisation on the climatology of thunderstorm initiation, QJRMS, https://doi.org/10.1002/qj.2499			2014
[RD.9.]	Walker, J. R., W. M. MacKenzie, Jr., J. R. Mecikalski, and C. P. Jewett, 2012: An enhanced geostationary satellite-based convective initiation algorithm for 0/2-h nowcasting with object tracking. J. Appl. Meteor. Climatol., 51, 1931-1949"			2012
[RD.10.]	Autonès, F., Moisselin, J.-M., **2019**, Validation Report for Convection Products, available on http://nwc-saf.eumetsat.int	NWC/CDOP3/GEO/MFT/ SCI/VR/Convection	1.0	2019
[RD.11.]	https://www.vaisala.com/fr/products/data-subscriptions-and-reports/data-sets/gld360			
[RD.12.]	Brenguier, J.-L, Bouttier, F., Moisselin, J.-M., 2015, Les nouveaux services météorologiques pour l'aviation, La Météorologie 91, Novembre 2015			2015
[RD.13.]	Le Gléau, H., 2016, Algorithm Theoretical Basis Document for the Cloud Product Processors of the NWC/GEO, available on http://nwc-saf.eumetsat.int	NWC/CDOP2/GEO/MFL/ SCI/ATBD/Cloud		2016
[RD.14.]	Moisselin, J.-M., 2017, Scientific report on on future verification of CI product, available on http://nwc-saf.eumetsat.int	NWC/CDOP2/GEO/MFT/ SCI/RP/08		2017

Ref	Title	Code	Vers	Date
[RD.15.]	Setvák, M., Radová, M., Kaňák, J., Valachová, M., Bedka, K, Šťástka1, J., Novák, P., Kyznarová, H., 2014, Comparison of the MSG 2.5-minute rapid scan data and products derived from these, with radar and lightning observations, 2014 EUMETSAT Meteorological Satellite Conference, 22 - 26 September 2014, Geneva, Switzerland			2014
[RD.16.]	Roberts, N.M. and H.W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.,* **136**,78-97.			2008
[RD.17.]	Stein and Stoop, 2018, Neighborhood-based contingency tables including errors compensation, https://www.researchgate.net/publication/327808409_Neighborhood-Based_Contingency_Tables_Including_Errors_Compensation			2018

2   CONVECTION INITIATION (GEO-CI) VALIDATION

2.1   OVERVIEW

2.1.1   General objectives of the validation

The main objective of this section is to document Convection Initiation product accuracy. The method used has to compare POD and FAR scores to the threshold accuracies listed in the NWCSAF product requirements document [AD.4] : FAR<0.6, POD>0.4

2.1.2   Methodology outline

CI product verification is performed with both objective scores and case studies. The ground truth in objective scores is a radar-based product in object mode, providing key information about the birth of the cells.

The objective validation study focused on nine days with convection over France in June 2021. The verification is rather strict and the definition of the ground truth is derived from a Météo-France radar-based product.

2.2   SPECIFICITIES OF CI VERIFICATIONS

The verification method will have also a high impact on scores (see [RD.14.])

- Case study for convective days or large period,

- Double penalty issue taken into account or not,

- Threshold above which CI is verified (25%, 50% 75%),

- Day verification or day and night verification,

- Object or pixel-based verification,

- Verification of tracked clouds or verification of all clouds.

- CI below cirrus taken into account or not.

- Use of a rapid scan imager or not

For example in [RD.9] the POD varies from 0.32 to 0.72 depending on the way CI is verified.

Additionally high FAR are often attributed to difficulties inherent to CI problem. For example one CI object can dominate all other CI objects in the surrounding, as low-level convergence and upper-level divergence suppress other up-drafts. Very large FAR values can be found in literature.

2.3   OBJECTIVE VALIDATION

2.3.1   Context

Tropos (Leibniz Institute for Tropospheric Research) performed in Autumn 2018 a verification of CI v2018 in the framework of an Associated Scientist activity [RD.4] with German radar data

		Validation report of the Convection Product Processors of the NWC/GEO	Code: NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
NWC SAF	METEO FRANCE		Issue: 2.0.1 Date: 28th February 2021
			File:NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
			Page: 14/75

v2021 version is validated by MFT with radar data over several territories following some [RD.4] methods but introducing some novel approaches like the use of radar-based convective objects that are already produced. Following sections describe it.

2.3.2 Data and methods

The objective validation study focused on nine days with convection over France in June 2021. The verification is rather strict and the definition of the ground truth is derived from a Météo-France radar-based product.

2.3.2.1 Data

2.3.2.1.1 Case studies meteorological context

Nine days in June 2021 are used to evaluate the CI. They offer both rather sparse or isolated short-living thunderstorms (20210606, 20210611, 20210618, 20210621, 20210622, 20210623) and long-living widespread ones (20210616, 20210619 and 20210620).

2.3.2.1.2 Ground truth used for the validation

Convective initiation is most of the time defined as the first occurrence of a radar precipitation echo of intensity reaching or exceeding 30 to 40 dBZ (RD.2), most often 35 dBZ (RD.9).

Météo-France produces routinely radar-based convective objects at the radar temporal frequency (5 minutes) and at the radar spatial resolution (1 km) over France. The radar objects delineate convective cells with an outline at 32 dBZ radar echo. Thus, the product includes cell tracking, so the identification of trajectory of cells and the birth of convective cells, namely the radar convection start.

The radar-derived convection start used to validate the NWC SAF CI product is then defined as the footprint of the first convective cell belonging to a trajectory. The trajectory has to meet requirements in terms of severity (reflectivity threshold and / or lightning impacts paired with the trajectory) and lifetime to meet the conceptual scheme of the thunderstorm life cycle. The Figure 1 recaps the ground truth definition.
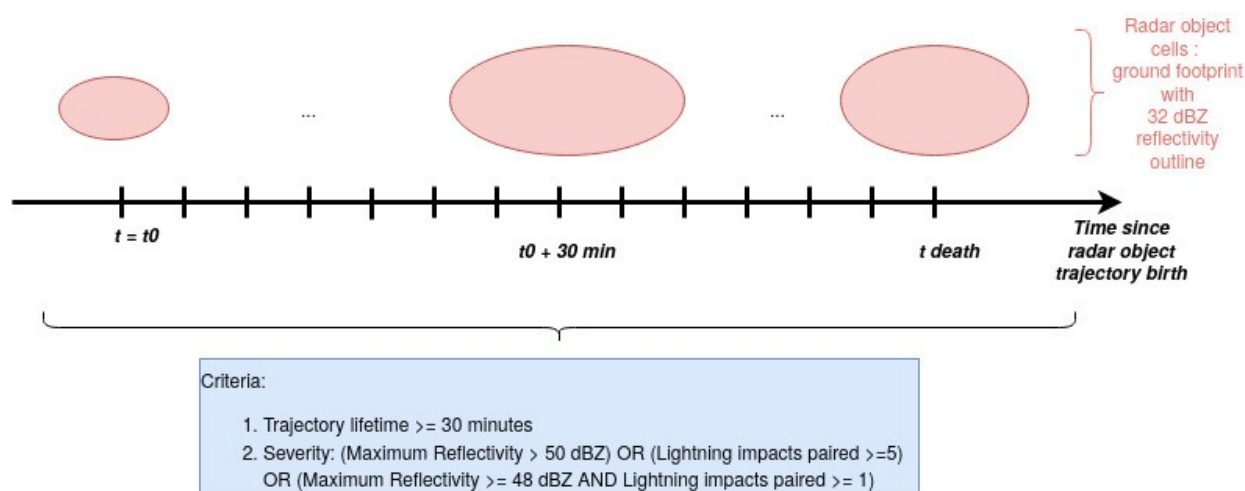
Figure 1: Ground truth defined from radar-derived convective objects to validate NWC SAF CI product.

2.3.2.1.3 NWC SAF CI product description

It is operated with data from MSG Primary service mode (MSG4 satellite) and from MSG Rapid Scan Service mode (MSG3 satellite for the first seven days and MSG2 for the last two days - 20210622 from 0900Z onwards and 20210623). In order to compare the new v2021 version with the previous one, CI is run in nominal configuration with v2018.1 code and v2021 code. Finally, CI is produced over quite a small domain covering France and most of the French meteorological radar coverage (Figure 2).
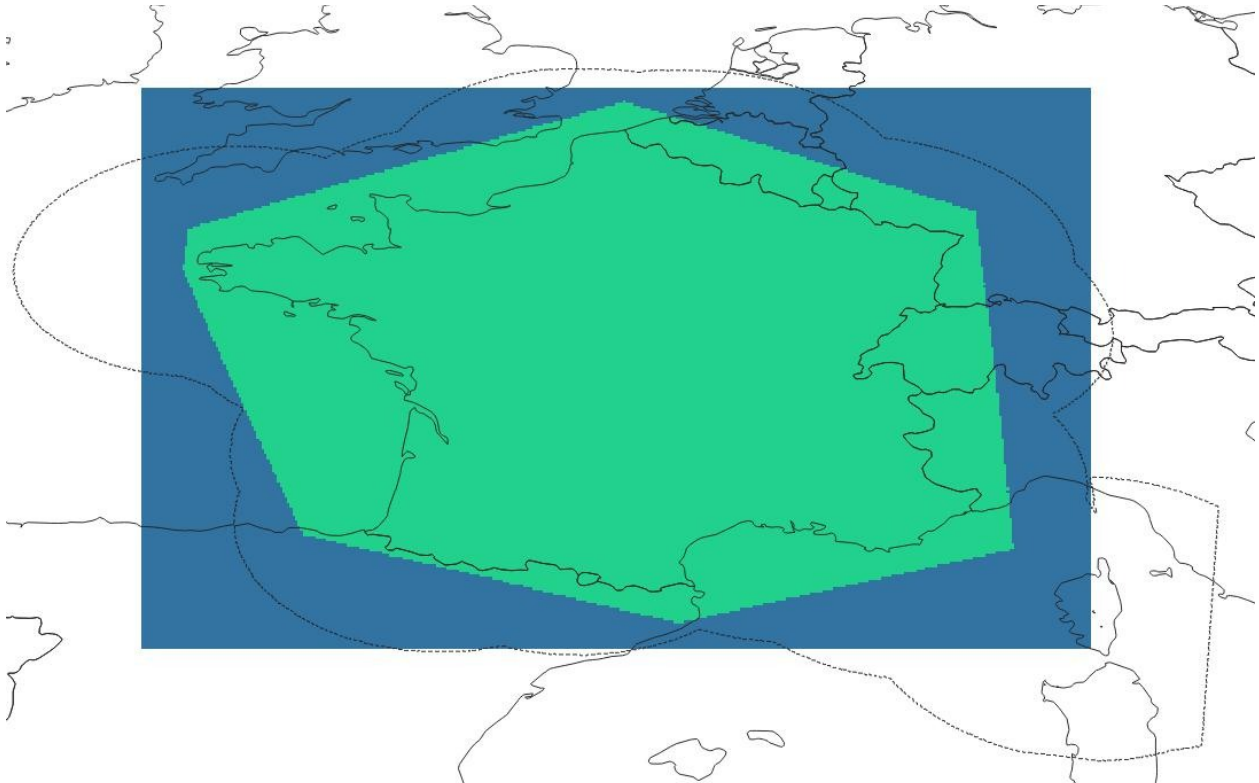
Figure 2: Blue rectangle depicts the MSG4 CI production domain, dotted black line the radar coverage and green polygon the area where statistics are computed.

2.3.2.2 Method

Ground truth footprints (radar polygons) are rasterized as binary events and remapped on the satellite grid (both MSG Rapid Scan Service and MSG Primary Service grid). 'ci_prob30' container of CI product is also transformed in a yes/no product.

As radar and satellite signals are quite different and as CI is quite a small localized phenomenon, it is necessary to take pixel neighbourhood into account in the computation of the statistical scores. Neighbourhood-based contingency tables (RD.15) are implemented for that purpose. The main advantage of these tables is to offset the double penalty issue by counting the observed and the forecast frequencies inside a window of n*n pixels.

What's more, the Fraction Skill Score (FSS, RD.16) is implemented to score the CI product against the ground truth and to compare previous and current CI versions. On top of that, bootstrap techniques are used to evaluate the confidence interval of the scores.

The validation is performed every 15 minutes. Ground truth data are then aggregated on a 15-min basis, so as the CI product run with data from the MSG Rapid Scan Service mode.

2.3.2.3 Results

2.3.2.3.1   Balance between the four levels of 'ci_prob30' probability

v2018.1 CI version suffered from a lack of medium and high levels of CI probability (levels 3 and 4 corresponding to medium and high level), with a probability of both levels below 1 %. V2021 CI version partially corrects the imbalance between the four different classes of probability (Figure 3). The low level (level 2) is still prevailing at the expense of, in particular, the medium level (level 3). This behaviour is especially true for the CI run with MSG Rapid-Scan service mode data. The adaptation to all satellite scan update rates and the improvement made in the pixel trends computation in CI v2021 version code are the principal reasons explaining the reduction in the imbalance.
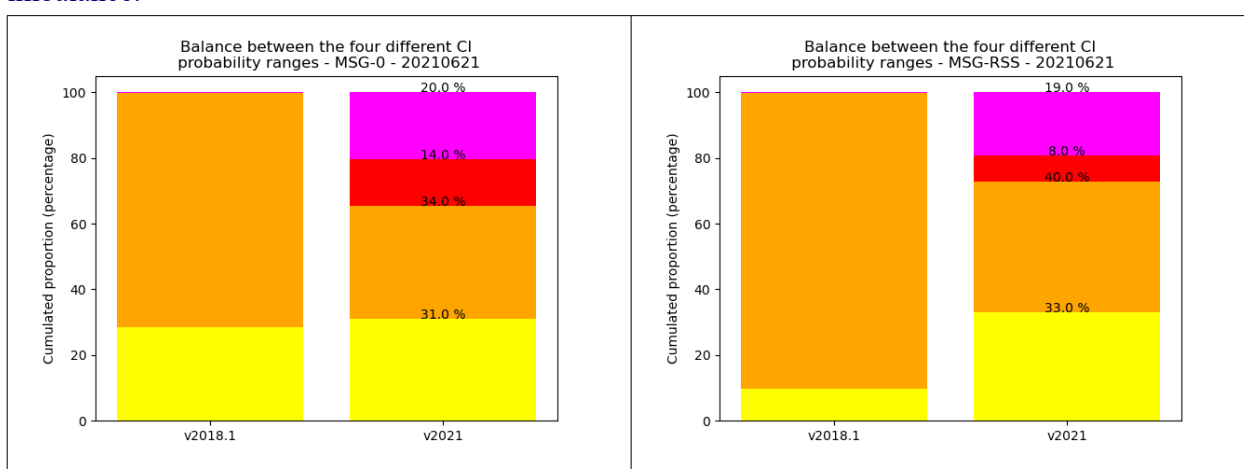
Figure 3: Balance between the four different levels of CI probability on the 20210621 case study, with MSG Primary Service data (left) and MSG Rapid Scan data (right). CI v2018.1 (left bar) and v2021 (right bar) are compared. Yellow, orange, red and magenta colours correspond to the very low, low, medium and high levels of CI probability.

Table 3: Proportion of CI v2021 levels of probability for each case study and for each satellite data (left values: MSG Primary service data, right values: MSG rapid scan service data)

Case Study / level of probability	Very low (yellow / 1)	low (orange / 2)	medium (red / 3)	high (magenta / 4)
20210606	43 % / 39 %	38 % / 43 %	07 % / 07 %	12 % / 11 %
20210611	23 % / 23 %	38 % / 45 %	18 % / 12 %	21 % / 20 %
20210616	29 % / 28 %	38 % / 41 %	09 % / 07 %	24 % / 23 %
20210618	32 % / 32 %	39 % / 43 %	14 % / 09 %	15 % / 16 %
20210619	32 % / 31 %	36 % / 41 %	11 % / 08 %	21 % / 19 %
20210620	32 % / 32 %	37 % / 41 %	09 % / 06 %	22 % / 21 %
20210621	31 % / 33 %	34 % / 40 %	14 % / 08 %	20 % / 19 %
20210622	35 % / 37 %	36 % / 39 %	14 % / 09 %	16 % / 15 %
20210623	35 % / 33 %	38 % / 39 %	13 % / 11 %	15 % / 16 %

2.3.2.3.2   Statistical scores.

Next Figure summarizes the False Alarm Ratio, the Probability of Detection and the fraction skill score for the 20210621 case study.
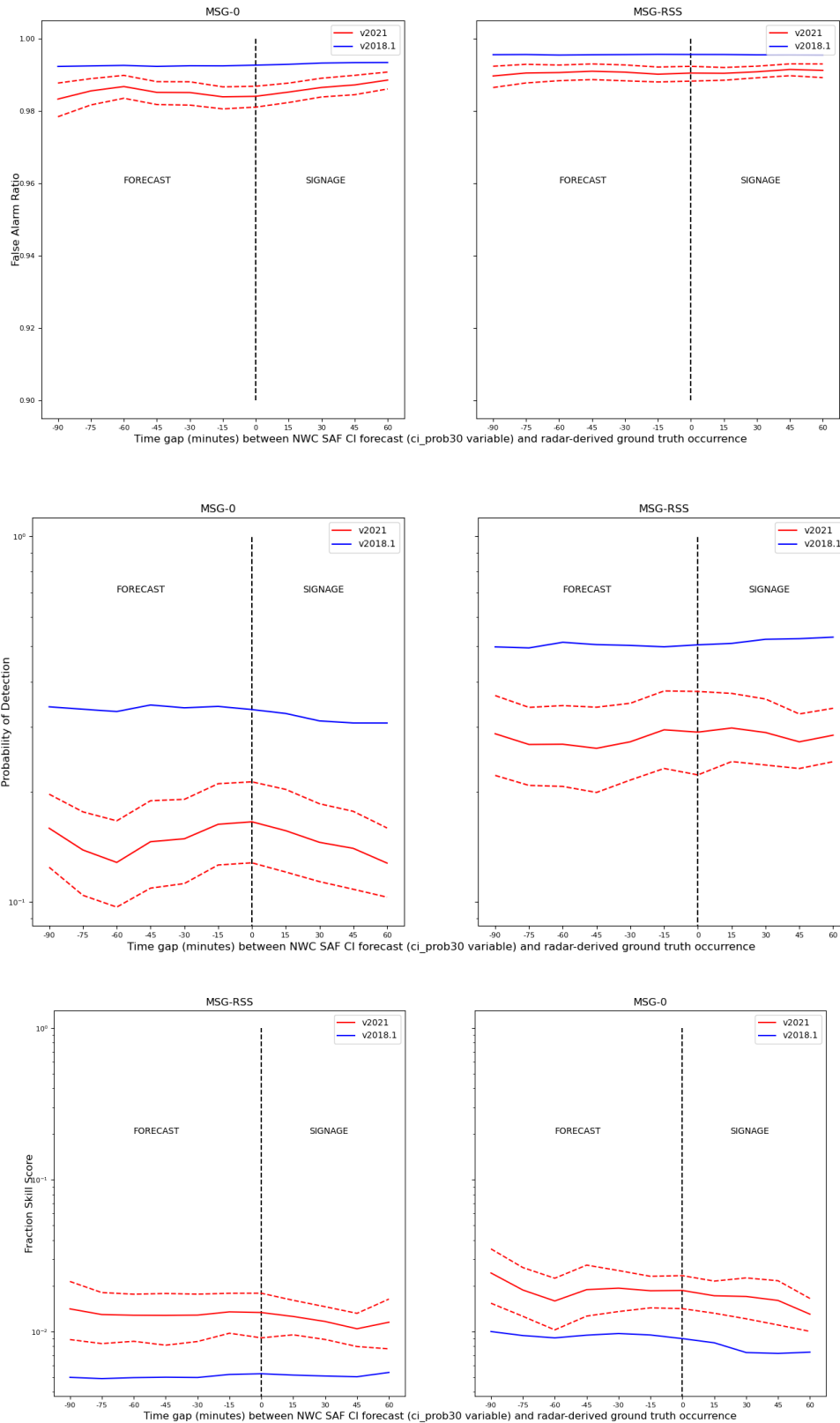
Figure 4: Neighbourhood-based statistical scores for the 20210621 case study. From top to bottom: FAR (False Alarm Ratio), POD (Probability of Detection), FSS (Fraction Skill Score). The red curve corresponds to the CI v2021 version and the blue curve to the CI v2018.1 one. Dotted red lines give the 95 % confidence interval (bootstrap technique). Window size is 11 pixels.

2.3.2.3.3 FAR False Alarm Ratio.

New v2021 version significantly reduces the false alarm ratio compared to the v2018.1 version whatever the satellite data used to produce CI. A ratio of 3 to 8 times less CI pixels is observed in the v2021 version compared to the v2018.1 version.

Anyway the FAR is still very high, over 0.98. In other words, frequency biases range from 4 to 22 with MSG-0° data between CI run in v2021 version and the ground truth, and from 12 to 95 with MSG-RSS data (not shown).

The window size impacts the FAR differences between v2021 and v2018.1 version: the bigger the window, the more significant the differences (not shown).

2.3.2.3.4 POD Probability of Detection

New v2021 version significantly degrades the probability of detection compared to the v2018.1 version whatever the satellite data used to produce CI. POD ranges between 0.15 and 0.3 with v2021 version and between 0.3 and 0.5 with v2018.1 version depending on the time gap between the CI and the ground truth. The comparison between RSS curves and FDSS ones clearly exhibit that the use of RSS goes toward an improvement of detection.

POD values' differences between v2021 and v2018.1 tend to remain stable (rarely bigger) when the window size increases (not shown).

2.3.2.3.5 FSS Fraction Skill Score

New v2021 version improves the FSS compared to the v2018.1 version whatever the satellite data used to produce CI. FSS is above 0.01 for v2021 version and well below 0.01 for v2018.1 version for the 20210621 case study. FSS is strongly affected by FAR. The improvement in FSS is bigger when running v2021 with MSG RSS data than with MSG FDSS data, compared to v2018.1 runs.

The FSS value is quite similar for all time gaps between CI product and radar ground truth occurrence. There is no more forecast skill than signage skill (when convection has started).

FSS values' differences between v2021 and v2018.1 tend to increase when the window size increases (not shown). It follows the FAR values' differences.

Same conclusions can be drawn from the other case studies (not shown). Only POD and FSS values may differ, depending on the case study. For some case studies (20210611, 20210616, 20210618), FSS improvement is not significant when running v2021 version with MSG FDSS data.

2.3.2.3.6 Time evolution of CI and ground truth

The first two hours of the day should not be taken into account (warm-up period). The number of CI pixels tend to be too large in particular during the night until the dawn (included) and at the end of the day (Figure 5). This is a common feature for days observing a diurnal cycle of convection. For days with long-living widespread thunderstorm systems, the evolution is more stable (not shown).
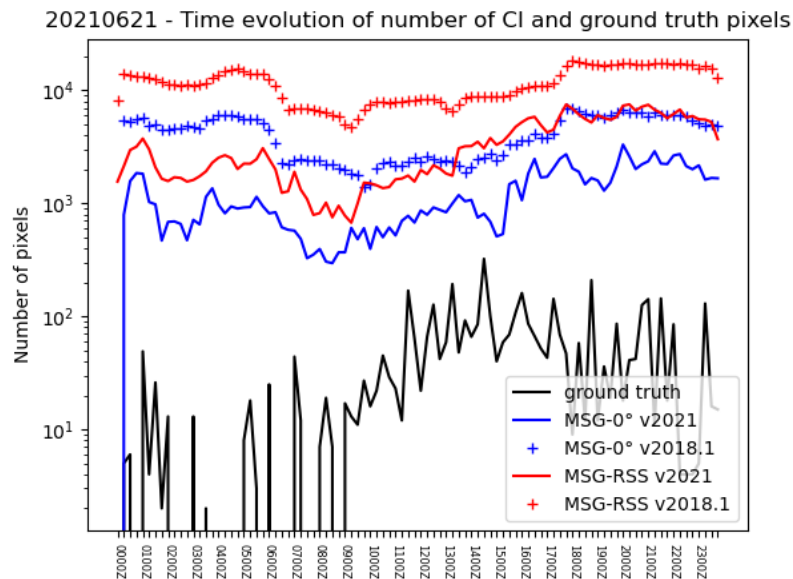
Figure 5: Time evolution of number of CI and ground truth pixels for 20210621 case study

2.3.2.3.7 Summary

As it was already highlighted in the previous CI version, false alarms in CI are the main issue. A work was first needed and done to reduce the number of false alarms but there is still room for improvement particularly with the CI diagnosis. The reduction in false alarms has the drawback of degrading the POD. Overall the fraction skill score seems better with the new v2021 version, but the statistical significance is not reached.

2.3.2.4 Assumptions and limitations

The method assumes that satellite and radar convective initiation signals might be compared, even though they are intrinsically different. Spatial resolutions differ by a minimum factor of 3 (so 9 in surface) with MSG data (1km with radar data vs 3km with satellite data at satellite sub-point). Moreover, radar signal focuses on the convective core of the convective cell. By construction, CI satellite signal may appear at successive times, while radar ground truth only once. A 15-min basis data integration is used to mitigate this issue with MSG rapid-scan data.

CI product sets a diagnosis on cloudy pixels belonging to the categories ranging from very low to medium clouds ; ground truth may occur in already developed thunderstorms and might be the result of the split of large-size long-living thunderstorm systems. So some ground truth footprints may be large or not consistent with what we are looking for in CI product. Case studies with isolated unorganized thunderstorms are preferred.

No minimum distance is considered in the ground truth; because of the remapping on the satellite grid at a coarser resolution than radar resolution, two radar cell footprints may appear on the same pixel of the grid.

2.4 CASE STUDIES

2.4.1   20210621 MSG Case Study

Figure 6 illustrates a MSG case study. Early in the morning (0530Z, top), already existing large-scale convective storms affect northern parts of France while small-scale cells hit southern part. New radar cells (ground truth) form in the southwestern part of France. Some CI good detections appear in the vicinity of them. CI false alarms are detected in the western part of France. Additional information about the meteorological context (trough at the edge of the Atlantic coast) may help the forecaster not to take into account those signals.

Later in the morning (0930Z, bottom), CI product offers good skill to detect diurnal evolution of convection over land. CI pixels are quite well collocated with new developing radar convective cells. False alarms are detected, especially over Spain (it is to note that radar data in the validation only marginally cover Spain). Over Spain, in a north-westerly flow (colder air mass), CI product may have difficulties to catch the quick displacement of low levels clouds, hence wrong values of growth parameters and an unexpected CI diagnosis. We may notice that CI pixels appear at the edge of cloud systems.

Figure 7 illustrates the same case study with MSG rapid scan data as the one with MSG data. Early in the morning (0530Z, top), the two CI products behave similarly. More false alarms appear on the edge of the cold large-scale convective system over northern part of France and United Kingdom. Again, cloud movement tracking and then growth parameters' values are probably responsible for that. In contrast, later in the morning (0930Z, bottom), less false alarms are detected over Spain. With the diurnal evolution, good CI detections are observed over southern and western part of France.
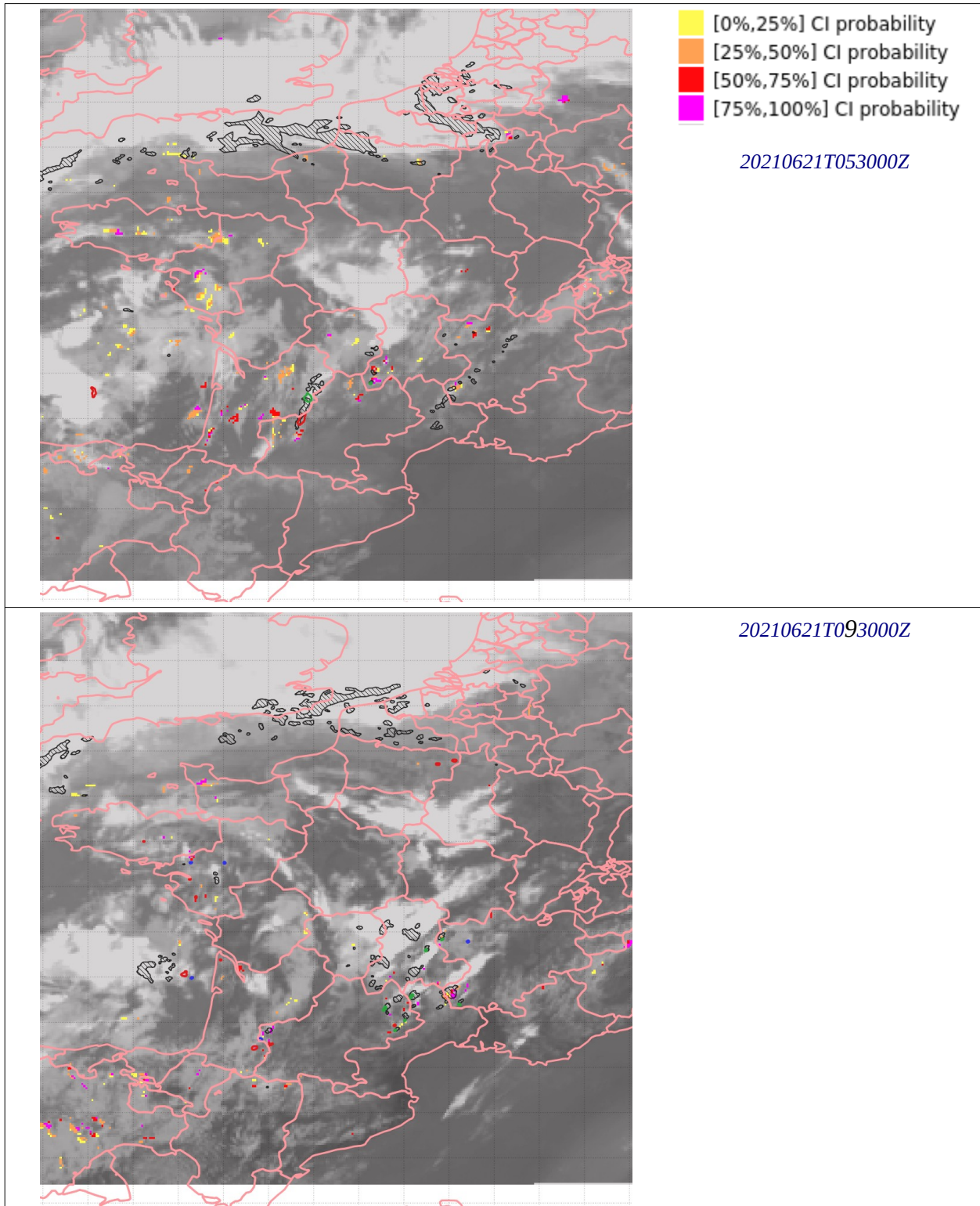
Figure 6: NWC SAF CI superimposed on SEVIRI IR10.8µm MSG data. Radar-derived convective objects (32 dBZ contours) are represented with black hatches. Ground truth is represented with red (birth within the next [0;30] minutes time interval) or green contours (birth within the last [0;30] minutes time interval.

Figure 7: Same as Figure 6 with MSG Rapid scan data (CI product and IR10.8µm)
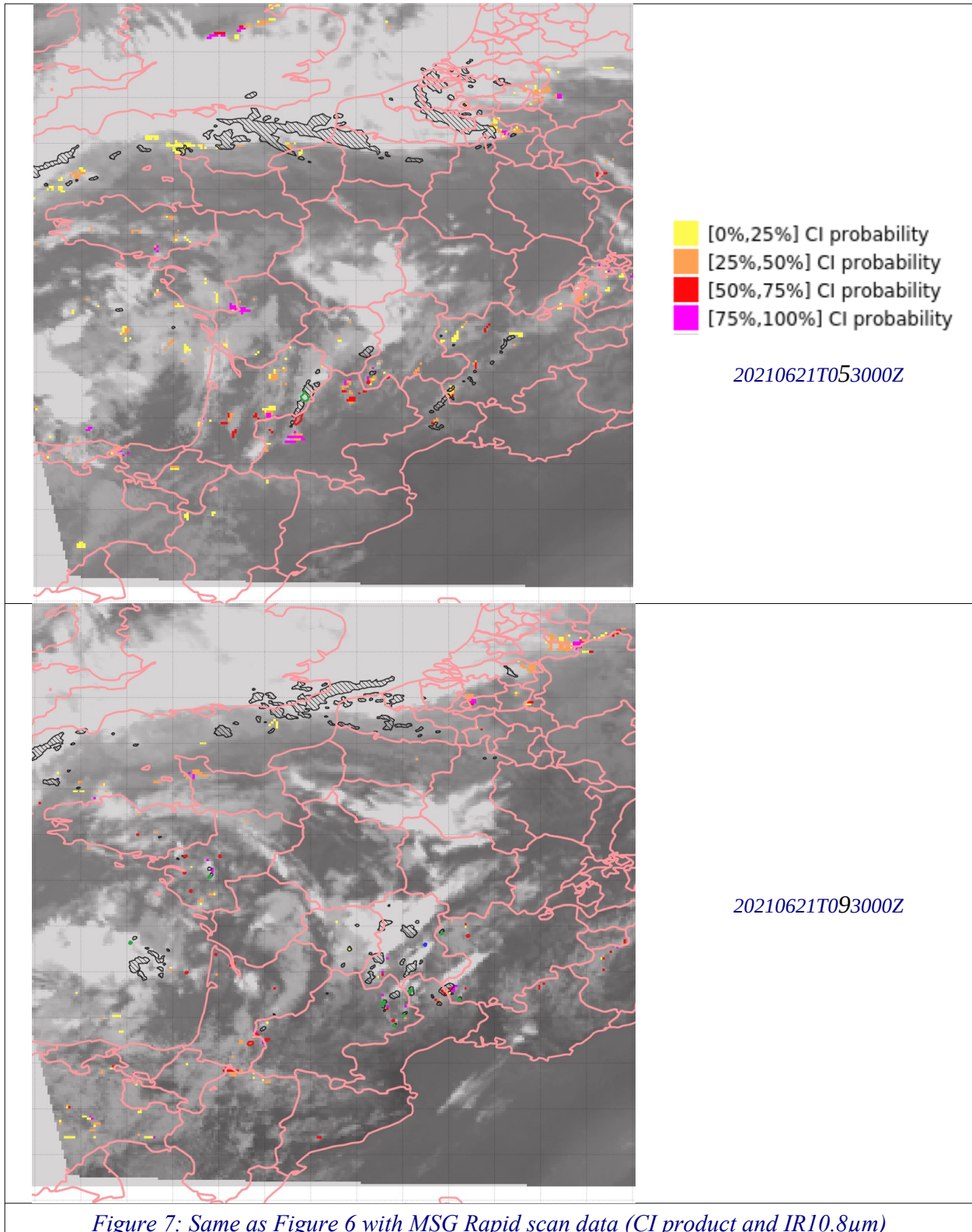
2.4.2 20210207 MSG-OI Case study

Figure 15 illustrates a case study over the Indian Ocean, with a focus on La Reunion Island (located at the centre of the image, pink contour). It highlights the difficulty to gain information from CI product on some situations with ground truth embedded in cold convective systems. The appearance

and evolution of new Radar ground truth is already hard to grasp in such situations; CI pixels are not linked at all with the ground truth. Cloud movement is probably hard to track, giving a lot of CI false alarms. Some good detections may be noticed on the southern part of La Reunion Island, at 1400Z.

Figure 8: Same as Figure 6 for 20210207 case study and MSG-IODC data

2.4.3   20210914 GOES16 Case Study

Validation of GOES16 CI product relies on the analysis of 20210914 case study over French Guyane.

Figure 9 depicts the diurnal evolution over French Guyane with the forecaster workstation. First radar convective cells (radar echos above 32 dBZ) are concomitant with the first CI signals at 1430Z if we accept a 2-3 pixel tolerance. The convective cells evolve first in low clouds. Lightning impacts occur one hour and a half after the first CI signals (1600Z). When the storm south of Cayenne becomes mature, CI pixels appear downstream, in the convergence area between the density current of the storm and the environment, where thunderstorm is about to move forward (1630Z to 1800Z). Subjectively, it looks like a good detection even if in the objective validation, it would have been considered as a false alarm.
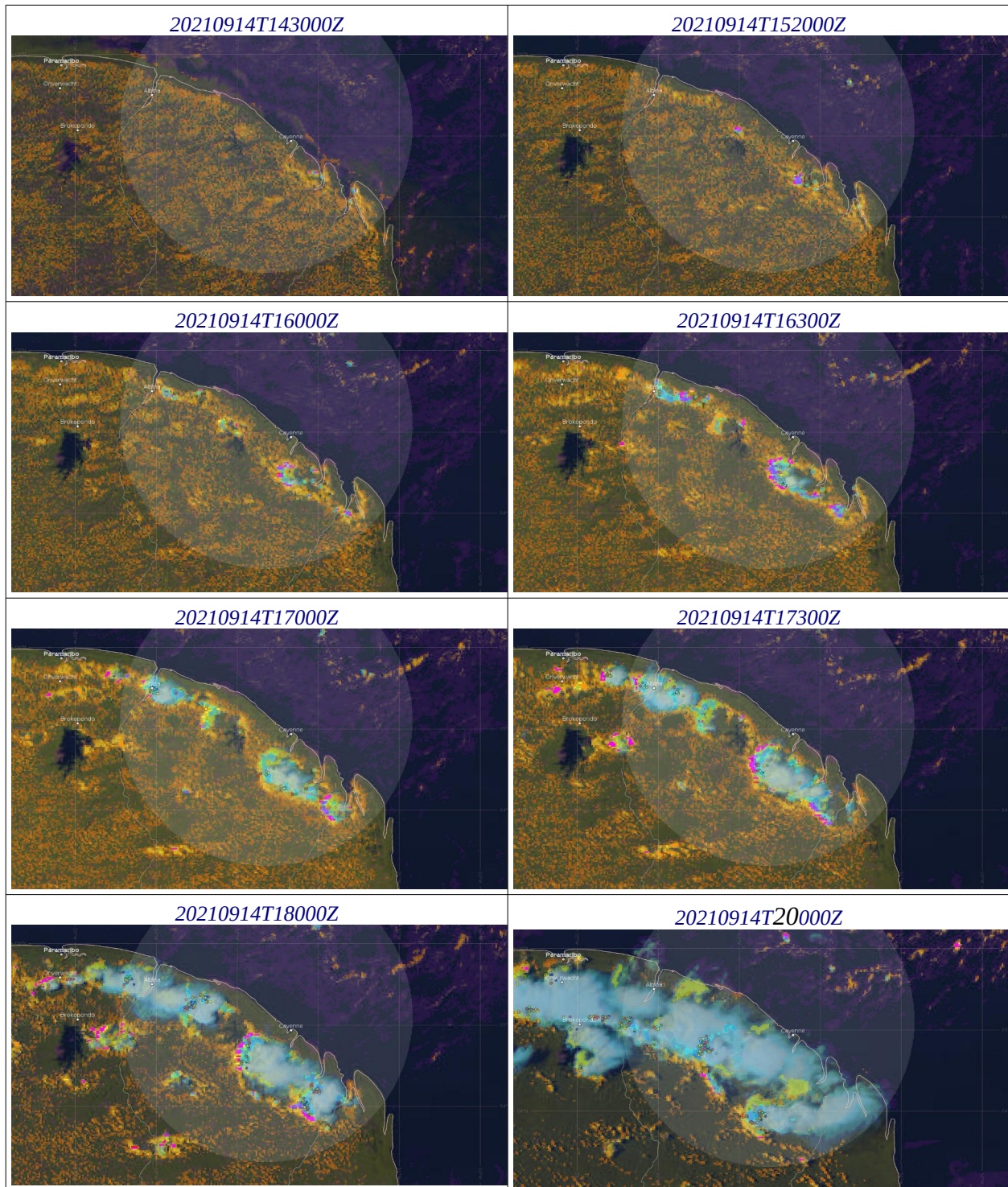
Figure 9: 20210914 case study over Guyane. Cloud type superimposed with radar convective objects (radar echos over 32 dBZ, in blue) and v2021 CI product with standard colours (yellow, orange, red and magenta for the 4 levels of probability)

The balance between the four different CI classes of probability is fulfilled over the whole day (Figure 10).
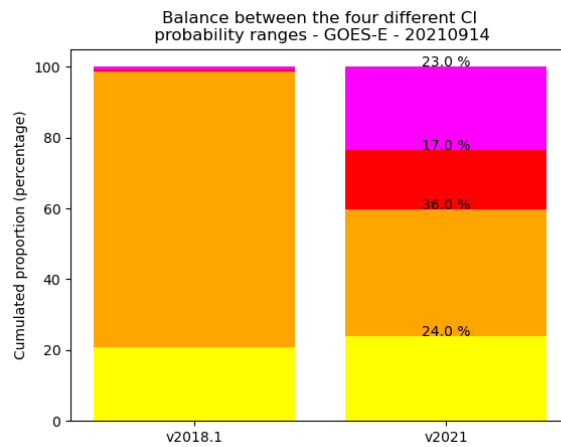
Figure 10: Same as Figure 7 for 20210914 case study with GOES16 data over French Guiana.

2.4.4 20210215 HIMAWARI Case Study

Figure 12 gives a focus on this case study centred on New Caledonia, where Météo-France operates meteorological radars. At 1000Z (top), CI product gives interesting features near existing radar convective cells and ground truth. It is worth noting that over tropical regions, convective cells trigger with less cloud extension than in mid-latitudes.

1700Z picture (bottom) depicts an important feature of CI. North of New Caledonia over the Pacific ocean, CI pixels are diagnosed on the northern part of a large convective system, where it regenerates. Attention should be paid in the future to keep such signals when further trying to remove CI false alarms on the edge of low cloud systems, detected as CI because of bad-assessed cloud movement and then wrong growth parameters' values.
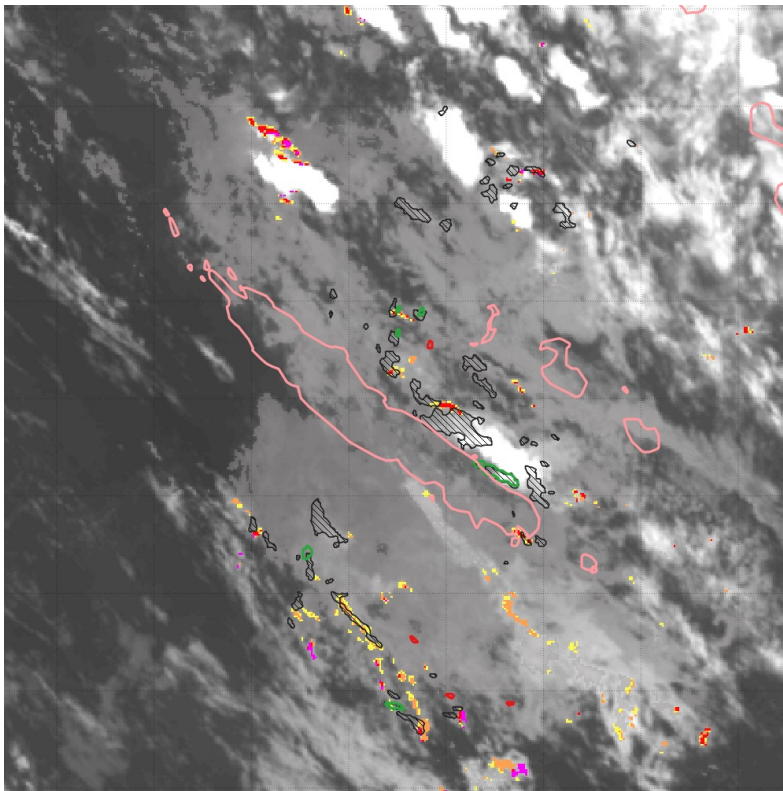
[0%,25%] CI probability
[25%,50%] CI probability
[50%,75%] CI probability
[75%,100%] CI probability

20210215T100000Z
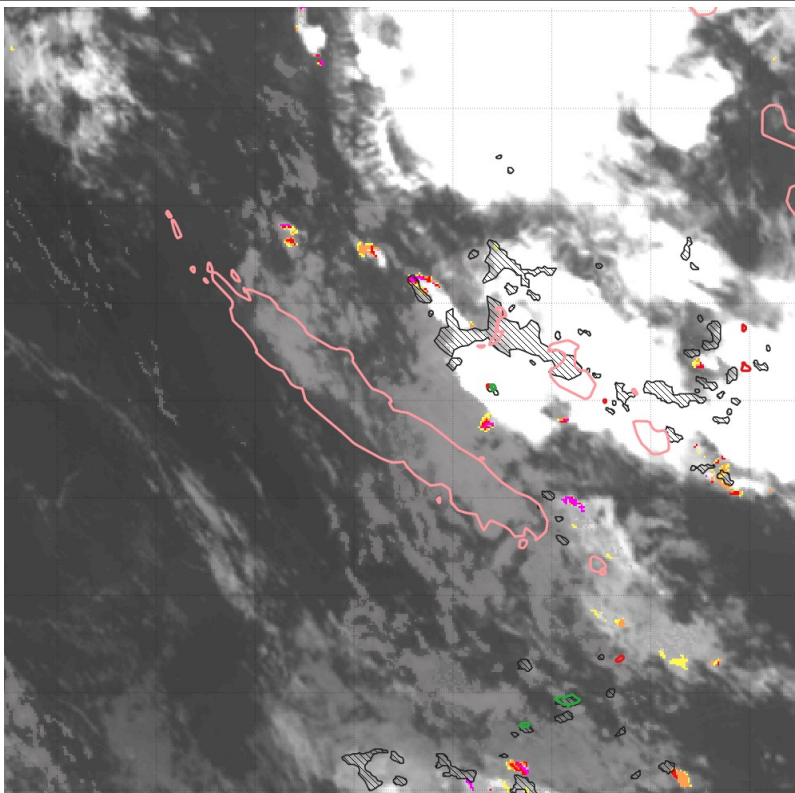
20210215T170000Z

Figure 11: Same as Figure 6 applied to 20210215 case study with Himawari data

	Validation report of the Convection Product Processors of the NWC/GEO	**Code:** NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
NWC SAF / **METEO FRANCE**		**Issue:** 2.0.1 **Date:** 28th February 2021
		File:NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
		Page: 28/75

2.4.5 20210914 GOES17 Case Study

Concerning GOES-W mission (GOES 17 at the moment of the validation) coverage, there are no territory where Météo-France operates radar. Thus the methods explained in previous paragraphs can not be used to assess the quality of CI product. The case study makes a verification of CI based on WWLLN lightning data over the United States (Colorado). It highlights the good capability of the CI to signal initiation of convection at the beginning of the diurnal cycle of convection, when the scene is clear or progressively fills up with growing cumulus.

On this case study, CI is produced with GOES17 ABI data at a 30 minutes frequency. As a consequence, and due to the way the decision tree is organised and the pixel trends are computed, CI diagnosis can not reach the two highest levels of probability (medium and high).
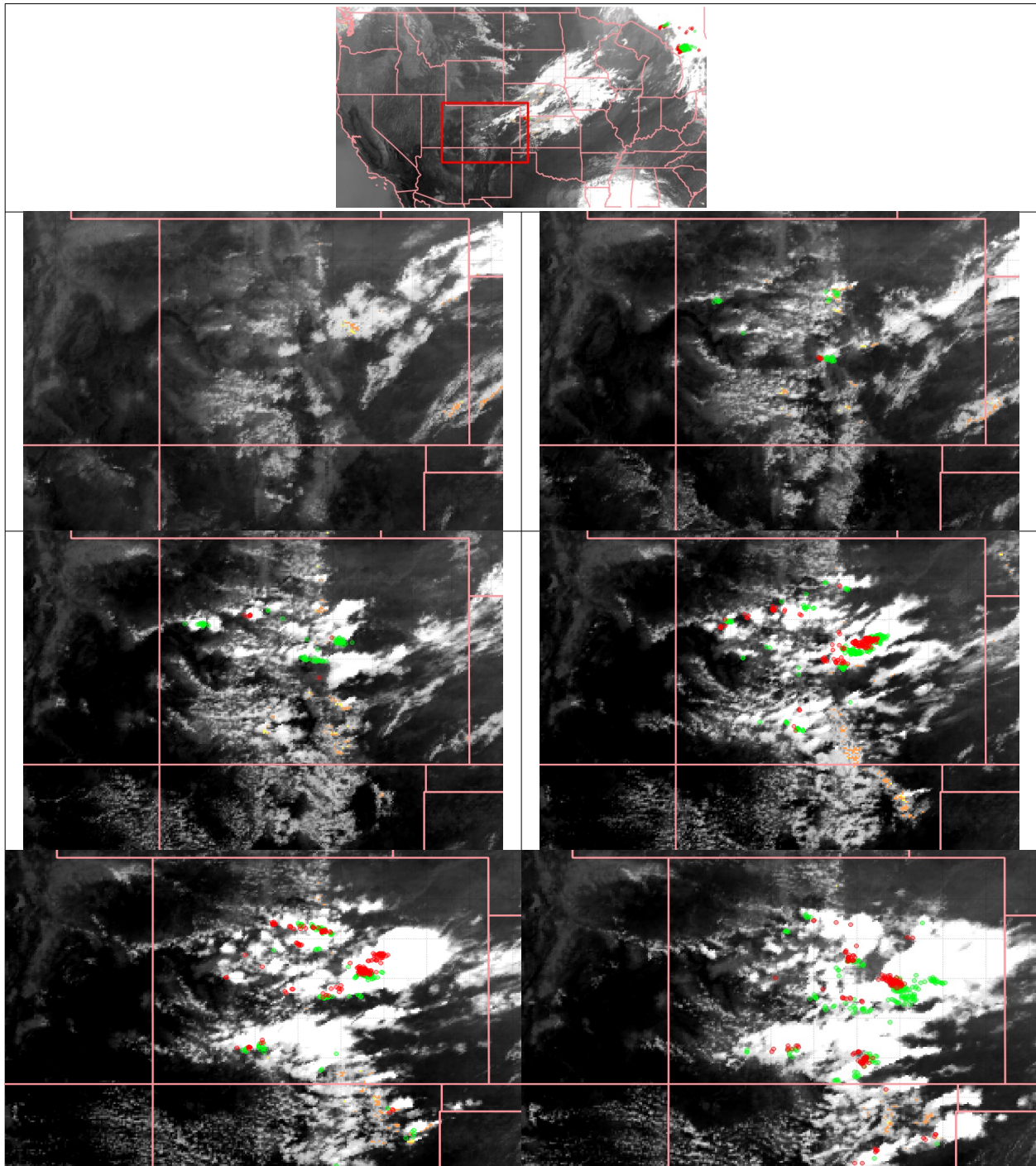
Figure 12: 20210914 GOES17 Case Study. First image describes the area on which we zoom in next images (red rectangle, covering the Colorado state). Then from left to right and top to bottom in 6 zoomed images, slots 20210914T1700Z, 20210914T1800Z, 20210914T1900Z, 20210914T2000Z, 20210914T2100Z, 20210914T2200Z. CI product overlaid on ABI IR11.2 channel and WWLLN lightning data (red circles, last 30 minutes occurrences and green circles, next 30 minutes).

2.5 CONCLUSION AND COMPLIANCE REQUIREMENTS

After CI v2016 first delivery of the product, v2018 exhibited major improvement and an effort on validation has been made. Improvements are illustrated in [RD.5] and the v2018 reached the status

	Validation report of the Convection Product Processors of the NWC/GEO	**Code:** NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
NWC SAF METEO FRANCE		**Issue:** 2.0.1 **Date:** 28th February 2021
		File:NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
		Page: 30/75

"pre-operational". v2018.1 and v2021 brought new improvements of CI products and a larger set of committed satellite.

The change of objective validation method has an impact on scores lowering them compared to [RD.5]. But the scores based on new method exhibit some improvements when v2021 is compared to previous version looking scores like FSS. The high sensitivity of CI validation scores to the method used has been quoted from the first stage of CI development [RD.14]. High FAR are often attributed to difficulties inherent to CI problem. For example one CI object can dominate all other CI objects in the surrounding, as low-level convergence and upper-level divergence suppress other up-drafts. Very large FAR values can be found in literature.

The unbalance in probabilities in previous versions made impossible the use of CI in a probabilistic way, forced users to use CI as a Yes or No product and made almost meaningless the verification with probability scores. The unbalance issue is now almost fixed with version v2021 but the verification with probability scores has not started yet.

Case studies with isolated unorganized convection clearly help user to understand the behaviour of the product and exhibit some cases where CI can be useful with bad scores, sometimes for a too-early diagnosis. It illustrates also the difficulty for CI to provide behaviour in case of quick displacement (perturbation of trends computation). Lastly a case illustrates how a basic knowledge of the meteorological context helps to easily identify some false alarms pixels, an argument that tends to favour the use of CI v2021 by forecasters and suggests some directions for the improvement of future versions.

In conclusion, MFT considers that CI has been improved in v2021 and that the strength and weakness of the product are identified. There is still room to improve the product in terms of objective validation targets as the scores remains below the accuracy level.

3 RAPIDLY DEVELOPING THUNDERSTORM – CONVECTION WARNING (GEO-RDT-CW) VALIDATION

3.1 OVERVIEW

The main objective of this section is to document RDT convective discrimination accuracies and compare them to the threshold accuracies listed in the NWCSAF product requirements document [AD.4]. As the RDT discrimination scheme has been enriched with an additional calibrated scheme, new graphs and comments in this report will refer to this new approach (CAL). One still can refer to the previous validation of so-called generic (GEN) scheme over European areas, which can be still used in specific configurations (un-calibrated satellite, invalidated statistical model).

Concerning the forecast capabilities (forecast products) included in RDT-CW code (CW part) it is to note that thunderstorm conceptual models often show a rapid morphological evolution and intensity variability, for which satellite data doesn't bring enough information. A subjective evaluation based on cases study is undertaken for an analysis of the localization of extrapolated cloud cells, depending on moving speed estimate accuracy and morphological evolution of the cloud systems. An objective validation has been made in a scientific report [RD.3].

3.2 VALIDATION OF GEN DISCRIMINATION DIAGNOSIS

3.2.1 Context

The GEN discrimination diagnosis has been fully described in previous validation report and results are still available as this choice of validation is still proposed to end-users. Hereafter a summary of the validation.

3.2.2 Data and methods

The configuration of discrimination verification is following

- Domain Europe

- Ground truth given by EUCLID network

- Period including intermediate season: June-August 2008 + April-October 2009

- Section approach where trajectory is split in convective and non convective sections

- Trajectories with light electric activity suppressed of the sample

- H2 hypothesis where RDT-CW early diagnosis (before an electric section) and continued diagnosis (after an electric section) have a positive impact on scores. It meets the requirement of "30 minutes detection after first lightning occurrence"

- RDT v2011 operated without lightning data to force the discrimination process.

- Results applicable to this release with GEN discrimination option

3.2.3   Results

Table 4: RDT v2011 Discrimination skill table

POD	77 %
POFD	4 %
FAR	28 %
TS	59 %

The figure below points out that more than 50% of good detection are already classified at the time of the first lightning occurrence, 80 % thirty minutes after. Nevertheless, only 25% are classified before the first flashes stroke (15 min before).
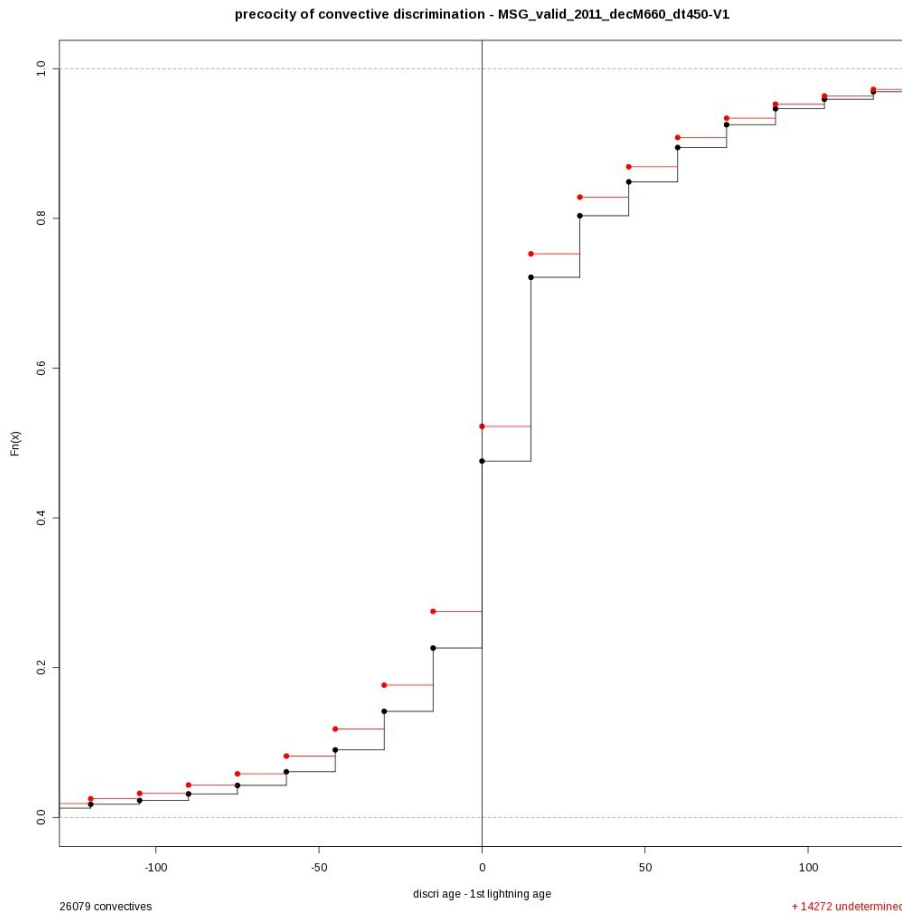
Figure 13: Precocity of RDT v2011 discrimination for moderate (black) and low (red marks) ground truths.

3.2.4   Conclusion

Threat Score is above 50% (above threshold accuracy of overall thunderstorm detection skill of 50%) and POD of 77% (above target detection 30 minutes after first lightning occurrence 50%). For this option GEN, the early diagnosis is still of 25% (threshold of the target requirement).

We consider that RDT meets the requirement. The versions associated with GEN discrimination scheme only are "operational" in EUMETSAT sense. Details and complete verification are in scientific report [RD.7] and previous verification report [RD.6].

3.3 THE VALIDATION OF CAL DISCRIMINATION SCHEME

3.3.1 Context

The purpose of this release is to provide updated calibrated discrimination for each geostationary satellite at its nominal scan rate. As mentioned in ATBD ([AD.11.]), long series of operational runs have been taken into account for that purpose: MSG-FDSS (MSG4 at that time) and MSG-IODC (MSG1 at that time) / 15 minutes, MSG-RSS (MSG3 at that time) / 5 minutes, GOES16 / 10 minutes. Himawari-8 being taken into account at Météo-France every 20 minutes only, and GOES17 every 30 minutes only, no calibrated discrimination has been attempted for those satellites. GOES16 models will be used instead.

In this report, only the nominal default configuration for the latest release is analysed, it means that calibrated discrimination is used. For GOES and HIMAWARI satellites, the channel #15 is taken as main channel. For GOES17 and HIMAWARI satellites, only subjective validation is presented hereafter. For MSG satellites and GOES16, both objective and subjective validations have been undertaken.

Lightning data remain the ground truth used to tune and to validate (on objective or subjective basis) RDT-CW. Those data are usually provided by a ground lightning network. Meteorage lightning network data allow to get reliable information over Europe, providing a high-quality ground truth. GLD360 network is operationally used by Météo-France on a limited region in southern part of Indian Ocean including Madagascar and La Reunion island, and allows also a reliable verification for RDT operated with MSG-IODC. GLM data are used for the RDT operated with ABI radiometer data from GOES16 and GOES17. For HIMAWARI, WWLLN global network is available. Compared to other networks, the quality of detection and precision of localization are however limited in that case. The RGB image used in some figures of RDT-CW validation is a Météo-France production based upon channels 0.6 µm (day), 3.7 µm (night), 10.8 µm and 12.0 µm.

3.3.2 Validation and Cases study

Those cases study rely on RDT-CW processing whose discrimination scheme is based on satellite data only, after NWP filtering (focusing on unstable areas only). Lightning data are paired with cloud cells on a passive "mode" (data not used for the convective diagnosis). Satellite image and lightning data will be hereafter visualized to synthesize ground truth.

3.3.2.1 RDT-CW discrimination using MSG4 0°

Runs for MSG4 case have been undertaken over Europe, for several dates out of the period used for the tuning. Those dates range from May to October 2018.

3.3.2.1.1 Objective validation

3.3.2.1.1.1 Data

Only datasets included in extended Meteorage and partners network coverage area are taken into account (see Figure 14).
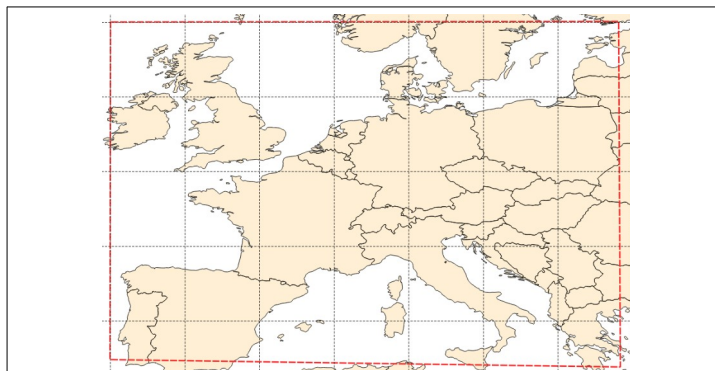
Figure 14: METEORAGE and partners network coverage area taken into account for objective validation

3.3.2.1.1.2 Methodology

RDT-CW was operated with lightning flashes as input data, but without changing the convective diagnosis. This diagnosis is evaluated against a "Ground Truth" based on electrical activity of the trajectory. Two "levels" of Ground Truth are considered: moderate and severe, regarding the electrical activity:

- Number of flashes per cloud cell of a given trajectory
- Total number of flashes during the trajectory
- Continuity of activity during the trajectory

For example moderate trajectories are assumed convective if they match with 5 flashes strokes at least. Moreover, non electric trajectories are defined if their minimum distance to nearest flash during lifetime is over 200km. This value is set to eliminate ambiguous cases and to take into account limited geographical extension of RDT-CW cell contours when focusing on a convective tower part of a larger convective cloud system.

Non electric trajectories helps to define correct rejection (CR). Hit (HI), False alarm (FA) and miss (MI) case are determined for the whole trajectory.

3.3.2.1.1.3 Results

Contingency tables and corresponding scores are presented below. They are determined independently for each full day.

Scores against moderate ground truth below highlight the ability of RDT-CW updated discrimination scheme to limit false alarm ratio in most cases, keeping correct values for POD. Scores are of course dependent on the situation, generally better for most actives situations. PODs range from 50% to 80%, and FAR from 4% to 33%. Average values are 67% for POD and 9% for FAR with a Threat Score (TS) of 63%.

Table 5: Contingency tables and scores for RDT-MSG4 runs over Europe, inside Meteorage coverage area. RDT discrimination diagnosis against Moderate Ground Truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	CR (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20180513	262	13	212	3153	55,3	4,7	0,4	0,6	53,8
20180526	368	38	173	4334	68,0	9,4	0,9	0,8	63,6
20180621	148	6	83	372	64,1	3,9	1,6	0,7	62,4
20180702	186	12	66	1493	73,8	6,1	0,8	0,8	70,5
20180703	232	18	80	2201	74,4	7,2	0,8	0,8	70,3
20180725	347	21	81	1054	81,1	5,7	2,0	0,9	77,3
20180807	364	28	144	3234	71,7	7,1	0,9	0,8	67,9
20180808	616	12	237	2782	72,2	1,9	0,4	0,7	71,2
20180809	480	15	168	2855	74,1	3,0	0,5	0,8	72,4
20181007	111	7	164	4328	40,4	5,9	0,2	0,4	39,4
20181008	163	15	165	3929	49,7	8,4	0,4	0,5	47,5
20190615	151	82	103	4822	59,4	35,2	1,7	0,9	44,9
20190619	482	27	214	3885	69,3	5,3	0,7	0,7	66,7
20190806	227	114	165	4264	57,9	33,4	2,6	0,9	44,9

Those results become of course even better when lightning data is considered also for the diagnosis, when necessary. In this configuration (100% for POD by construction), FAR value is automatically reduced (not shown here). In that case the gain is limited, to few situations : FAR drops from 35% to 26% on the 15th June 2019, from 33% to 22% o the 6th August 2019).

Scores against severe ground truth implies a reduced dataset, and show very logically a strong decrease of misses. One can note a light increase for PODs and FARs, sometimes significant like in the less favourable situations. Nevertheless, average values are improved with a POD of 72.6% , a FAR around 8.6%, and a Threat Score (TS) of 68%.

Table 6: Contingency tables and scores for RDT-MSG4 runs over Europe, inside Meteorage coverage area. RDT discrimination diagnosis against Severe Ground Truth

	HI (nb)	FA (nb)	MI (nb)	CR (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20180513	203	13	139	3153	59,36	6,02	0,41	0,63	57,18
20180526	282	38	112	4334	71,57	11,88	0,87	0,81	65,28
20180621	116	6	52	372	69,05	4,92	1,59	0,73	66,67
20180702	154	12	43	1493	78,17	7,23	0,80	0,84	73,68
20180703	190	18	42	2201	81,90	8,65	0,81	0,90	76,00
20180725	251	21	43	1054	85,37	7,72	1,95	0,93	79,68
20180807	304	28	91	3234	76,96	8,43	0,86	0,84	71,87
20180808	530	12	137	2782	79,46	2,21	0,43	0,81	78,06
20180809	402	15	100	2855	80,08	3,60	0,52	0,83	77,76
20181007	87	7	116	4328	42,86	7,45	0,16	0,46	41,43
20181008	107	15	111	3929	49,08	12,30	0,38	0,56	45,92
20190615	115	82	62	4822	64,97	41,62	1,67	1,11	44,40
20190619	363	27	123	3885	74,69	6,92	0,69	0,80	70,76
20190806	174	114	82	4264	67,97	39,58	2,60	1,13	47,03

		Code: NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
		Issue: 2.0.1 **Date:** 28th February 2021
NWC SAF METEO FRANCE	Validation report of the Convection Product Processors of the NWC/GEO	**File:** NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
		Page: 36/75

3.3.2.1.2   Case Study 20180621 over Europe

Following figures illustrate a convective situation in the Southern part of Europe. A cold front associated to a synoptic perturbation splits the region in two parts, convective systems developing in the warm air mass. Comparing the results for v2018 and v2021 releases, one can note that:

- Most electric phenomena are identified as convective clouds with RDT-CW, whatever the release.

- v2021 discrimination scheme does slightly lower the number of false alarms on this situation, when compared to v2018. This is illustrated at 15h00Z in Central Europe, but also in the Northern part of the perturbation.

- Some misses can be observed in embedded cloud systems especially in the cold front, but also with some isolated low active clouds.
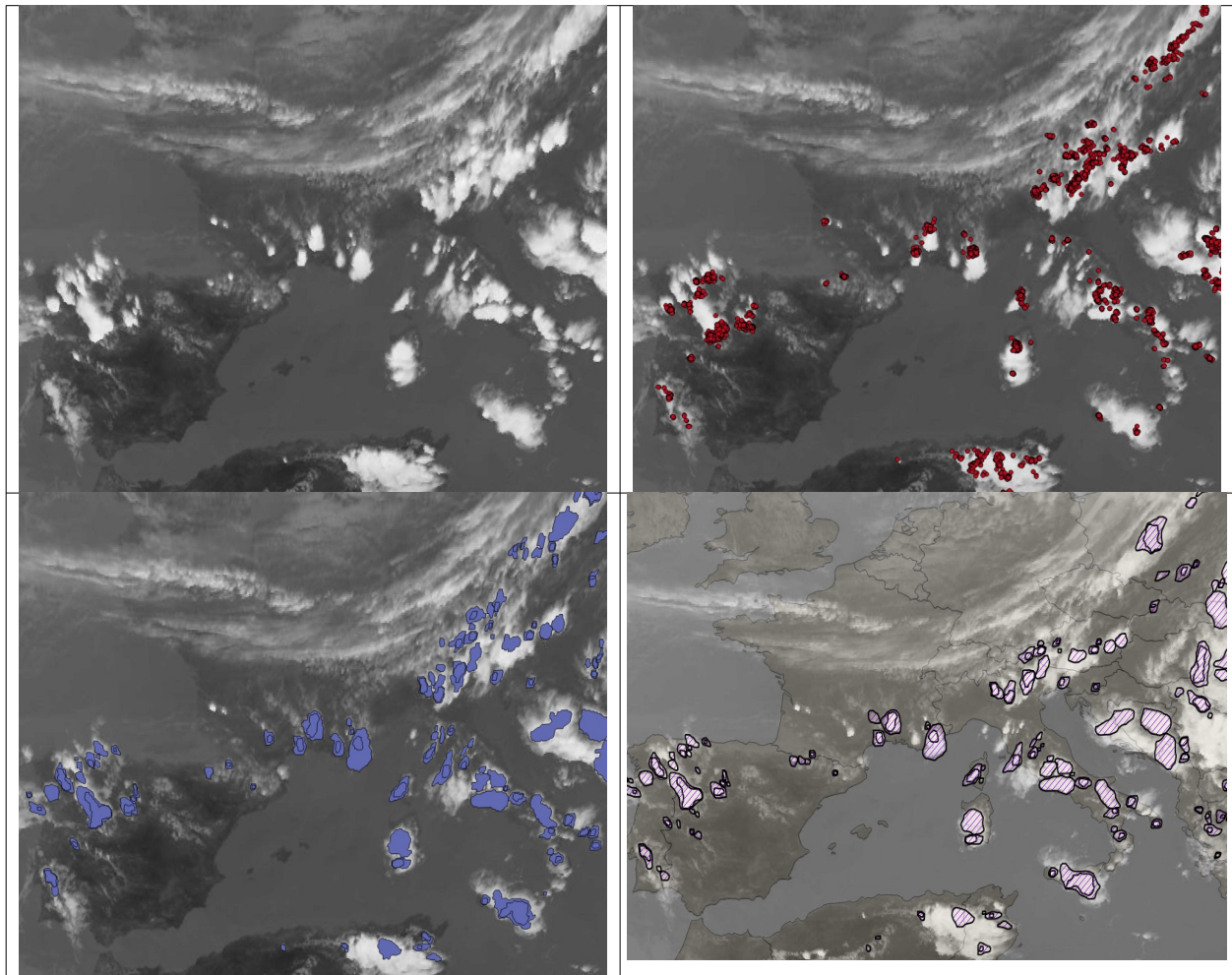
Figure 15: MSG4 case study for 12h00-15h00Z on 20180621. 15h00Z IR image (top left), 30min accumulated METEORAGE impacts around 15h00Z (top right), v2018 (bottom left) and v2021 (bottom right) results for 15h00Z.

Embedded cloud systems are difficult to distinguish, and lead either to numerous cloud cells among which some can be false alarms, either to limited cloud cells, at the risk of missing some apparent electric cells. This latest v2021 release is aimed to limit the number of false alarms, with the possibility to eventually use external lightning data to complement diagnosis.
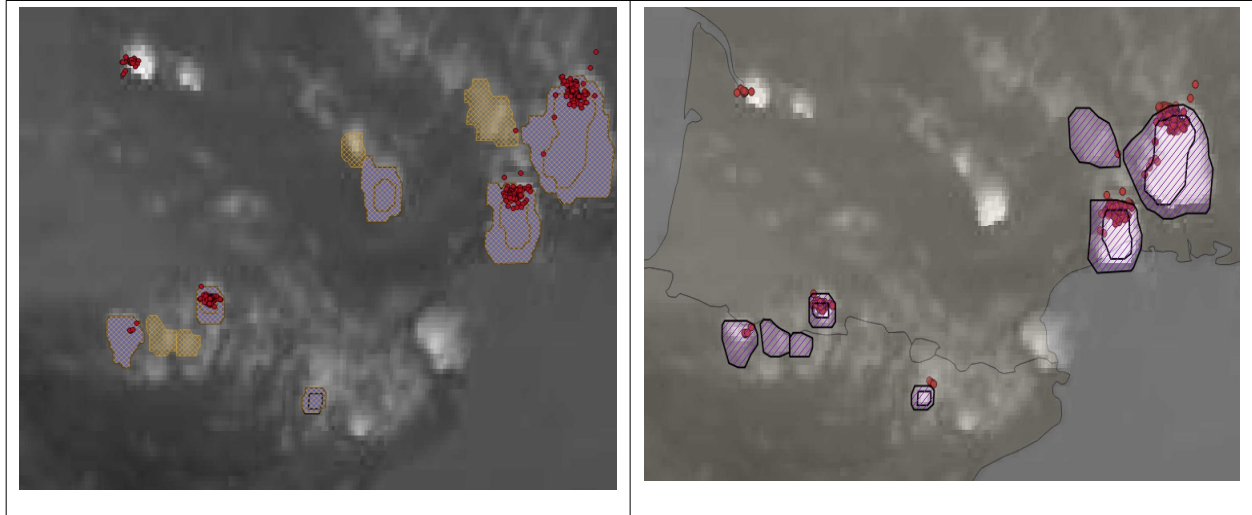
Figure 16: MSG4 case study for 20180621 15h00Z, zoom on South of France. v2018 (plain blue cells) and GEN (orange dashed) on the left, v2021 on the right. One can note a miss near Bordeaux (top left cloud) in any configuration, and a non-electric cloud system ignored by v2021.
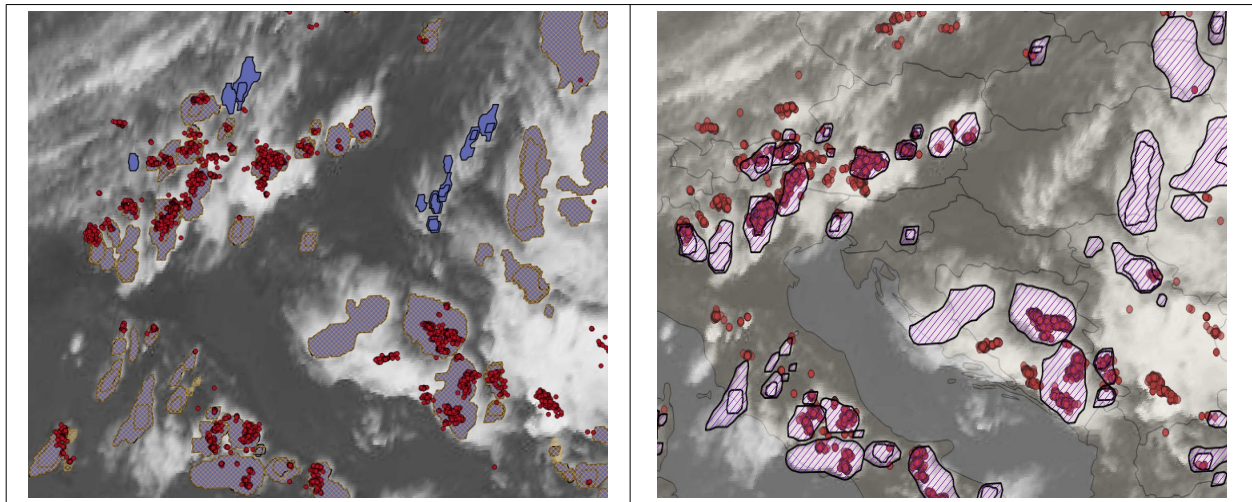
Figure 17: MSG4 case study for 20180621 15h00Z, zoom on Central Europe. v2018 (blue cells) and GEN (orange dashed) on the left, v2021 on the right. One can note with v2021 the suppression of False Alarms seen by v2018 on the edge of cloud systems
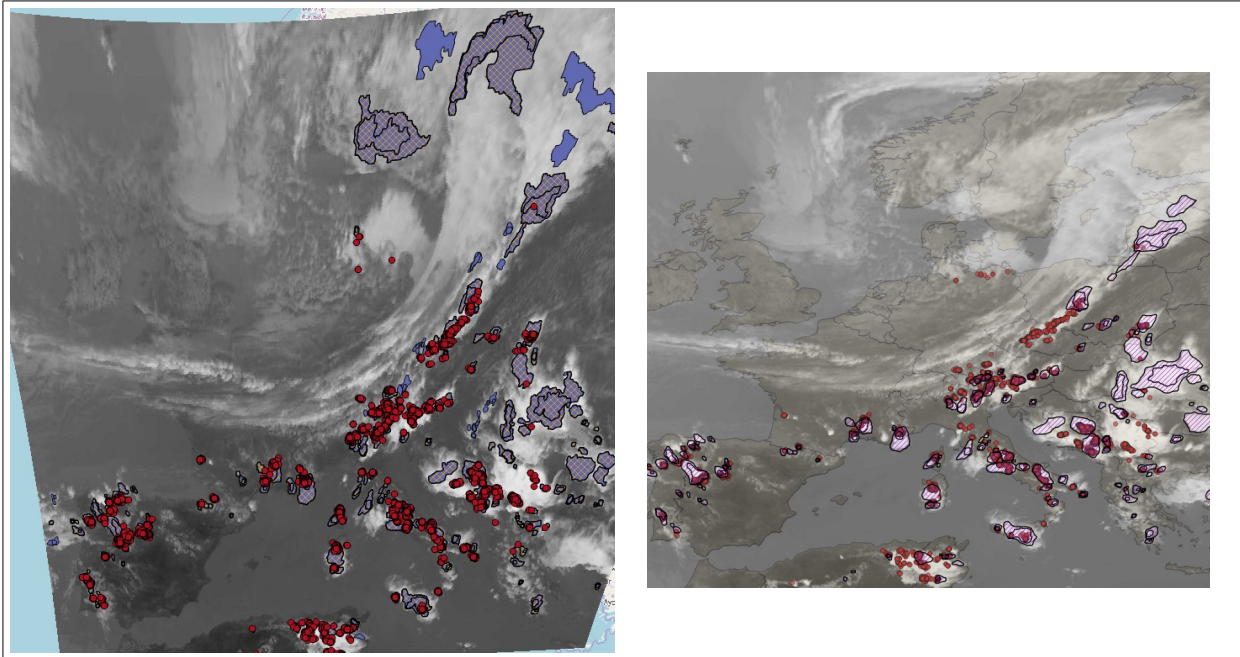
Figure 18: MSG4 case study for 20180621 15h00Z, large view. False Alarms suspicion in the North of the domain with v2018 (left), correctly rejected with v2021 release (right), in an area which is out of lightning coverage area.

3.3.2.1.3 Case study 20180702 over Europe

Figures hereafter illustrate a large and zoomed view of this situation.

Embedded and also isolated diurnal convective cells can be observed. The behaviour of RDT convective diagnosis appears here pretty efficient regarding electrical activity, especially if we have a look on the temporal evolution. Most of lightning strokes have been, are, or will be associated with a RDT-CW cloud cell diagnosed as convective.

In the middle part of France, RDT-CW cells appear sometimes prior to electrical activity. On the other hand, embedded cloud systems, as on the top left of the image, seem more difficult to separate and identify. Electrical activity is observed in several places, RDT-CW cells do not fully correspond at that time for this complex system. The focus of RDT-CW on the identification of cloud towers limits the representation of convective activity.
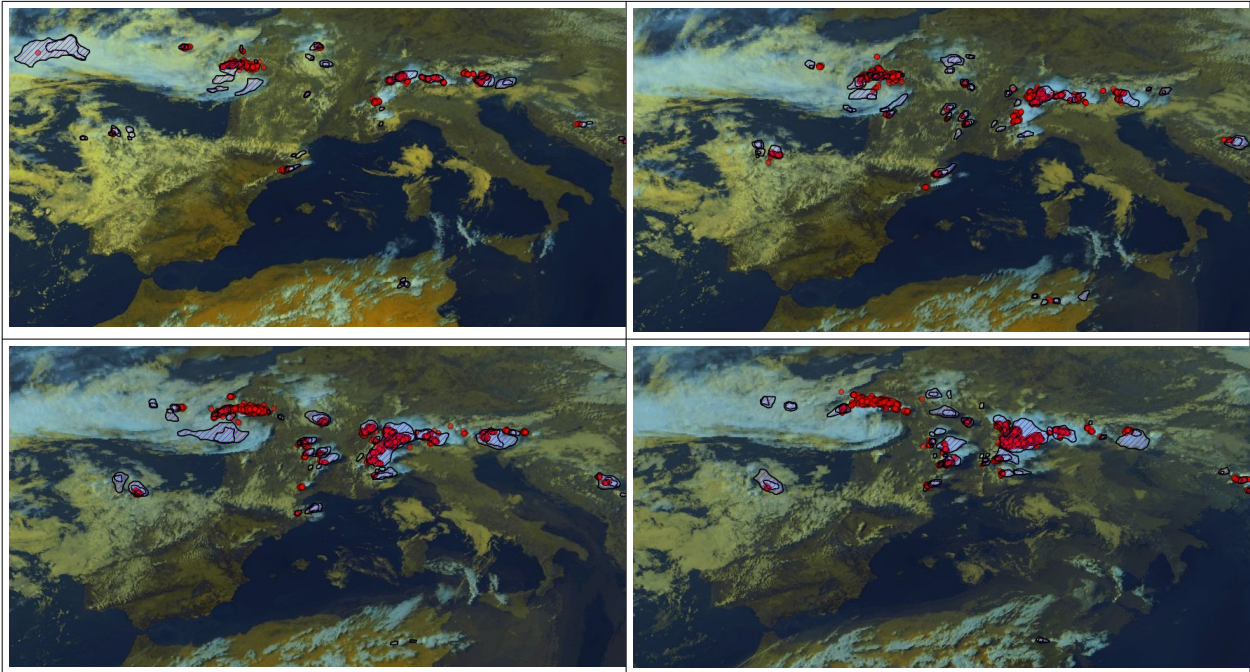
Figure 19: MSG4 case study for 20180702. 15h00Z, 16h00Z, 17h00Z and 18h00Z from top left to right bottom. RGB images with METEORAGE impacts, RDT-CW v2021 cells.

And if we consider RDT-CW from a temporal point of view, like in Figure 20, one can conclude that all electrical activity can be associated with a RDT-CW cloud systems, even if the flashes can be in the vicinity of the cells. Few false alarms are noted. Even an apparent miss at a given time (15h00Z on the French Alps) corresponds to a case where cells were diagnosed earlier (14h to 14h45Z) in the vicinity, and part of a larger system later on (17h00Z).
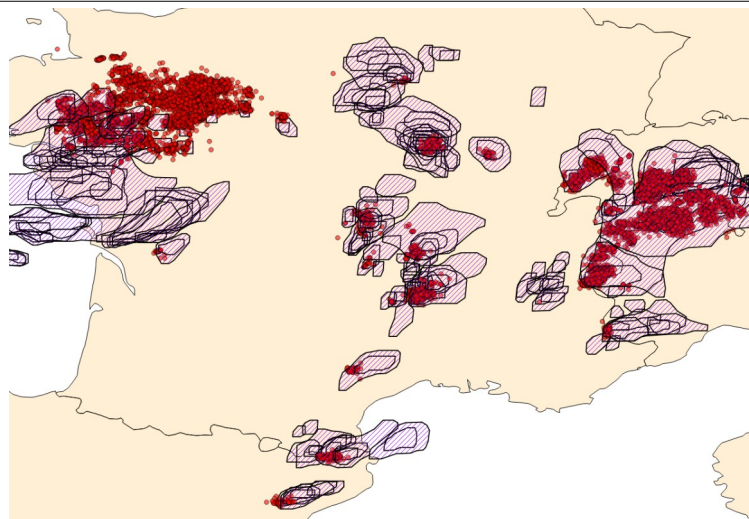
Figure 20: MSG4 case study for 20180702. Zoom over France, with accumulated Lightning and RDT data from 14h00Z to 18h00Z

3.3.2.1.4 Case study 20190419 over Africa

When regarding RDT-CW behaviour over Africa, one can note that lots of convective systems, even the smallest ones, are linked to strong signatures of predictors, and make the statistical models adapted and relevant.

For this region, WWLLN data are used to check a RDT-CW convective diagnosis based on satellite characteristics only. Following figures illustrate RDT-CW performances for a 20190419 situation in the end of afternoon.

A global overview show a good agreement between electrical activity and identification of convective cells by RDT-CW. both diurnal intensification of convection over land, and activity of ITCZ over ocean are highlighted, even if electrical activity over land is obviously stronger.

One can note rare misses (mainly around 40°S in a cold air mass behind synoptic perturbation, red circle). Few false alarms are suspected (blue circles). Most of them occur over sea surface. A deeper analysis looking at previous and following periods reveal that for some of those suspected false alarms, RDT-CW cells are located in areas where electrical activity was, or will be present.
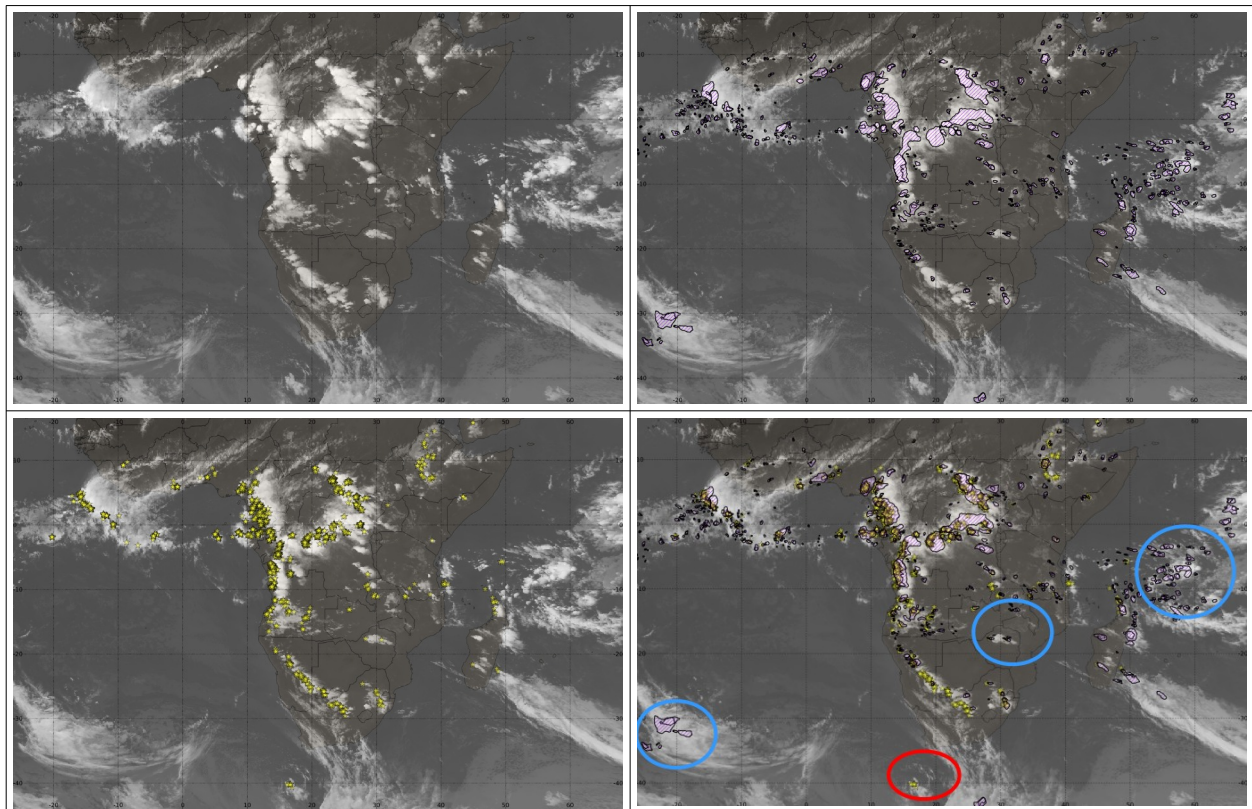
Figure 21: MSG4 case study for 20190419 17h00Z over Africa. IR image (top left) with RDT-CW black-dashed cells (top right), with WWLLN data as yellow stars (bottom left), all data overlaid (bottom right).

A detailed zoom over inland (Democratic Republic of Congo, Tanzania, Zambia), taking into account previous (H-2h) and following (H+2h) electrical dataset is illustrated in following figure. RDT-CW maintains the identification of cloud systems which have been electrically active (yellow circles), and identifies cloud system which will become electrically active (magenta circle). It highlights and confirms the fact that RDT-CW cells are located in active regions, even if not always synchronous with electrical activity.
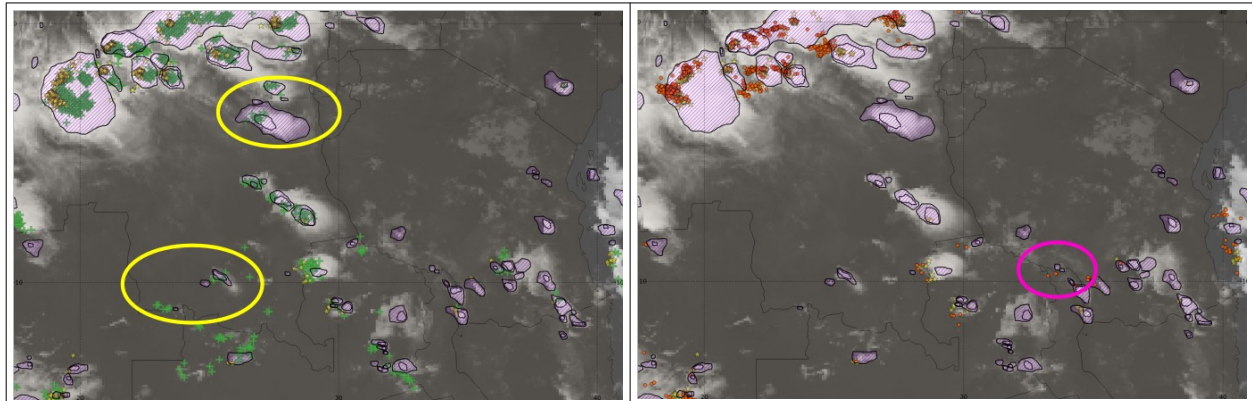
Figure 22: Zoom of MSG4 case study for 20190419 zoomed. 17h00Z IR image overlaid with synchronous RDT-CW black-dashed contours and electric data (yellow stars). Left: adding previous 2h WWLLN data (green crosses). Right: adding following 2h of WWLLN data (orange dots).

This coherence regarding a temporal tolerance is also found in Indian Ocean north-east of Madagascar, where RDT-CW points lots of convective cells. Finally we can consider the behaviour of RDT-CW relevant on those tropical / equatorial regions during a warm season.

3.3.2.2 RDT-CW discrimination using MSG3 - 9.5E° RapidScan mode

RDT runs with MSG3 have been undertaken over Europe, for 6 dates out of the period used for the tuning, common with but less numerous than for MSG4. Those dates range from May to August 2018.

3.3.2.2.1 Objective validation

The same methodology has been applied than for RDT-MSG4, for quantifying the results against Meteorage&Partners ground truth.

Contingency tables and corresponding scores are presented below. They are determined independently for each full day.

Two situations among six reveal higher false alarms with RDT-MSG3-RSS compared to RDT-MSG4. PODs are sometimes much better than RDT-MSG4, sometimes slightly lower or equivalent. Scores are of course dependent on the situation, generally better for most actives situations. PODs range from 60% to 68%, and FAR from 6% to 43%.

Even if less significant because based on lower dataset, average values are 65% for POD and 20% for FAR with a TS of 56%.

Table 7: Contingency tables and scores for RDT-MSG3-RSS runs over Europe, inside Meteorage coverage area, considering moderate ground truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	GND (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20180526	351	53	214	7885	62,12	13,12	0,67	0,72	56,80
20180702	190	12	104	2533	64,63	5,94	0,47	0,69	62,09
20180703	293	28	136	2828	68,30	8,72	0,98	0,75	64,11
20190615	218	164	123	8871	63,93	42,93	1,82	1,12	43,17
20190619	590	59	277	5829	68,05	9,09	1,00	0,75	63,71
20190806	266	178	168	6726	61,29	40,09	2,58	1,02	43,46

Regarding severe ground truth, one can note again an increase of PODs and FARs, more significant for the worse situations. Average POD is higher with 71%, but FAR lower with 25% and Threat Score (TS) remains equivalent.

Table 8: Contingency tables and scores for RDT-MSG3-RSS runs over Europe, inside Meteorage coverage area, considering severe ground truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	GND (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20180526	266	53	126	7885	67,86	16,61	0,67	0,81	59,78
20180702	155	12	60	2533	72,09	7,19	0,47	0,78	68,28
20180703	229	28	81	2828	73,87	10,89	0,98	0,83	67,75
20190615	162	164	74	8871	68,64	50,31	1,82	1,38	40,50
20190619	447	59	153	5829	74,50	11,66	1,00	0,84	67,83
20190806	202	178	93	6726	68,47	46,84	2,58	1,29	42,71

The higher update rate of MSG-RSS implies a higher number of diagnosis attempted, which could bring an explanation to the higher number of false alarms compared to RDT-MSG. But we also have to consider that a higher number of splits and merges can be diagnosed. Situations with embedded convection in particular do probably not benefit from a higher update rate.

3.3.2.2.2 Case study 20180703 over Europe

With this situation, we focus on the problematic of false alarms which could be observed in areas without any electrical activity, more frequently with the tuning for RapidScan mode.
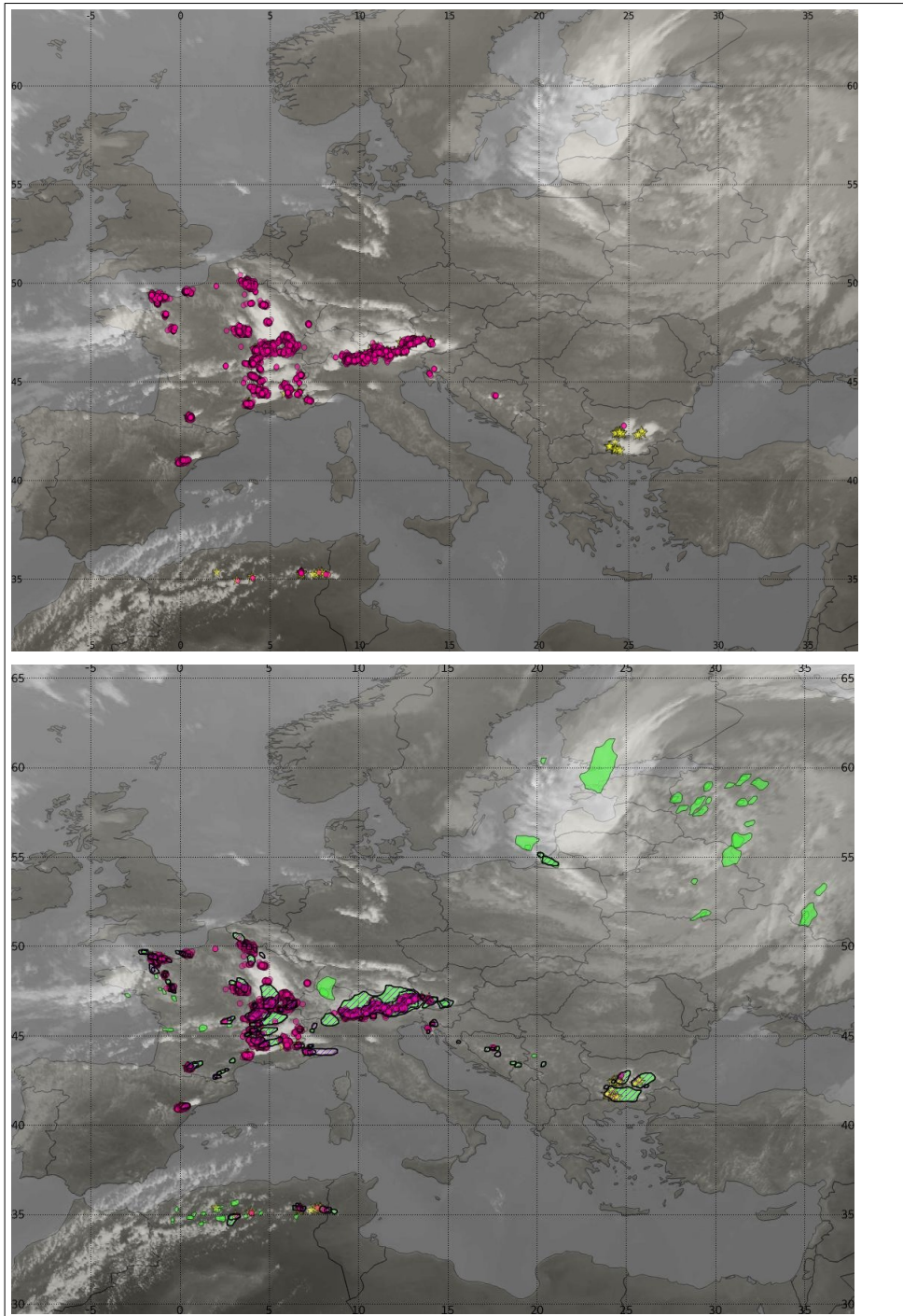
Figure 23: MSG3-RSS case study for 20180703 14h15Z. Top: IR image overlaid with METEORAGE (magenta circles) and WWLLN strokes (yellow stars). Bottom: same with MSG3-RSS RDT-CW v2018 cells (green) and MSG3-RSS RDT-CW v2021 cells (magenta dashed)

Despite a relative good agreement between v2018 and v2021 results over the central convective zone, v2021 tuning obviously improves this point, regarding tree different regions: North-east of Europe where almost all v2018 cells disappear with v2021 tuning, North Algeria where 2021 tuning seems more focused on active clouds, and in the central and west of France, where no convective cells appear in non electric areas with v2021 tuning.

As for other cases, RDT-MSG3-RSS with updated tuning shows improved performances.

3.3.2.3 RDT-CW discrimination using MSG1 - 41.5°E

RDT runs with MSG1-IODC have been undertaken for February 2019 over a limited domain, over which GLD360 data are operationally available at Meteo-France (see next figure). The period is representative of the warm and convective season, and has been extended to catch enough convective systems over the oceanic part of the domain.

3.3.2.3.1 Objective validation

The same methodology has been applied than for RDT-MSG4, for quantifying the results against a ground truth. But in that case GLD360 data have been taken into account in the area illustrated in Figure 24. This ground lightning network is officially considered to be reliable in this region. Full domain and a reduced domain limited mainly to land surfaces have been considered.
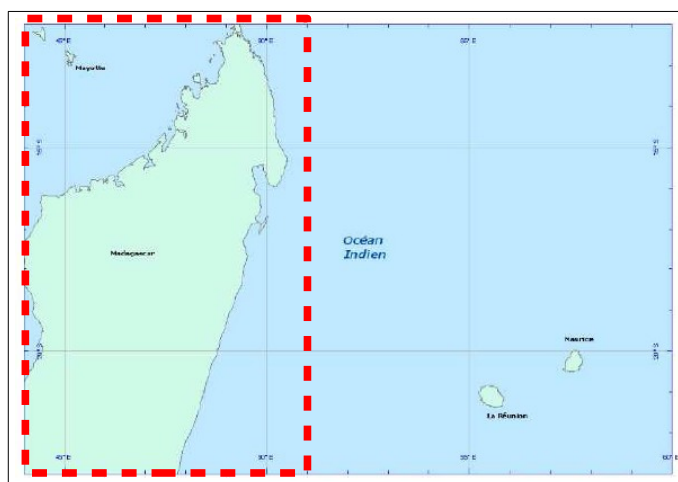
Figure 24: Coverage area of GLD360 for Météo-France, and reduced domain (dashed red) to focus on land area

Contingency tables and corresponding scores are presented below. They are determined independently for each full day.

PODs are very high, between 65 and 90%, with low variability (mean value 75%). FAR on the contrary vary between 5 and 78%, and have an average value of 45%. It is to note that, regarding the season, the number of convective systems is not so high. But the length of the period makes the results significant.

Table 9: Contingency tables and scores for RDT-MSG1-OI runs inside GLD360 MF coverage area, considering moderate ground truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	GND (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20190131	168	123	41	2546	80,38	42,27	4,61	1,39	50,60
20190201	133	288	30	3485	81,60	68,41	7,63	2,58	29,49
20190202	115	329	31	3063	78,77	74,10	9,70	3,04	24,21
20190203	85	284	34	3395	71,43	76,96	7,72	3,10	21,09
20190204	127	170	43	3295	74,71	57,24	4,91	1,75	37,35
20190205	123	43	43	2423	74,10	25,90	1,74	1	58,85
20190206	229	100	53	1990	81,21	30,40	4,78	1,17	59,95
20190207	187	125	75	2047	71,37	40,06	5,76	1,19	48,32
20190208	295	44	96	2058	75,45	12,98	2,09	0,87	67,82
20190209	206	68	83	2891	71,28	24,82	2,30	0,95	57,70
20190210	156	32	59	1003	72,56	17,02	3,09	0,87	63,16
20190211	127	48	32	850	79,87	27,43	5,35	1,10	61,35
20190212	102	76	30	741	77,27	42,70	9,30	1,35	49,04
20190213	90	19	40	272	69,23	17,43	6,53	0,84	60,40
20190214	157	9	68	1811	69,78	5,42	0,49	0,74	67,09
20190215	196	11	108	4538	64,47	5,31	0,24	0,68	62,22
20190216	244	41	84	3240	74,39	14,39	1,25	0,87	66,12
20190217	173	53	73	1411	70,33	23,45	3,62	0,92	57,86
20190218	151	64	56	1994	72,95	29,77	3,11	1,04	55,72
20190219	103	126	53	3013	66,03	55,02	4,01	1,47	36,52
20190220	124	105	61	4187	67,03	45,85	2,45	1,24	42,76
20190221	115	155	21	1870	84,56	57,41	7,65	1,99	39,52
20190222	229	128	46	1859	83,27	35,85	6,44	1,30	56,82
20190223	213	160	47	3129	81,92	42,90	4,86	1,43	50,71
20190224	147	252	22	4121	86,98	63,16	5,76	2,36	34,92
20190225	142	466	17	3722	89,31	76,64	11,13	3,82	22,72
20190226	135	490	11	3514	92,47	78,40	12,24	4,28	21,23
20190227	142	514	17	4148	89,31	78,35	11,03	4,13	21,10
20190228	98	137	12	2614	89,09	58,30	4,98	2,14	39,68

Regarding the number of false alarms, higher than expected, the validation has been estimated over a smaller domain mainly covered by land, over Madagascar (area in Figure 24, limited to 51°E). The comparison will help to estimate the impact of oceanic surfaces on RDT-CW discrimination performances.

Results are presented in the table below. FARs are considerably reduced in that case (average value 21%), with same values of POD. It suggests that a large part of false alarms are located over ocean. Results appear much better over land. Case study below will confirm the fact that with a lower electrical activity, convective systems over ocean are difficult to assess. The tuning in particular is highly influenced by the electrical activity over land. And an unknown part of "false alarms" could probably be cancelled because obviously convective without electrical activity.

Table 10: contingency tables and scores for RDT-MSG1-OI runs inside "reduced" GLD360 MF "land" area, considering moderate ground truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	GND (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20190131	135	17	29	823	82,32	11,18	2,02	0,93	74,59
20190201	105	37	29	1117	78,36	26,06	3,21	1,06	61,40
20190202	91	49	26	1309	77,78	35	3,61	1,20	54,82
20190203	64	75	31	1144	67,37	53,96	6,15	1,46	37,65
20190204	102	75	33	1269	75,56	42,37	5,58	1,31	48,57
20190205	97	22	29	964	76,98	18,49	2,23	0,94	65,54
20190206	201	22	34	782	85,53	9,87	2,74	0,95	78,21
20190207	155	15	61	570	71,76	8,82	2,56	0,79	67,10
20190208	226	20	71	405	76,09	8,13	4,71	0,83	71,29
20190209	162	19	73	713	68,94	10,50	2,60	0,77	63,78
20190210	156	28	59	655	72,56	15,22	4,10	0,86	64,20
20190211	127	42	32	731	79,87	24,85	5,43	1,06	63,18
20190212	102	67	30	541	77,27	39,65	11,02	1,28	51,26
20190213	90	19	40	259	69,23	17,43	6,83	0,84	60,40
20190214	157	7	68	934	69,78	4,27	0,74	0,73	67,67
20190215	196	10	108	1734	64,47	4,85	0,57	0,68	62,42
20190216	215	9	78	940	73,38	4,02	0,95	0,76	71,19
20190217	95	20	46	799	67,38	17,39	2,44	0,82	59,01
20190218	92	18	47	980	66,19	16,36	1,80	0,79	58,60
20190219	71	49	38	1642	65,14	40,83	2,90	1,10	44,94
20190220	82	37	44	1802	65,08	31,09	2,01	0,94	50,31
20190221	105	14	19	646	84,68	11,76	2,12	0,96	76,09
20190222	205	21	37	488	84,71	9,29	4,13	0,93	77,95
20190223	199	68	36	1287	84,68	25,47	5,02	1,14	65,68
20190224	138	75	20	1657	87,34	35,21	4,33	1,35	59,23
20190225	129	167	17	1395	88,36	56,42	10,69	2,03	41,21
20190226	127	135	11	1405	92,03	51,53	8,77	1,90	46,52
20190227	136	124	15	1756	90,07	47,69	6,60	1,72	49,45
20190228	89	74	10	1923	89,90	45,40	3,71	1,65	51,45

Regarding a severe ground truth will lead to the same conclusions: higher POD, slightly higher FAR, with a more positive impact over land (high gain for POD, low loss for FAR).

As a consequence of RDT-MSG1-OI objective validation, one can consider that the tuning phase of RDT-CW discrimination should take into account the nature of the surface (land or sea), with a kind of tolerance still to determine, in order to establish two different statistical approaches. This should also be the case for real-time RDT-CW processing. But the main difficulty will remain the adaptation of the ground truth to take into account over ocean.

3.3.2.3.2   20190221 over South-west part of Indian Ocean

This case study illustrated below in Figure 25 highlights two major features:

1) On one hand RDT-CW seems to be in good agreement with convective clouds as they appear in MSG1-OI IR image

2) On the other hand, this agreement is supported by lightning data over land than over ocean. Out of the GLD360 coverage area (grey shaded), WWLLN lightning data shows weak activity over ocean at latitudes closer to equator.

It confirms how difficult it is to consider a ground truth, depending on sensors (coverage area, detection efficiency), but also on the region concerned: land or ocean surface, equatorial / tropical / mid latitudes.

Nevertheless this RDT-CW discrimination scheme is considered to be satisfactory in most cases
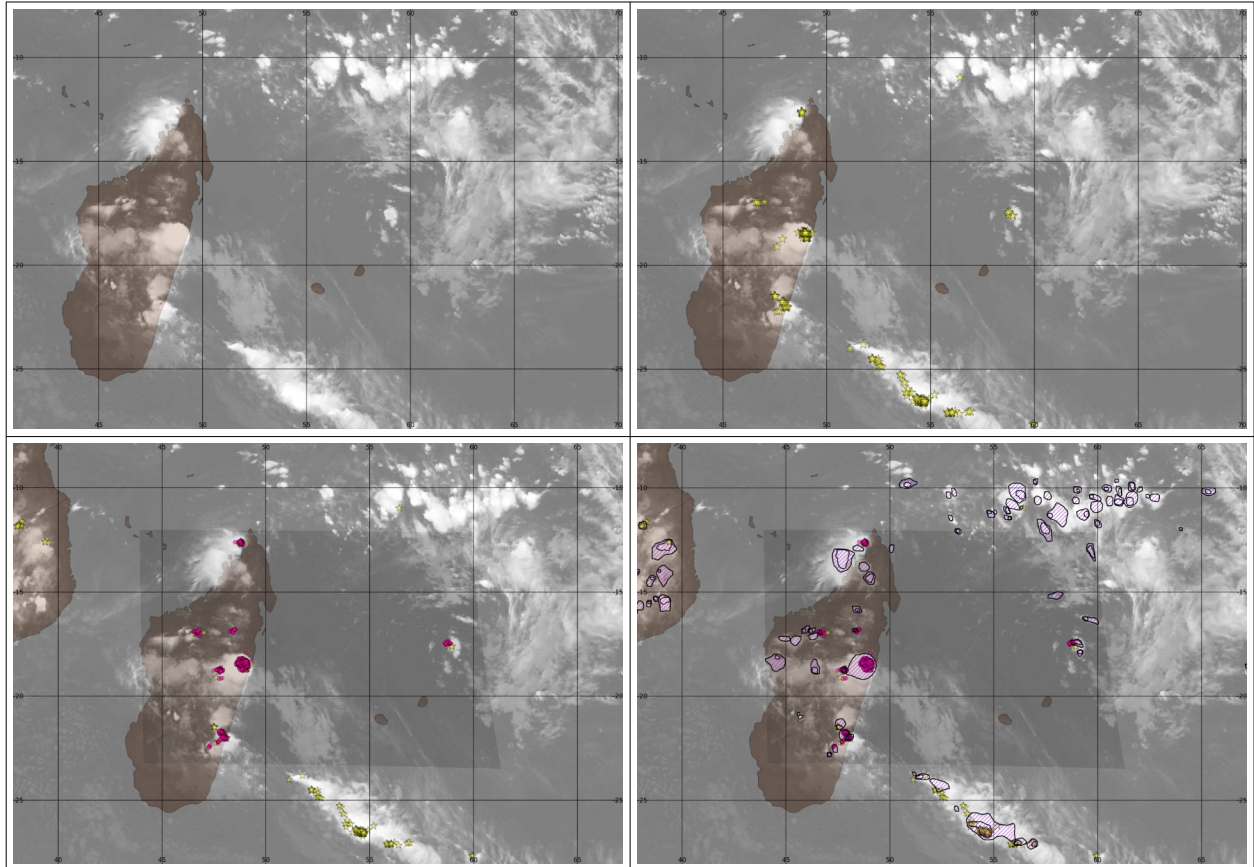
Figure 25: MSG1-41.5E case study for 20190221, 18h00Z slot. MSG1-IR image (top left), overlaid with WWLLN data (top right), with GLD360 data (bottom left, coverage area grey shaded), and with RDT-CW cell contours (black dashed contours)

3.3.2.4 RDT-CW applied to GOES16

NOAA has undertaken until summer 2020 a high improvement of GLM-GOES16 data, for identifying and qualifying regular and punctual artefacts of lightning data. For that reason, as mentioned in ATBD, those data have been used for a new tuning of RDT-CW discrimination scheme. Those data are also used for validation on cases study.

Following nominal configuration is chosen: 10 minutes update rate and IR10.3µm as main channel, to process RDT-CW with GOES16 ABI data.

3.3.2.4.1 Objective validation

Runs have been undertaken for about 15 days during September 2020, in a large GOES16 sub domain including Tropical and Equatorial regions (Caribbean islands, Central America and North of South-America), continental and oceanic surfaces, as illustrated in Figure below.
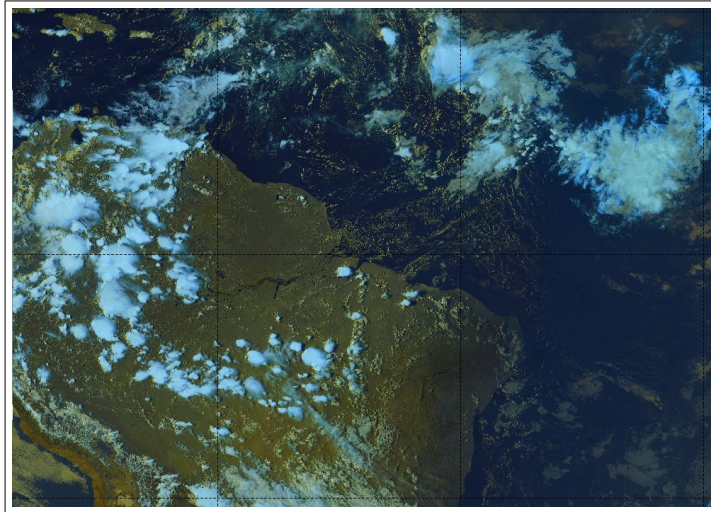
Figure 26: Sub-domain taken into account for replays and objective validation

The same methodology has been applied than for RDT-MSG4, for quantifying the results against a ground truth. But in that case GLM flash data have been taken into account. They are paired with cloud cells within a period centred on the given slot.

Hit (HI), False alarm (FA) and miss (MI) case are determined for the whole trajectory.

Contingency tables and corresponding scores are presented below. They are determined independently for each full day.

Scores against moderate ground truth below highlight the ability of RDT-CW discrimination scheme to reach high PODs. But one can note a relative high level of false alarm ratio, with a high variability. Scores are of course dependent on the situation, generally better for most actives situations. PODs range from 60% to 80%, and FAR from 16% to 50%.

Average values are 73% for POD and 38% for FAR with a Threat Score (TS) of 50%. With a modified sample of ground truth focusing on the most severe activity, average POD can rise to 80%.

Table 11: Contingency tables and scores for RDT-GOES16 runs over chosen sub-domain, inside GLM coverage area. RDT discrimination diagnosis against Moderate Ground Truth

runs/scores	HI (nb)	FA (nb)	MI (nb)	GND (nb)	POD (%)	FAR (%)	POFD (%)	B (%)	TS
20200901	1642	1550	427	17883	79,36	48,56	7,98	1,54	45,37
20200902	1493	1190	342	12113	81,36	44,35	8,95	1,46	49,36
20200904	1168	908	326	9826	78,18	43,74	8,46	1,39	48,63
20200906	915	1180	338	12922	73,02	56,32	8,37	1,67	37,61
20200908	1415	1530	569	14362	71,32	51,95	9,63	1,48	40,27
20200912	1551	1179	308	10990	83,43	43,19	9,69	1,47	51,05
20200914	1343	857	527	15013	71,82	38,95	5,4	1,18	49,25
20200916	966	663	434	18223	69	40,7	3,51	1,16	46,83
20200919	1632	834	574	14654	73,98	33,82	5,38	1,12	53,68
20200921	1655	607	1135	14370	59,32	26,83	4,05	0,81	48,72
20200923	1590	675	727	15965	68,62	29,8	4,06	0,98	53,14
20200926	1739	522	741	13841	70,12	23,09	3,63	0,91	57,93
20200928	2963	592	968	12946	75,38	16,65	4,37	0,9	65,51

Despite some situations with relative high false alarm ratio, those results can be generally considered as good, because catching a large amount of convective systems. Moreover, as for MSG-IODC, we face the difficulty to validate a convective characteristic with an electrical ground truth over sea

surface that may lead to an apparent high level of false alarm even if some suspected features are detected on satellite images.

The possibility with GOES16 to benefit in real time from GLM data with channels from ABI imager, ensures that, with a specific active pairing (GLM data used for convection diagnosis), all convective and interesting cloud systems will be identified.

3.3.2.4.2   Case study 20200919

This situation illustrates the performances of RDT-CW over continental area, and some limitations over sea surfaces.

The overall preview in figure below in the end of afternoon highlights a very good agreement between RDT-CW cells, GLM flashes, and apparent convective clouds in RGB image. Though, oceanic surfaces are associated with RDT-CW detection but very few flashes. And one can also note some obvious misses over land.
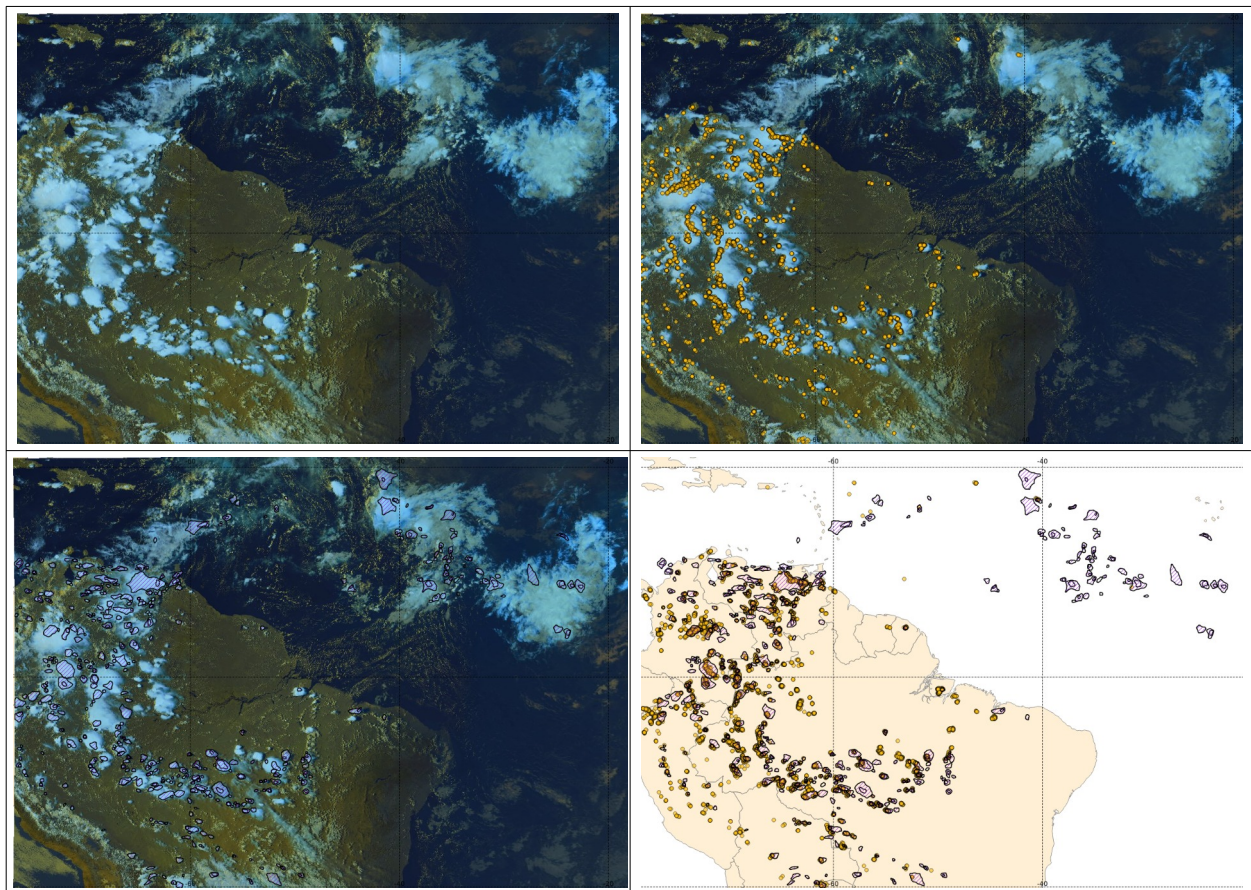
Figure 27: GOES16 case study for 20200919 at 20h00Z. RGB image (top left) overlaid with synchronous GLM (top right, orange dots), with RDT-CW (bottom left, dark shaded contours), and RDT-CW overlaid with GLM (bottom right).

Following steps focus on misses and supposed false alarms, over land and sea. With a zoom on a land area displayed in Figure 28, it appears that, even over land, it is not so easy to qualify with a high level of confidence some misses or false alarms. Some misses are obvious , but we know that the use of real-time GLM data can now compensate this weakness. Other misses are doubtful, because never caught by RDT neither GLM, but getting almost same appearance than neighbouring

confirmed convective clouds. On the other hand, RDT-CW identifies clouds which never become electric, but are obviously convective and close to electrical activity.

red = miss
dashed red = good non detection or double miss ?
blue = supposed false alarms
green = late hit (RDT after GLM)
dashed green = early hit (RDT before GLM)
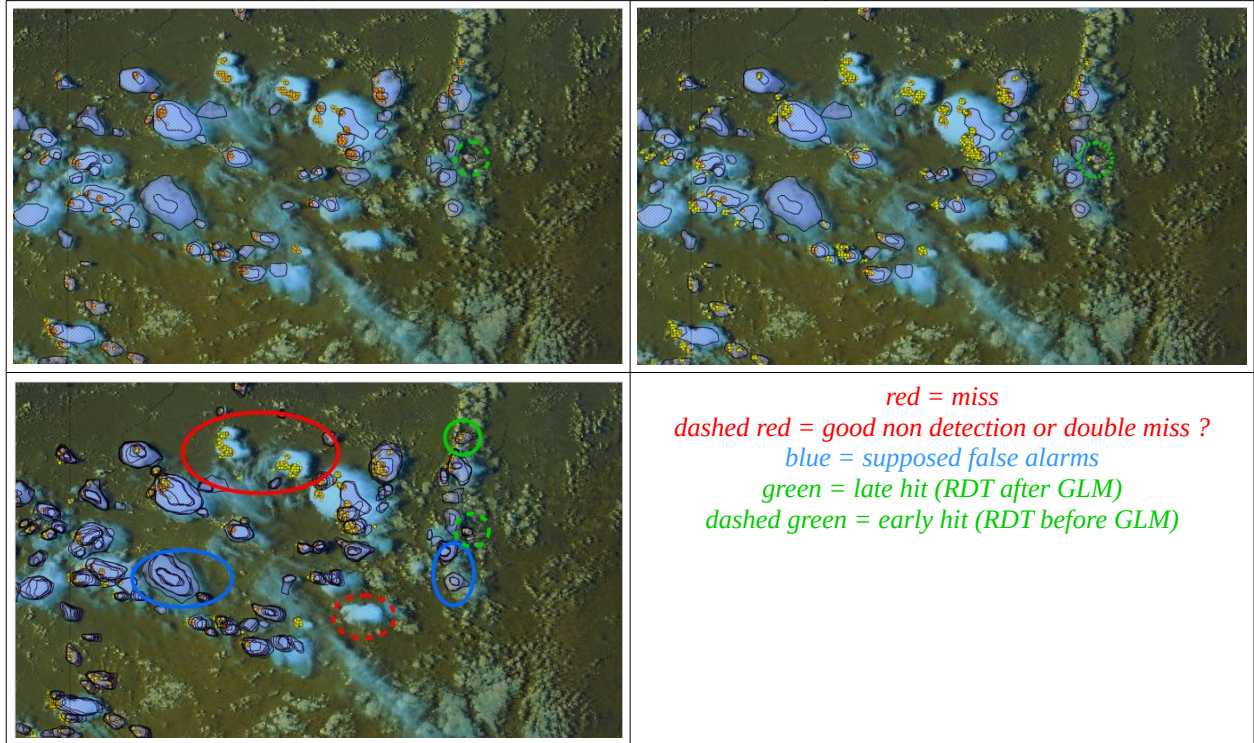
Figure 28: GOES16 case study for 20200919 at 20h00Z, over land. RGB image overlaid with synchronous GLM (top left , orange dots), with RDT-CW + next following 30min GLM data (top right, yellow dots), then with cumulated RDT-CW+GLM between 19h50 and 20h30Z (bottom left).

Over sea surface, checking RDT-CW performances with electrical activity is difficult: weak electric signature of diurnal convection over ocean, high distance of the nearest lightning detection point for some ground-based network. With a zoom on a zone concerned by ITCZ, few electrical activity appears and RDT-CW identifies significant cloud systems. Despite some false alarms in complex embedded cloud systems, RDT-CW seems here the best tool to focus on significant phenomena, compared to lightning data only or RGB only.
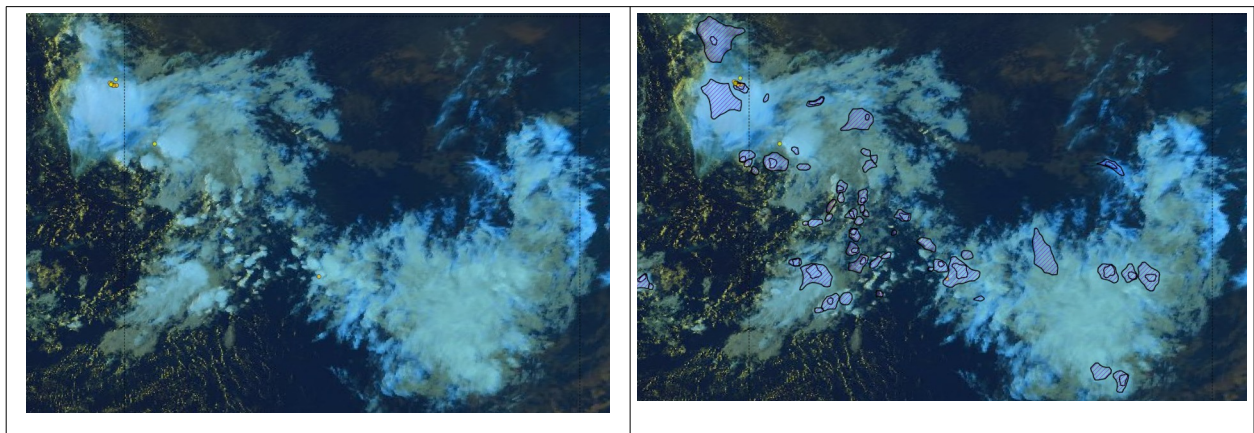
Figure 29: GOES16 case study for 20200919 at 20h00Z, over sea. RGB image overlaid with 40min cumulated GLM (left, orange/yellow dots), with RDT-CW (right).

3.3.2.4.3 Case study 20201123, Caribbean Sea

Here is presented a case study on a slightly different domain over sea surfaces, with moderate convective activity, illustrated in Figure 30 for a beginning of afternoon. Ten-minute GOES16 GLM flashes centred on 20h00Z slot are of limited extent. RGB image like RDT-CW cell contours highlight some other cloud systems obviously of interest. Very few flashes are orphans from RDT-CW cells, and RDT-CW subjectively seems to present good performances.

Supposed false alarms (blue circles) are identified in Figure 30. Moreover, if we have a look on the following electrical activity in Figure Erreur : source de la référence non trouvée, misses (red circles) and early good detections confirmed by GLM (green circles) are highlighted.
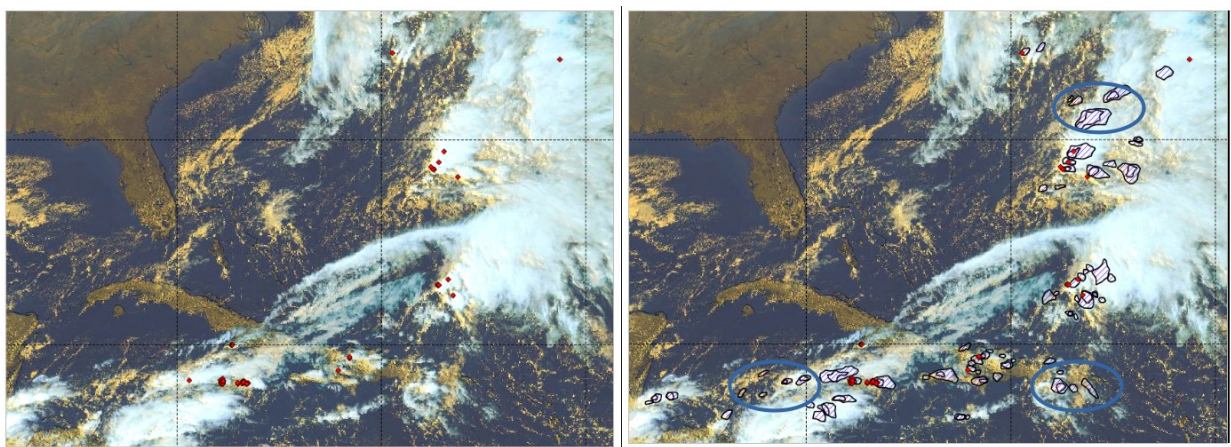
Figure 30: GOES16 case study for 20201123 at 20h00Z. RGB image superimposed with synchronous GLM (left, red diamonds), and with RDT-CW (right, dark shaded contours). Blue circles are supposed false alarms
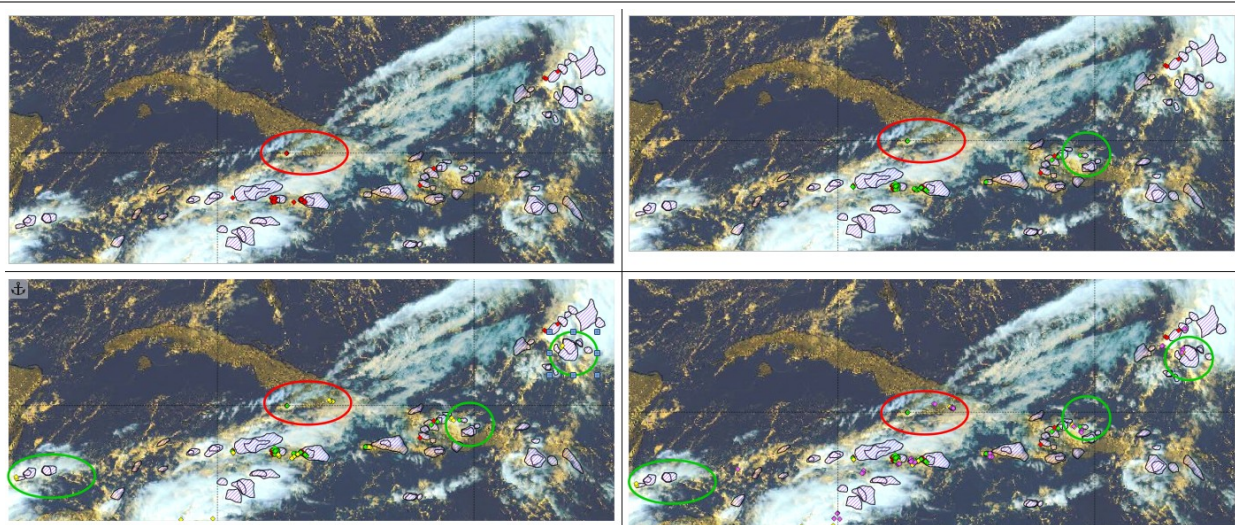
Figure 31: GOES16 case study for 20201123 at 20h00Z. RGB and RDT-CW , superimposed with synchronous GLM flashes (red diamonds), following 30 minutes GLM (green diamonds), following 60 minutes (yellow diamond) and 90 minutes (pink diamonds). Misses (red circles) and confirmed early detections (green circles) highlighted.

	Validation report of the Convection Product Processors of the NWC/GEO	**Code:** NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
NWC SAF METEO FRANCE		**Issue:** 2.0.1 **Date:** 28th February 2021
		File: NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
		Page: 52/75

For this situation, on this sub-domain, RDT-CW can be considered in agreement with GLM data, keeping in mind the necessity to have a look on the temporal evolution of the activity of ground truth.

Moreover, supposed false alarms (mainly above sea surface) seem associated with cloud cells far from GLM data, but with some texture or feature specific of convective clouds.

3.3.2.5 RDT-CW discrimination using Himawari-8

Validation process for RDT-CW applied to HIMAWARI has been done subjectively, and based on limited number of cases studies. It must be recalled here that Himawari data are made available in Meteo-France for NWC SAF processing at a sub-nominal 20 minutes update rate (instead of 10 minutes). For that reason, specific tuning has not been attempted. This lead us to consider GOES16 tuning, with the necessity to switch to the use of IR10.3 as main channel (instead of 11.2 previously). Lightning data from WWLLN data will be considered as an indication of convective activity, but hardly as a fully reliable ground truth.

3.3.2.5.1   Case study 20210326 over Micronesia region

We focus here on an almost full oceanic tropical domain. Next figure illustrates a mid-day situation, with a numerous apparent convective activity. Nevertheless, this activity is poorly confirmed by the electrical data from WWLLN networks.

The comparison between v2018 and v2021 RDT-CW results show how much both the adaptation of the tuning originally developed for GOES16 and the use of IR10.3 as main channel modify the identification of convective systems. The previous release was known to over-discriminate cloud systems as convective, producing a suspected large set of false alarms. Version v2021 appears as a mitigation of this point.

This release keeps an identification of convective systems which are confirmed by electrical activity (some few misses north of Solomon islands, but already missed with v2018 results), diagnoses most obvious convective clouds and MCSs, and seems to miss only some small growing convective clouds.

For this latter case, it is likely that the 20 minutes update rate at our disposal does not allow to take full benefit from a tuning made with 10 minutes data. However, we consider this release more efficient, keeping in mind the possibility to use lightning data as a possibility to change the the convection diagnosis from "No" to "Yes".
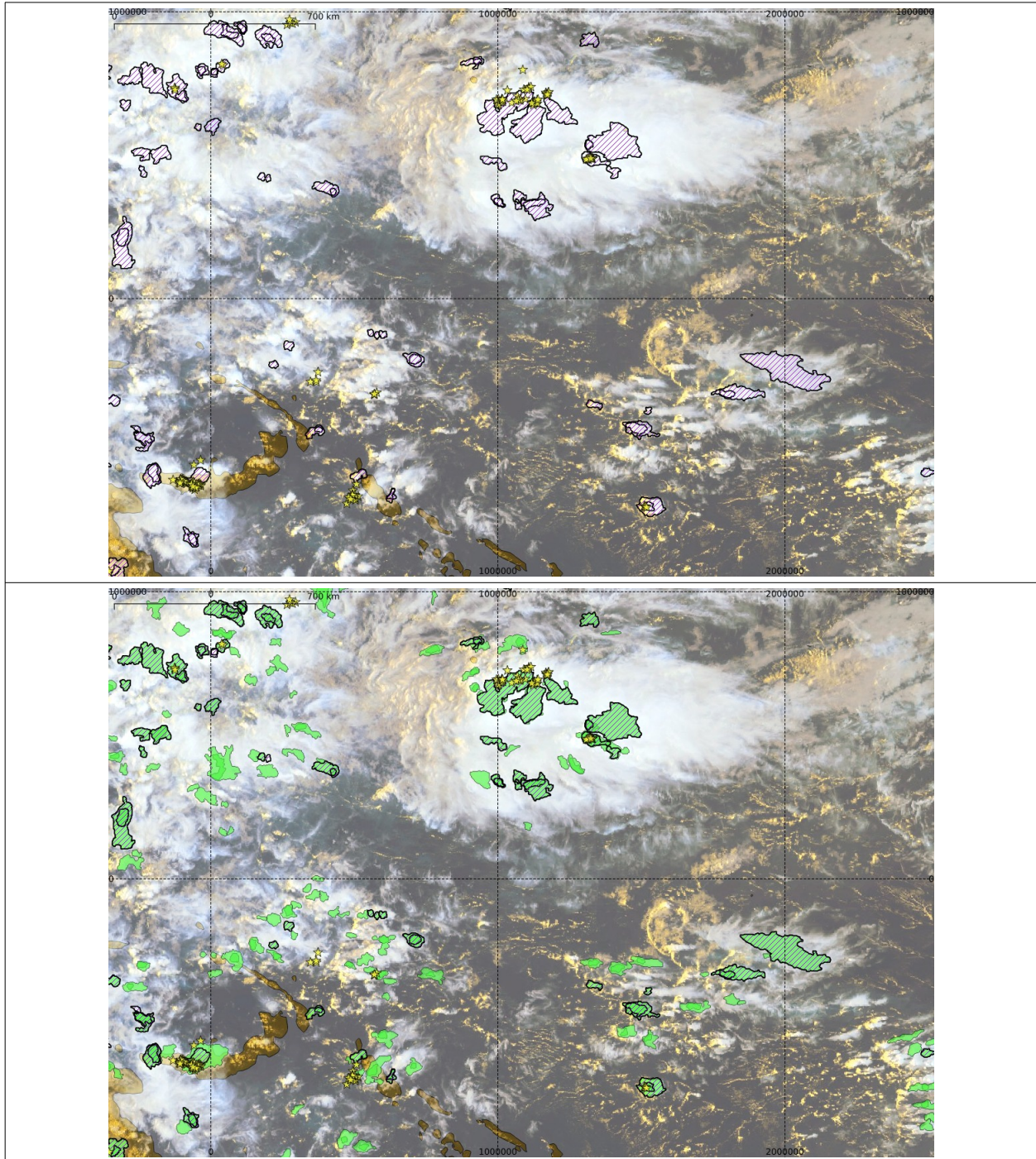
Figure 32: Himawari-8 case study for 06h00Z on 20210326. RGB image with 1h-accumulated WWLLN impacts around 06h00Z overlaid with RDT-CW v2021 black dashed contours (top), and with RDT-CW v2018 light green cells (bottom)

3.3.2.5.2 Case study 20180117 over Indonesia

Figure below illustrates with 06h00Z RGB image over Central Indonesia a large amount of convective clouds. Though, electrical activity is not so spread but rather concentrated in the South-Eastern part of the domain. The majority of RDT-CW cells for this slot are associated with WWLLN impacts and/or with suspected MCSs on RGB image. Highlighted are apparent convective systems

identified by RGB and RDT-CW without electrical activity (green circles), and low electrical activity missed by RDT-CW (red circles). One can note also some false alarms with RDT-CW (blue circles).

But with the use of IR10.3µm channel and GOES16 statistical models, those false alarms seem largely reduced when compared to previous release, and we consider this discrimination scheme brings valuable results when applied to HIMAWARI.
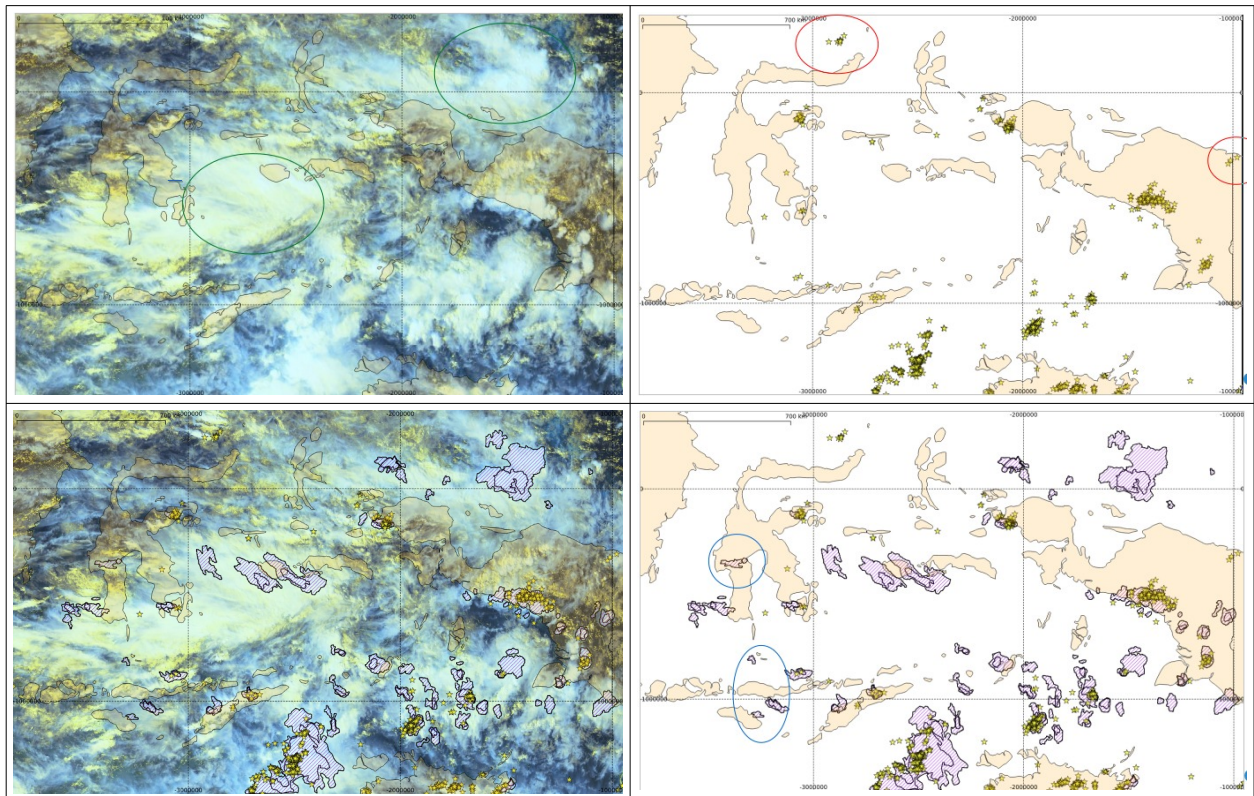
Figure 33: Himawari-8 case study for 06h00Z on 20180117. RGB image (top left), 30min-accumulated WWLLN impacts around 06h00Z (top right), RDT-CW overlaid with WWLLN (bottom right) and all data superimposed (bottom left). Supposed false alarms (blue circle), good detections (green circles), misses (red circles) are indicated

3.3.2.5.3 Case study 20180702 over East Asia

During this situation, convective cells develop over land on the Northern part of the domain, and a tropical disturbance is moving northward in the Southern part, mainly oceanic.

Next figure highlights some misses by RDT-CW, in particular over Mongolia. Even if on the edge of the domain chosen for this case study, those isolated convective clouds should have been identified by RDT-CW. We suspect here a lack due to the sub-nominal update rate at our disposal. Though, this release seems to be more efficient for focusing on convective activity, lowering the number of false alarms regarding previous v2018 product (highlighted with blue circles). This is the case for the head of the perturbation, but also for low clouds over China.

Finally, v2021 discrimination scheme for Himawari seems to help us to lower the number of false alarms, which was an issue.

Figure 35 illustrates an even better RDT-CW v2021 performance, with very few misses and false alarms, and a good identification of all obvious or apparent convective systems. Almost all WWLLN strokes are associated to or very close to a RDT-CW contour cell.

Here again, RDT-CW performs better than previous version.
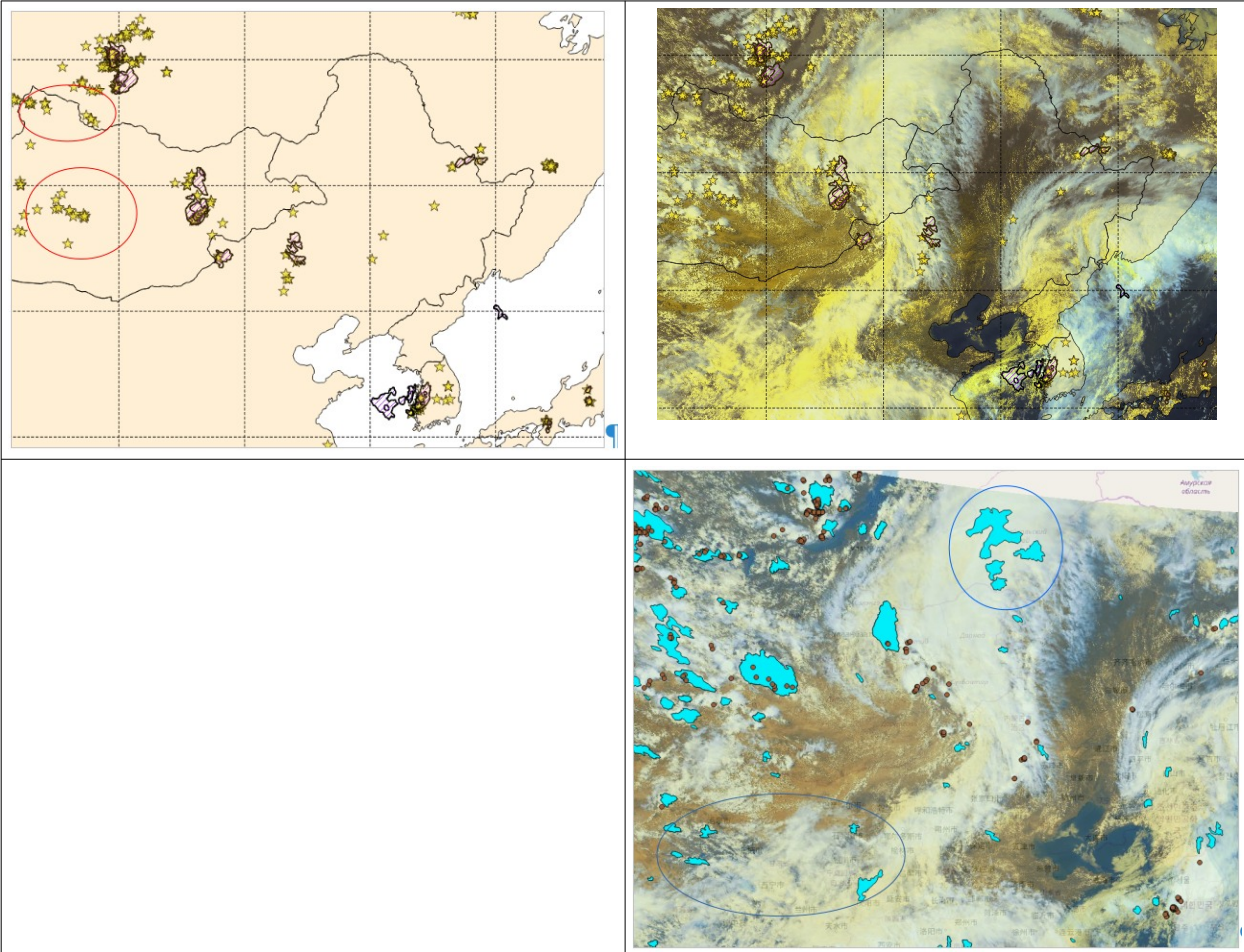
Figure 34: Himawari-8 case study for 20180702 06h00Z. Northern inner land Est Asian domain. RDT-CW v2021 black dashed contours with WWLLN strokes as yellow stars (top left), same overlaid with RGB (top right), RDT-CW v2018 blue cells (bottom right).
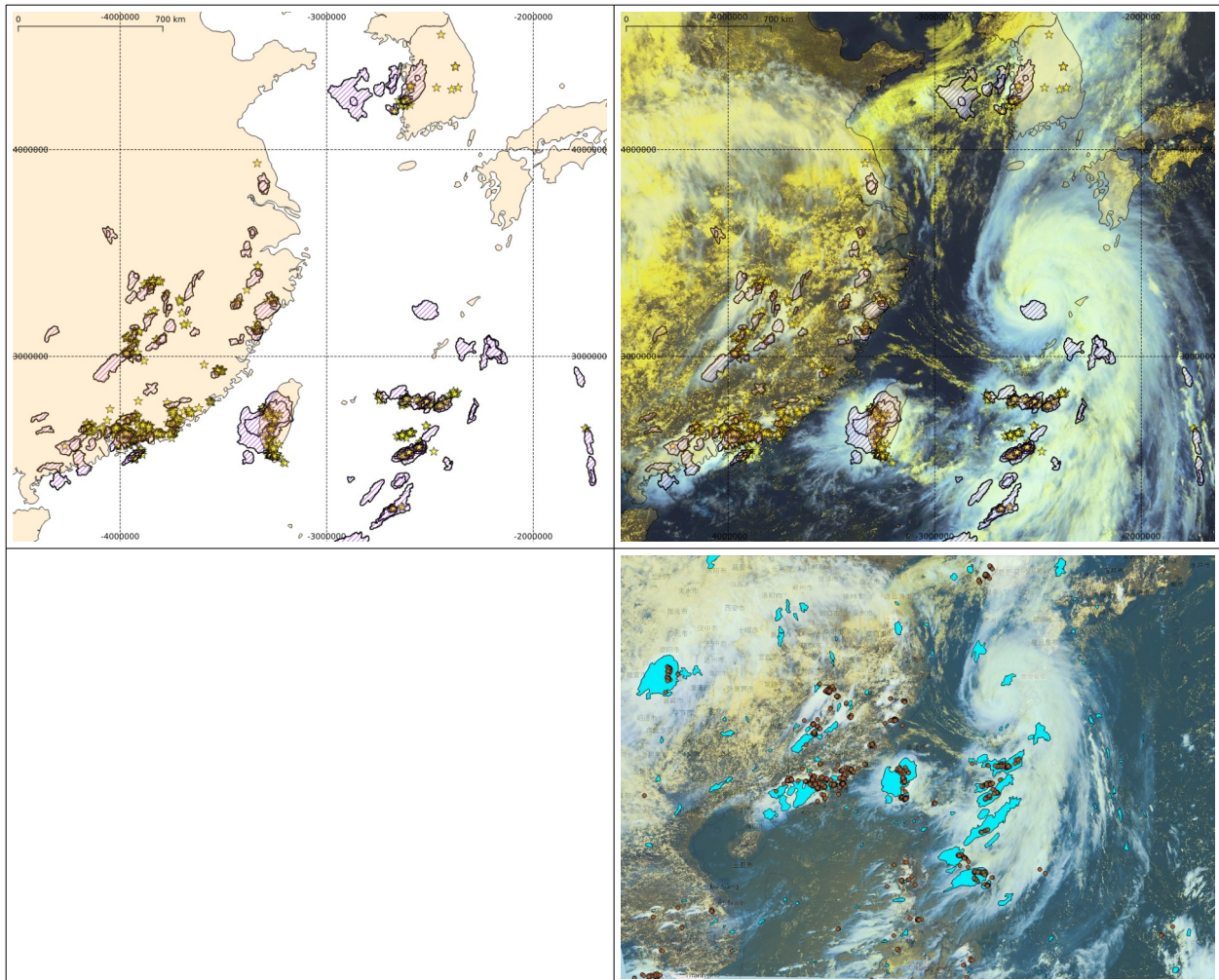
Figure 35: Himawari-8 case study for 20180702 06h00Z. Southern oceanic Est Asian domain. RDT-CW v2021 black dashed contours with WWLLN strokes as yellow stars (top left), same overlaid with RGB (top right), RDT-CW v2018 blue cells (bottom right).

3.3.2.6   RDT-CW applied to GOES17 ABI

As mentioned in ATBD, like for Himawari case, GOES17 data are taken into account at Météo-France at a sub-nominal update rate of 30min, instead of 10min. Moreover, regular ABI cooling problems with GOES17 lead us to suspend any tuning activities with this satellite. Because GOES16 and GOES17 are same generation satellites, it made sense to use GOES16 statistical models for generating RDT-CW with GOES17. For validation purposes, GLM-GOES17 data were used as ground truth for cases study when available in our production centre, WWLLN data otherwise.

With a full oceanic coverage area, difficulties were expected to formally validate RDT-CW against electrical data. Practically, whatever the source GLM or WWLLN, many RDT-CW cells are never paired with flash data, despite convective characteristics on images. Thus, many supposed "false alarms" are generated. On the other hand, one can observe that RDT-CW will highlight most supposed significant cloud systems.

This is illustrated in the following cases study for several dates, hours and geographical zones.

		Validation report of the Convection Product Processors of the NWC/GEO	**Code:** NWC/CDOP3/GEO/MF-PI/SCI/VR/Convection
			Issue: 2.0.1 **Date:** 28th February 2021
			File: NWC-CDOP3-GEO-MF-PI-SCI-VR-Convection_v2.0.1.odt
			Page: 57/75

3.3.2.6.1 Case study 20190520

Figure 36 presented below shows a weak electrical activity in a tropical/equatorial fully oceanic region. On top left image numerous convective clouds of various dimensions are displayed for 21h00Z, north and south of equator, but rarely electric regarding synchronous 30min-centred flash data. Bottom left image shows RDT-CW cells superimposed for all small and large, electric and non electric suspected significant clouds, even if some large cloud systems are identified by small towers, depending on their morphology. Of course, we also suspect real false alarms here and there. But a display of a 3 hours accumulation of electric data in top right image indicates the difficulty of considering a ground truth in this region, and also shows that RDT-CW can early identify significant clouds prior to this characteristic.
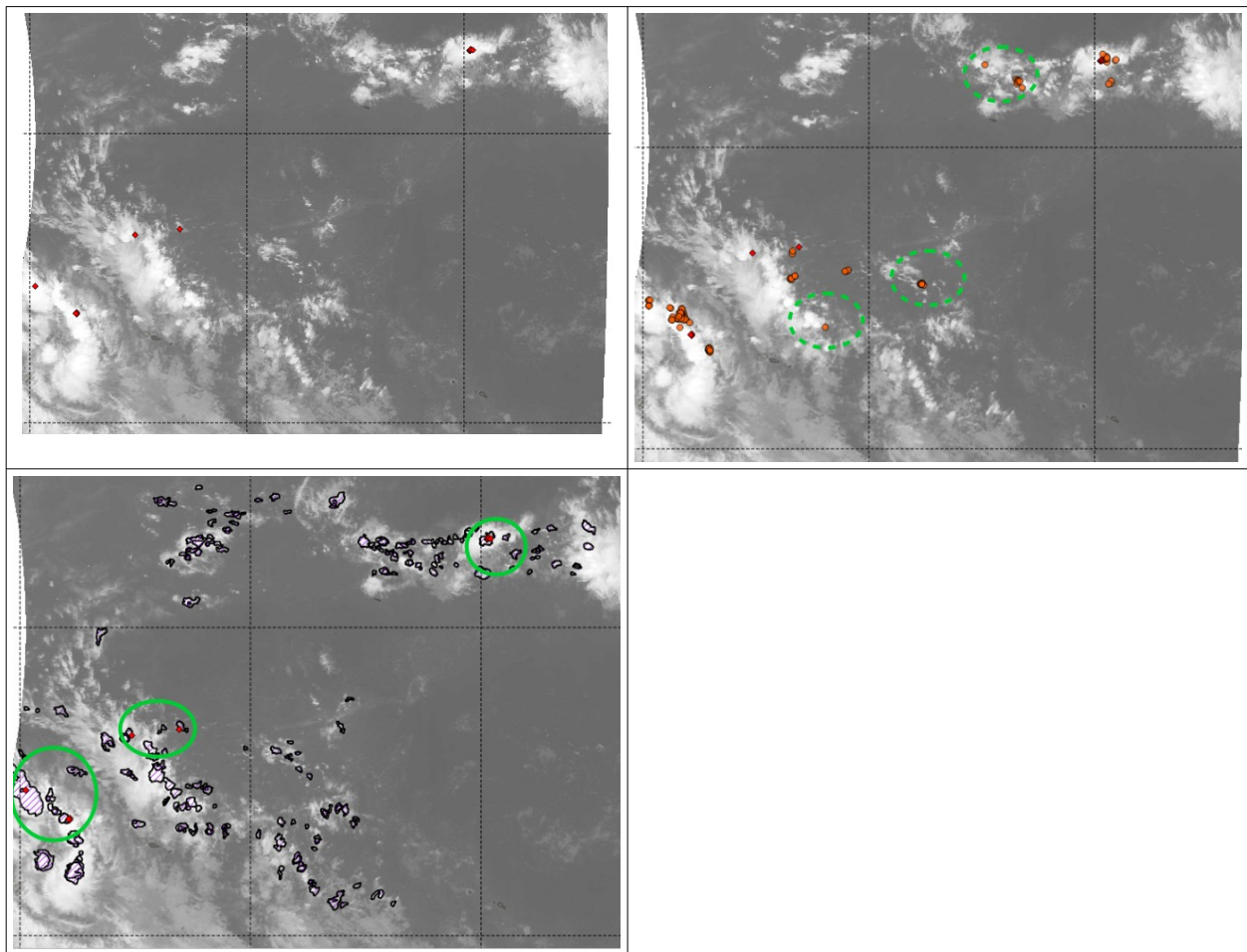
Figure 36: GOES17 case study for 20190520 , Mid Pacific region. 21h00Z IR image with synchronous WWLLN data (top left), same with following cumulated WWLN data between 21h00Z and 00h00Z on the 21st (top right), and synchronous IR, WWLLN data and RDT-CW black dashed contours (bottom left).green circle for hits, dotted green circle for early hits lately confirmed by flashes

3.3.2.6.2 Case study 20191211

Figure 37 below displays RDT-CW results for two different daytime periods (03h00Z and 21h00Z), and two different geographical areas (equatorial and tropical).

In the Equatorial area (Figure 37, lower panels) there is less electrical activity than in the Tropical area, but the number of RDT-CW is significant. Multiple cloud towers of seemingly convective

characteristics and linked to the same cloud system, are identified with RDT-CW in the centre of the image, causing a large number of suspected false alarms and poor hit ratio. Electrical activity is localized only in a small portion of the system and associated with several misses

Tropical area (upper panel) reveals here more favourable agreement between RDT-CW and electrical flashes with lots of hits, and little obvious false alarms The representation of a big convective system by several RDT-CW cells make an objective comparison difficult.

Figure 37: GOES17 case study for 20191211. IR image with synchronous WWLLN data (left), and with RDT-CW black dashed contours (right). 03H00Z on an area centred on the Samoa (top), 21h00Z in an area just north of equator (bottom)

3.3.2.6.3 Case study 20201123 over French Polynesia

This case gives two possible interpretations of RDT-CW behaviour :

> 1) The weakness of RDT-CW when identifying clouds never becoming electric.

> 2) The ability of RDT-CW concerning the identification of suspected significant convective cloud systems in the absence of electrical activity or radar information (example : cells in the Northern part of the domain in the figure herafter).

Second point is very sensitive and the comparison between this release and the previous illustrates the effort in lowering suspected false alarms (example: cells in the South-West part of the image, seen by v2018 and not v2021), keeping anyway a focus on what can be considered as cloud systems of interest.

The absence of trusted ground truth makes nevertheless difficult to evaluate the real relevancy of RDT-CW product for this area.
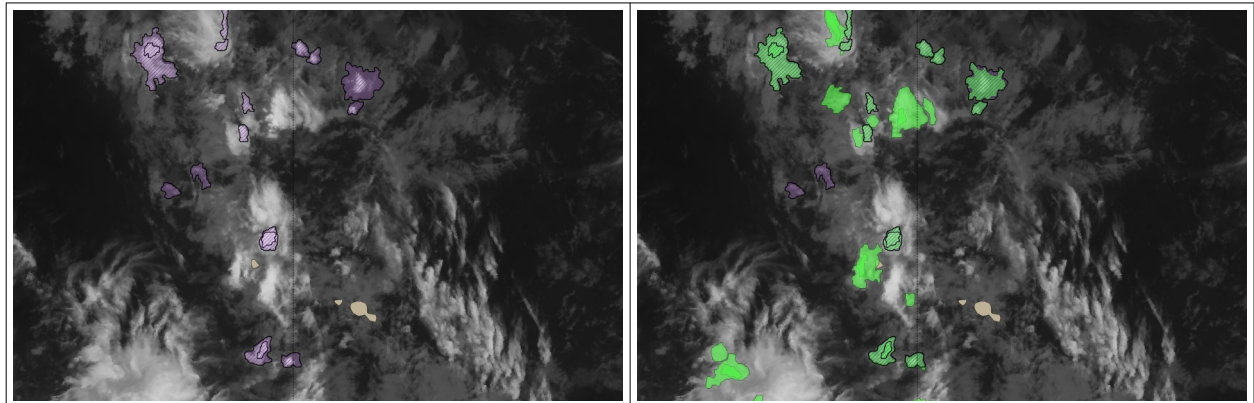
Figure 38: GOES17 case study for 20201123. 20h30Z IR image over French Polynesia with RDT-CW Left panel: v2021 cells in magenta, dashed. Right panel: same, overlaid with cells from previous version in green, cells detected by both versions appear in green and dashed. Tahiti Island in orange. Absence of electrical activity from GLM-GOES17.

3.3.2.6.4 Case study 20210418

This case study focuses on a reduced oceanic domain North of Samoa. For this situation, GLM has been used for validation. Figure hereafter highlights for 21h00Z, relatively numerous and equally-distributed flashes over the domain. Almost all flashes in the period [20h45-21h15] are associated with or close to a RDT-CW cell. Almost all RDT-CW cloud cells correspond on the RGB image to bright cold convective cloud. But here again, not all those clouds become electric. Consequently, despite an apparent very good subjective agreement between RDT-CW and RGB image, the number of false alarms remains high regarding electrical activity. One can note very few misses, all occurring with RDT-CW cloud cells in the vicinity.

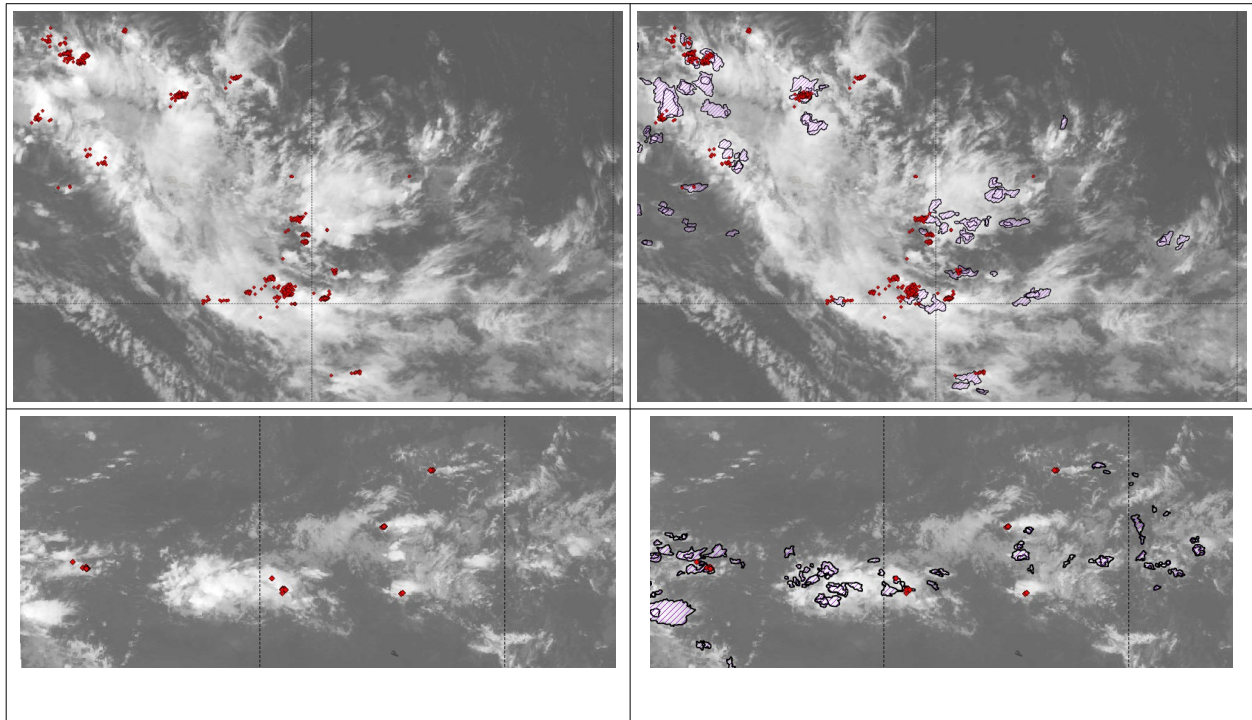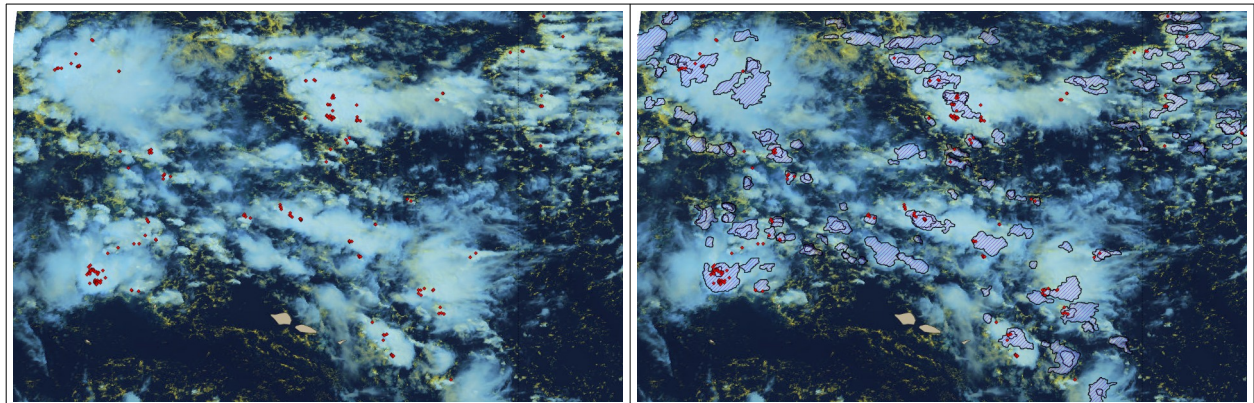*Figure 39: GOES17 case study for 20210418. RGB image with synchronous WWLLN data (left), and with RDT-CW black dashed contours (right).*

3.3.3 Conclusion about RDT-CW convection diagnosis validation

This validation approach, based on various situations, satellites, geographical regions, and periods, give us enough elements to consider RDT-CW v2021 discrimination scheme as relevant.

An objective validation provides scores reaching the requirements, especially in terms of POD. Some meteorological situations are sometimes associated with FAR under requirements, but average values remain acceptable.

It remains difficult to undertake an overall and global objective validation for all satellites, regions and seasons, regarding the methodology used. Thus a subjective analysis of various sampled situations is necessary to complete this approach.

Case studies cover several meteorological situations: land or sea surface, mid-latitudes, tropical regions, equatorial latitudes, seasons (except winter). An overview of those cases study reveals the ability of RDT-CW to identify most significant convective clouds. If some misses can be observed, the convection can be sometimes identified in the following image. In any case, theses misses would have been recovered in real-time mode if RDT-CW is operated with lightning data used for the diagnosis.

The main issue is the number of false alarms that is sometimes observed in some particular regions. This number seems difficult to lower. False alarms can be obvious or just suspected when no adapted ground truth is available. This is in particular the case for all tropical oceanic regions, where the RDT-CW seems rather adapted to identify all convective clouds, even if they don't evolute towards deep convection.

3.4 OVERSHOOTING TOP DETECTION

3.4.1 Overview

As detailed in [AD.11], Overshooting Top Detection (OTD) in RDT-CW code is undertaken in two steps.

- First, morphological analysis of cloud cells' top allows identifying cell's list of so-called "OT-candidates".

- Then, OT candidates are eliminated or confirmed considering a combination of thresholds of common BT or BTD available for all satellites (this condition has been set more restrictive in v2021), additional information from NWP (gap to tropopause), or from additional channels when available (high resolution visible, IR9.7 or IR13.4)

Criteria are inspired from existing bibliography about OTD, and have been adjusted and subjectively validated on case studies and regarding routine production. With OTD, we have the first use of visible channel in RDT algorithm (VIS 0.6), and possibly high resolution for this channel (HRV for MSG series).

3.4.2 Objective validation vs expert CHMI OT database

3.4.2.1 Context

A CHMI overshooting top database has been made available by Convection Working Group [RD.15]. This data base has been used to undertake a comparison with RDT-CW diagnosis of overshooting top detection. This database, processed on the 20th June and 29th July 2013, is issued from meteorologists' expertise with MSG1 on a 2.5 minutes super rapid-scan experiment from EUMETSAT over small regions in Central Europe.

The objective was first to assess the relevancy of RDT-CW OTD, and then to evaluate in this diagnosis the contribution of a high resolution visible channel as set in v2021 release.

3.4.2.2 Methodology

RDT-CW has been generated on those two periods with a configuration similar to v2021(input from additional optional PGEs, from lightning network and NWP), applied to MSG3 with 15min update rate and MSG2-RapidScan at 5min update rate.

As expected, for those two days, the number of expert OTD increased with diurnal convection. But, in the expert database, the benefit from a high scan rate of 2.5min, and the extended use of high resolution visible channel lead to a much higher number than with a default configuration of RDT-CW OTD in NWCSAF: about ten times more than with RDT-CW-MSG2-RapidScan mode, and almost twenty times more than with RDT-CW-MSG3 15 minutes scan. Moreover, the multiple buddings of an active convective system are all referenced in the expert database while the number of overshooting tops in a RDT cell is limited to two.

In such conditions, a quantitative comparison between expert OTs and RDT-CW OTDs will clearly lead to a huge number of misses. The quantification of RDT-CW OTDs false alarms will on the contrary be regarded with attention. For a reliable comparison, the exact SEVIRI dates have been taken into account. Regarding the mean latitude and size of the domain, we add 1 minute to expert OTs' dates, 3 minutes to RDT-CW-MSG2 OTs in RSS mode, and +11 minutes to RDT-CW-MSG3 OTs.

In the pairing process several tolerance thresholds have been tested , with following final choices:

- Temporal tolerance depending on update rate: 5 minutes for RDT-CW-MSG2 RapidScan, 15 minutes for RDT-CW-MSG3

- Spatial tolerance: 20 km

For a quantitative comparison, dataset without parallax correction have been taken into account, for experts OTs like for RDT-CW OTDs.

3.4.2.3 Example

Figure hereafter highlights the difference between an expert's analysis of OT presence (CHMI database) and an automated one (RDT process). The 16h40Z slot of MSG1 is here associated with 15 expert-assessed OT, most of them associated with obvious buddings. RDT-CW detection process of OT applied to slot 16h30Z (corrected radiometer date 16h41 for MSG FDSS over Europe) of MSG2 was able to detect one overshooting tops for each cloud system. Experts can detect several OT inside a cloud system but there is a limitation of the numbers of OT inside a RDT cell. The most spectacular case concerns the central system where experts have diagnosed eight OT while RDT has detected one.

Two RDT OT out of three are very closed to an expert-assessed OT. The third one is Southward the expert-assessed one, in the colder part of the system. The detection of relatively warm OT is difficult in a automatic process.

For other slots (not shown) conclusions remain the same: when a OT is detected by RDT this OT corresponds to a OT diagnosed by CHMI experts (given a reasonable space or time tolerance). Given the use of CHMI data base for reference, the number of misses by RDT remain significant.
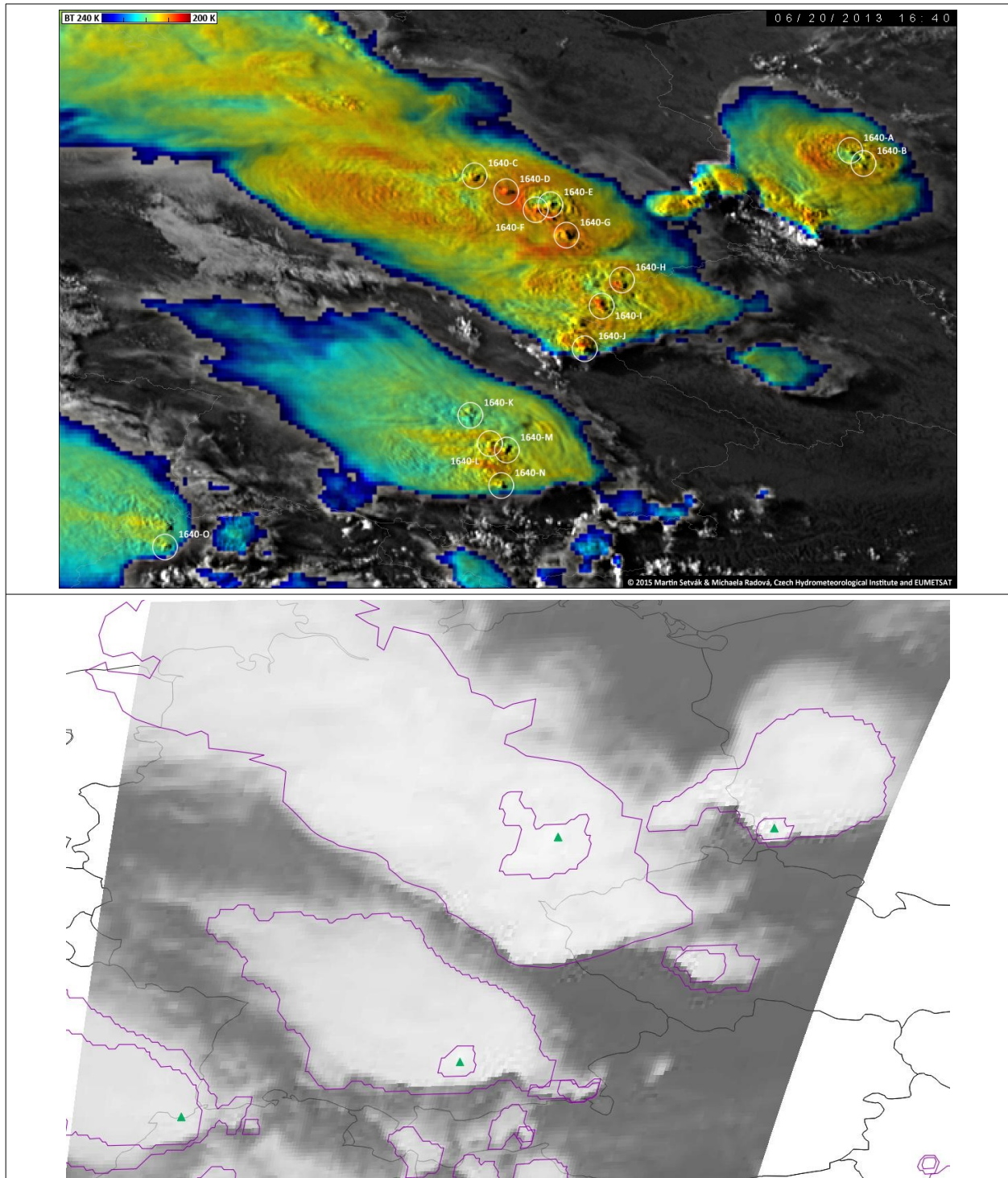
Figure 40: Sandwich image (HRV+IR10.8) with expertised OT for 16h40Z on 20/06/2013 (top). IR image with RDT-MSG cells from 16h30Z slot (+11minutes for exact radiometer date) and OTDs as green triangles (bottom)

3.4.2.4 Results of quantitative comparison

The 20130620 daytime period is much more active than for 20130729, with three times more expert-assessed areas with overshooting top. Nevertheless, the ratio between expert OTs and RDT-CW OTs

remains the same. In both cases, the overshooting top activity is concentrated in the afternoon, with a high rise of cases beyond 16h00Z.

For each RDT-CW OT, expert-assessed OTs are searched given a space and time tolerance. There are thus:

- Hits: the number of pairing between an expert-assessed OT and a RDT-CW OT, knowing that a given expert-assessed OT could be paired with several RDT-CW OT

- False alarms: the number of orphans RDT-CW OTD

- Misses: the number of expert-assessed OTs paired with any RDT-CW OT

For this kind of validation, the number of true negatives is of few interest. Thus, we mainly evaluate the POD and the FAR for this comparison.

Despite the good results of the subjective analysis comparing RDT-OT and expert-assessed OT as seen in the previous paragraph, the objective scores calculated for the whole dataset exhibit low PODs and significant FAR (table hereafter). POD are between 10% and 20% and FAR between 22 and 28% with respectively MSG-FDSS and MSG-RSS. It confirms the interest of high-frequency scan for the detection of short-lived phenomena like OT.

Given the ratio between RDT-CW OT and expert-assessed OTs (between 10 and 20 %), the comparison is clearly not in favour of RDT-CW .
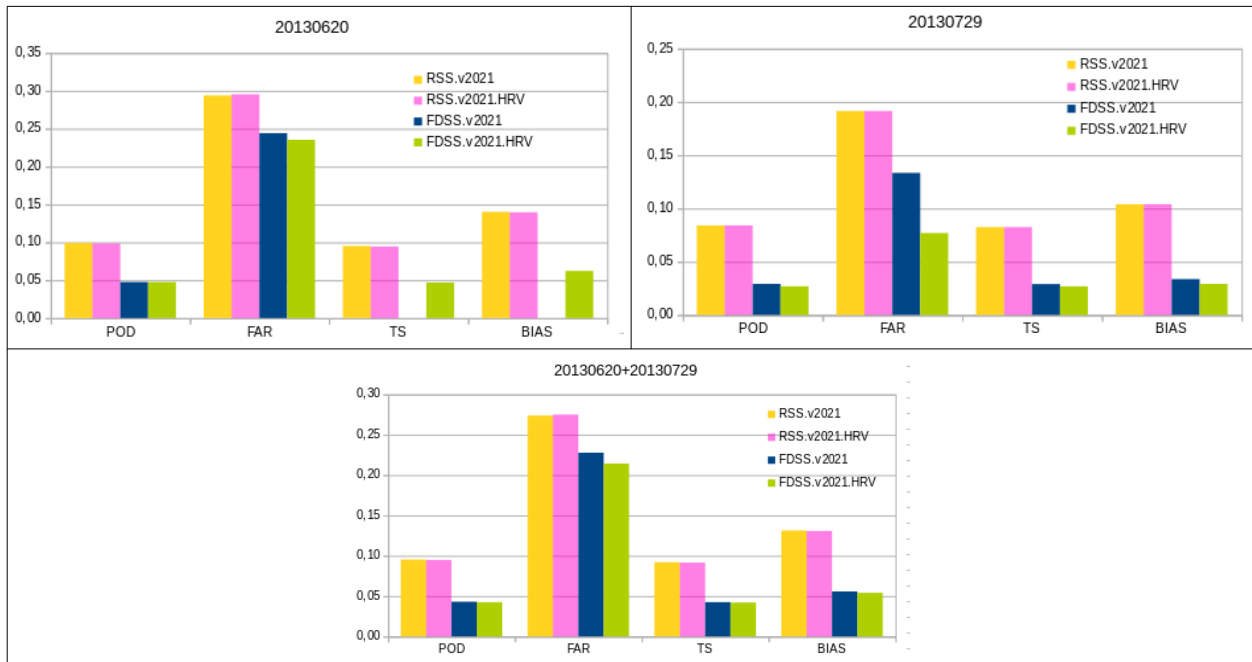
Figure 41: scores of RDT-CW OTD vs CHMI expert OTs for the two periods 20/06/2013 (top left) , 29/07/2013 (top right) and both day together (bottom)

3.4.2.5 Synthesis

The objective verification of overshooting top detection within RDT-CW revealed a significant number of false alarms compared to the number good detections. This was already suspected on operational productions with the previous release.

This release includes modifications in RDT-CW OTD algorithm, which aimed to lower those suspected false alarms, or excessive number of OTD. It appears that mechanically, the number of detections also decreased when compared to CHMI expert OTs database.

The use of high resolution for visible channel very poorly improves the scores of objective validation. Further work appears necessary to validate or improve this approach in the overshooting top detection algorithm.

3.5 LIGHTNING JUMP DIAGNOSIS

A lightning jump detection is implemented since version 2018 of RDT-CW. The algorithm takes benefit from full lightning activity (for details see [AD.11]), and relies on minute lightning analysis inside RDT cell for a period of 12 minutes with a condition on lightning rate and lightning rate trend.

The main objectives of this diagnosis are to contribute to severity index and to be used as a precursor of hazards like hail.

The assessment of this attributes needs to access time and localisation of hail events or other strong hazards, and ensure that those events occur on the path of a RDT-CW cell after or at the same time of a Lightning Jump (LJ) LJ diagnosis. It also needs to take benefit of an efficient and reliable lightning network.

Up to now, only some subjective validations of this attribute have been attempted on cases study over France and neighbouring countries, using Meteorage lightning network for RDT-CW, and following references as ground truth for validation:

- Météo-France HYDRE product over France, which includes an Hydrometeor diagnosis thanks to data fusion with data from radar, satellite, NWP and observations. This product is updated every 5 minutes, and provides reliable diagnosis of medium/large hail

- ESSL European Severe Weather Database (ESWD) over a larger domain, with reports of severe convective weather events, like hail, wind gusts, tornadoes, lightning damages, etc.

Hereafter are presented some cases study using those ground truths.

3.5.1 Case study 20180529 over France and Benelux

Figures below highlight RDT-CW cells associated with Lightning Jumps diagnosis prior to hail events from HYDRE and/or ESWD severe weather events.

A comparison with HYDRE product shows a subjective good collocation between hail diagnosis and RDT-CW cells associated with LJs. Even if it is still to be confirmed, LJ diagnosis seem here to be sometimes precursors of hail events.
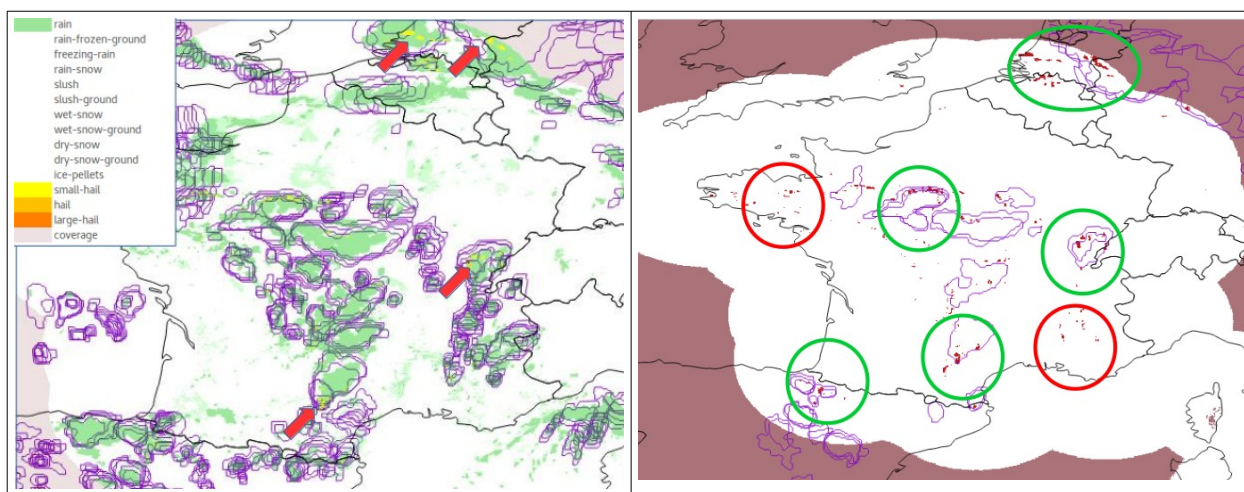
Figure 42. 20180529 Case study. Left: cumulated [15h30-16h00] RDT-CW cells overlaid with [16h00-16h15] HYDRE product. Right: filter on RDT-CW cells with LJ and on hail diagnosis with HYDRE (red pixels for small/medium/large hail classes)

The same situation is regarded below through ESWD severe weather events over Benelux.

Regarding RDT-CW cells associated with LJ vs ESWD convective reports (wind gusts, hail, lightning, tornadoes), one can estimate that most severe weather events find a correspondence with previous RDT-CW with LJ. There is a subjective good pairing, even if numerous RDT-CW cells with LJ are not paired with reports. It is however difficult to conclude if this is due to lack of observation or to RDT-CW false alarms.

Figure 43. 20180529 Case study. RDT-CW cells with LJs (left column) , consecutive ESWD severe weather reports (right column)

3.5.2 Case study 20190809 over France

In this situation, a south-west/nord-east axis of convection is observed in the end of the afternoon. Regarding the data accumulated over a 16h00Z-19h00Z period, pairing RDT-CW cells associated with LJ and medium/large hail diagnosed with HYDRE product, we can note a high level of matching.

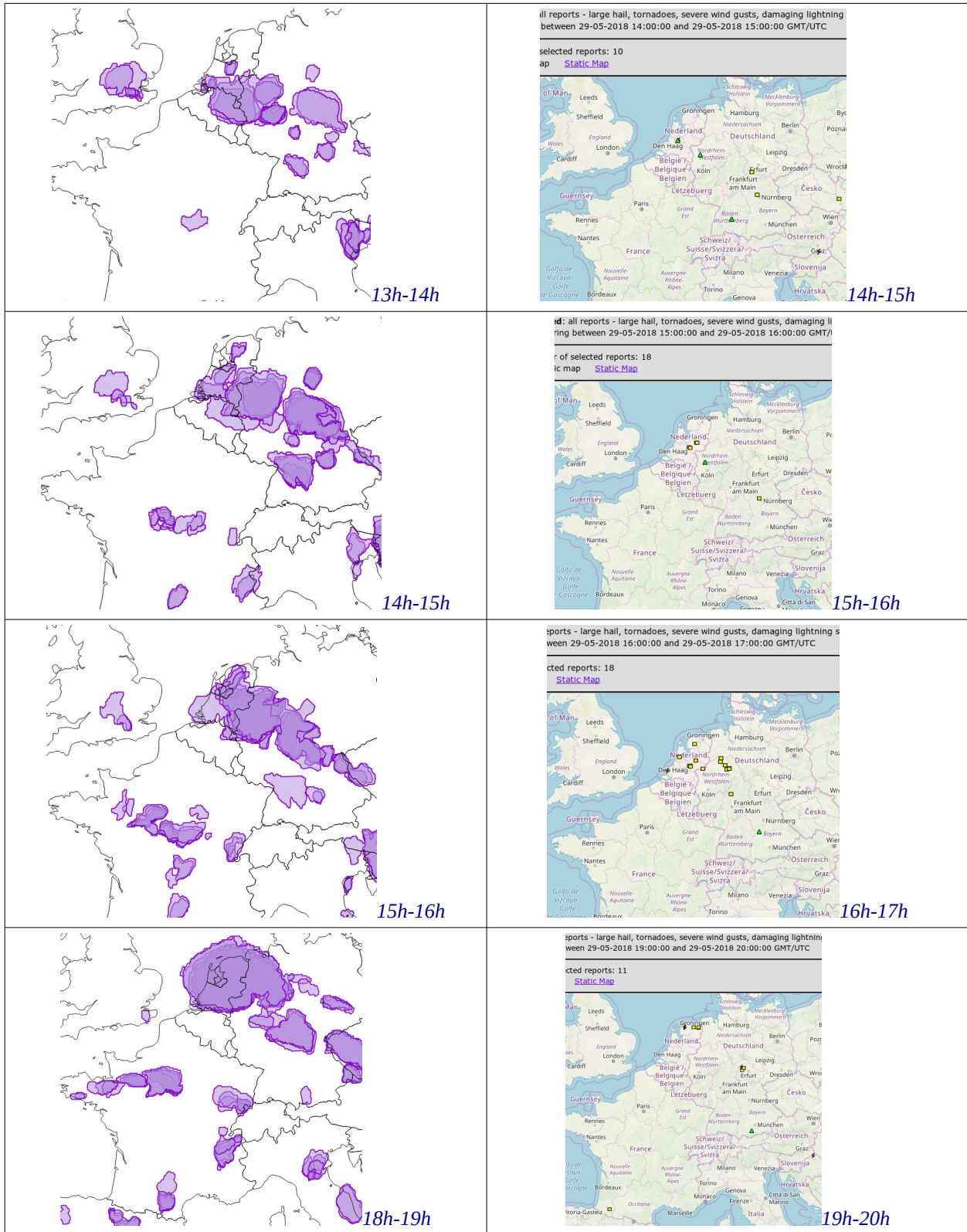Here again, despite some isolated hail pixels not paired with RDT-CW cells, and suspected false alarms South of Pyrenees, the situation appears favourable for considering LigntingJump diagnosis as relevant.
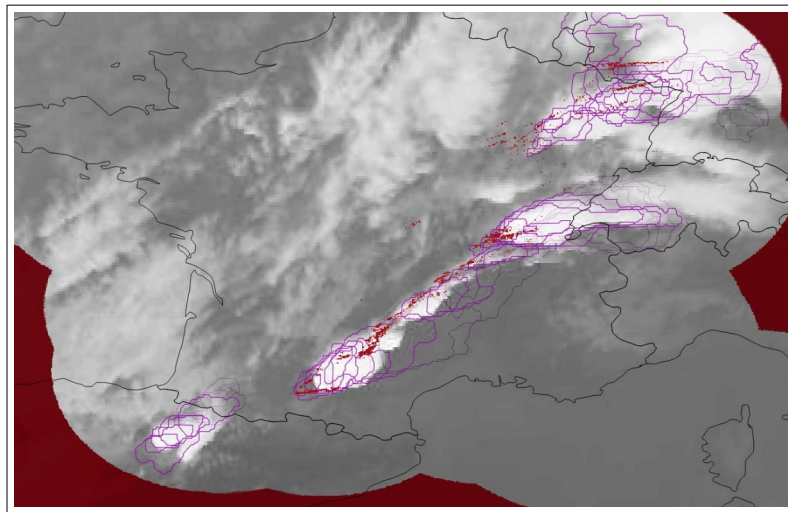

Figure 44: 20190809 16h-19hZ case study. Accumulated RDT-CW cells with LJ (magenta contours), and accumulated medium/large hail pixels from HYDRE (red pixels)

3.5.3 Applicability to GLM

Lightning sensors like GLM or LI provide a continuous and large-scale measurement of electrical activity of cloud systems. This activity is quantified in events, groups and flashes, but without discrimination of polarity.

RDT-CW is operated every 10 minutes with GOES16, with high rate netCDF GLM in input data. In order to eliminate artefacts, NOAA quality code is used for the pairing between GLM flashes and the cells. Lightning jump diagnosis is applied with the same configuration (parameters thresholds) as for ground lightning data.

In the absence of ground truth, this diagnosis with RDT-CW applied to GOES16 must be regarded with precaution. The number of co-located flashes can sometimes be very important with a lightning sensor, and the analysis of a 1minute lightning activity of a RDT-CW cell seems to provide a much larger number of lightning jumps.

In the example below, a RDT-CW cell over Bolivia lasts from 06h00Z to 16h00Z, is electrically active with multiple overshooting top detection, and many lightning jumps during its entire tracking (almost twenty).
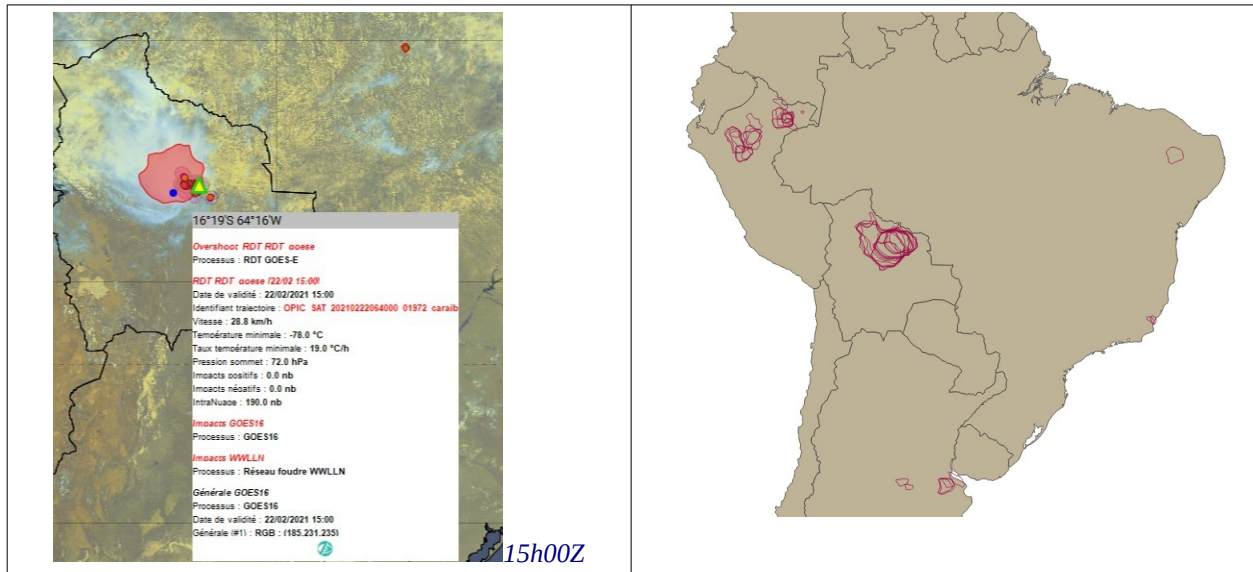
Figure 45: 20210222 case study with GOES16. RDT-CW cell over Bolivia at 15h00Z (left) with 190 paired flashes (here simulated as WWLLN) and OTD (yellow triangle). [06h30-16h00Z] cumulated RDT-CW cells with diagnosis of lightning jump during the period

The ten minutes time-series below (top graph of next Figure) shows two main peaks of increasing activity, but the RDT-CW lightning jump algorithm apparently "lights on" for more than twenty slots, almost continuously.

A deeper analysis of the 1 minute flash rate (or FR) highlights in the middle graph the necessity to adapt the configuration of the algorithm. Significant rises of electrical activity can be estimated once FR is above 20 min^{-1}, and not 10 min^{-1} as in the default configuration of the algorithm. With a stricter threshold of FR of 20 min^{-1}, the algorithm focuses on jumps which seem more relevant, as illustrated in the bottom graph.

3.5.4 Synthesis

First subjective analysis of cases study using Meteorage network has confirmed a good correlation between hail events and previous lightning jump (LJ) diagnosis in RDT-CW cells.

A wider monitoring of real-time LJ diagnosis at different periods with various couples of geostationary satellites and lightning networks (MSG-Meteorage, MSG-IODC/GLD360, GOES16/GLM) has pointed out an excessive number of diagnosis.

A look on 1 minute FR time series has confirmed this issue in the default configuration, and lead to double the FR threshold. This has been adopted for this release.
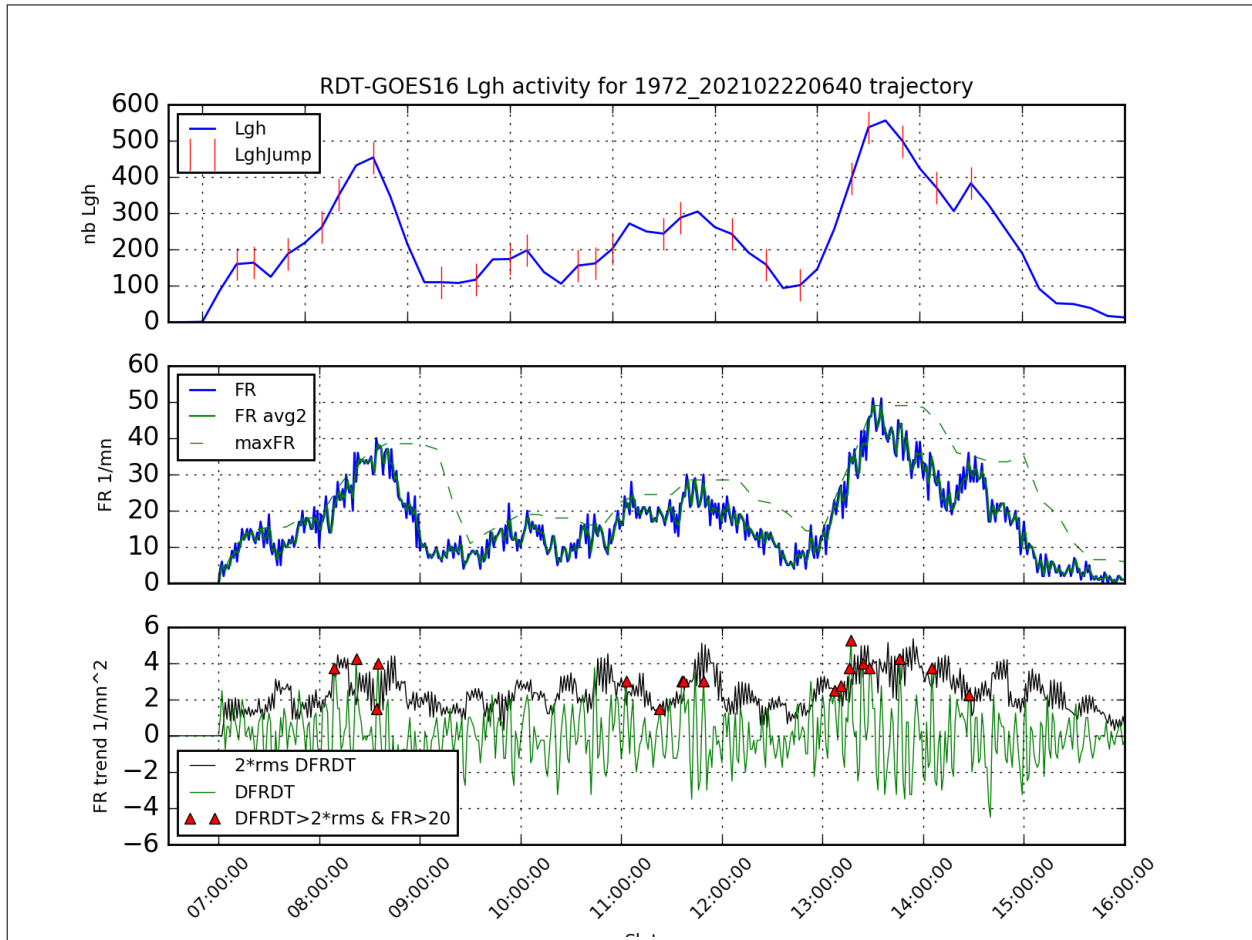
Figure 46: 20210222 case study with GOES16. [06h30-16h00Z] Time series of GLM-flashes paired with RDT-CW cell over Bolivia. Top: 10 minute flash count and diagnosed LJ with FR>10 (red bars). Middle: 1 minute FR and 2 minutes-average FR. Bottom: LJ algorithm parameters, and "corrected" LJ (red triangles) taking into account new threshold FR>20.

3.6 FORECAST OF CLOUD SYSTEMS

RDT-CW software provides the user the possibility to forecast cloud cell position using the diagnosed movement speed of the cloud cell.

Extrapolated cloud cells positions are obtained through Lagrangian forecast, i.e. the whole object is moved according to direction/speed of the diagnosed movement.

Consequently, quality of this extrapolation relies highly on the quality of movement diagnosis. The use of a pre-initialized guess of movement field from blending NWP wind data and HRW, and a final-step coherence checking has allowed to filter erratic speed and/or direction due to split/merge or to threshold temperature changes. Thus, confidence is Lagrangian extrapolation has become much higher, as illustrated in figure below.
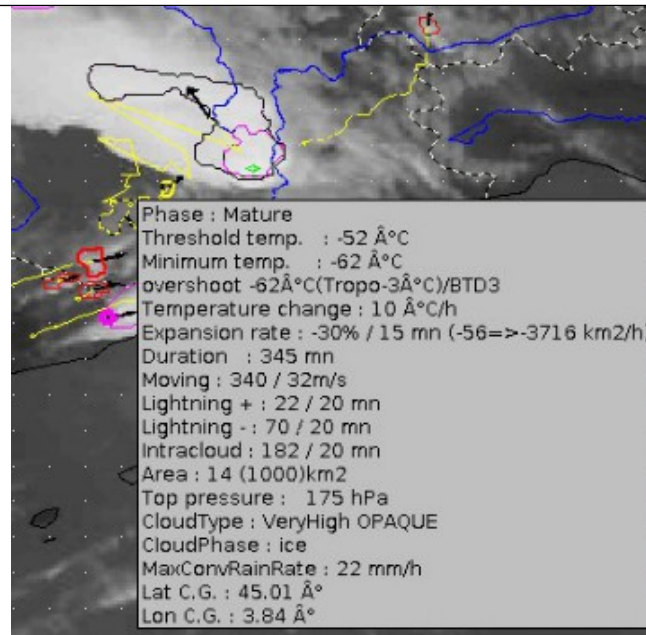
V2013 : erratic speed and direction following numerous splits/merges along trajectory
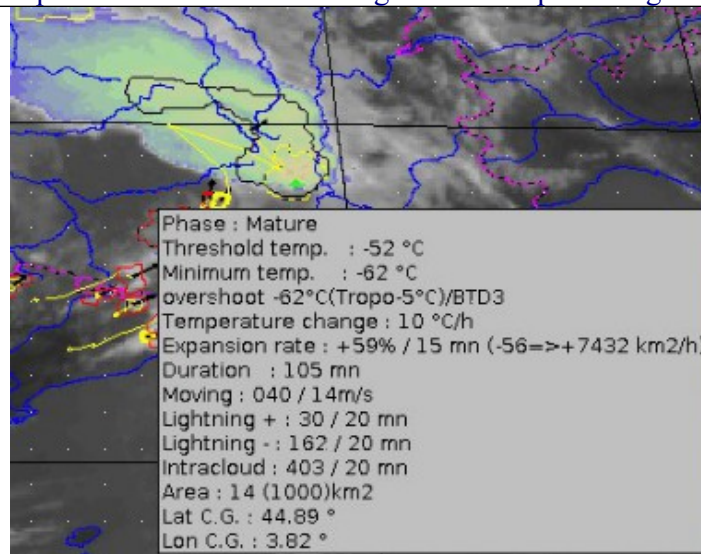
V2016 : coherent speed and direction despite numerous splits/merges along trajectory

Figure 47: v2013 vs v2016 illustration of RDT motion vectors improvement

Concerning forecast positions, it must be reminded that a cloud cell definition/contour corresponds to a given threshold temperature, and that this threshold changes dynamically/automatically from one slot to the other. ***A forecast contour remains based on the same temperature threshold than the observed/analysed object***. It will consequently be very difficult to assess the position of a forecast contour when compared to the following corresponding observed contours, because it is likely that those contours will not correspond to the same threshold temperature. Moreover, the longer the forecast ranges, the less precise the localization of forecast convective object is.

Assessment of forecast products position should ideally take into account for each cloud cell the stage of development, the morphological evolution or the expansion rate. Forecast cloud cell position can consequently only subjectively be regarded.

The objective validation of the forecast part has been carried out in a scientific report in 2018 [RD.3.], some of the results are still available for this version. Around 20% of cells are new at each

slot of MSG (15' scan) over Europe. Since the forecast scheme doesn't see these new cells, it gives the ratio of misses at step +15'. The ratio of false detection is also around 20%. Concerning the error on gravity centre position, the error value of 50 km is reached after the forecast range of 60 minutes.

Trends attributes of observed cloud cells are used to give indication of possible values changes of some parameters in the first forecast ranges (trend's values used for 15min range, half values for 30min). Here again those values can only be regarded as estimation and not pure forecast:

- Temperature change of coldest part of the cell (up to tropopause temperature limit if this value is available thanks to NWP data) and severity

- lightning trend for lightning activity and severity

- Top pressure trend for top pressure estimation

- Expansion rate for amplification or limitation of contour dilatation
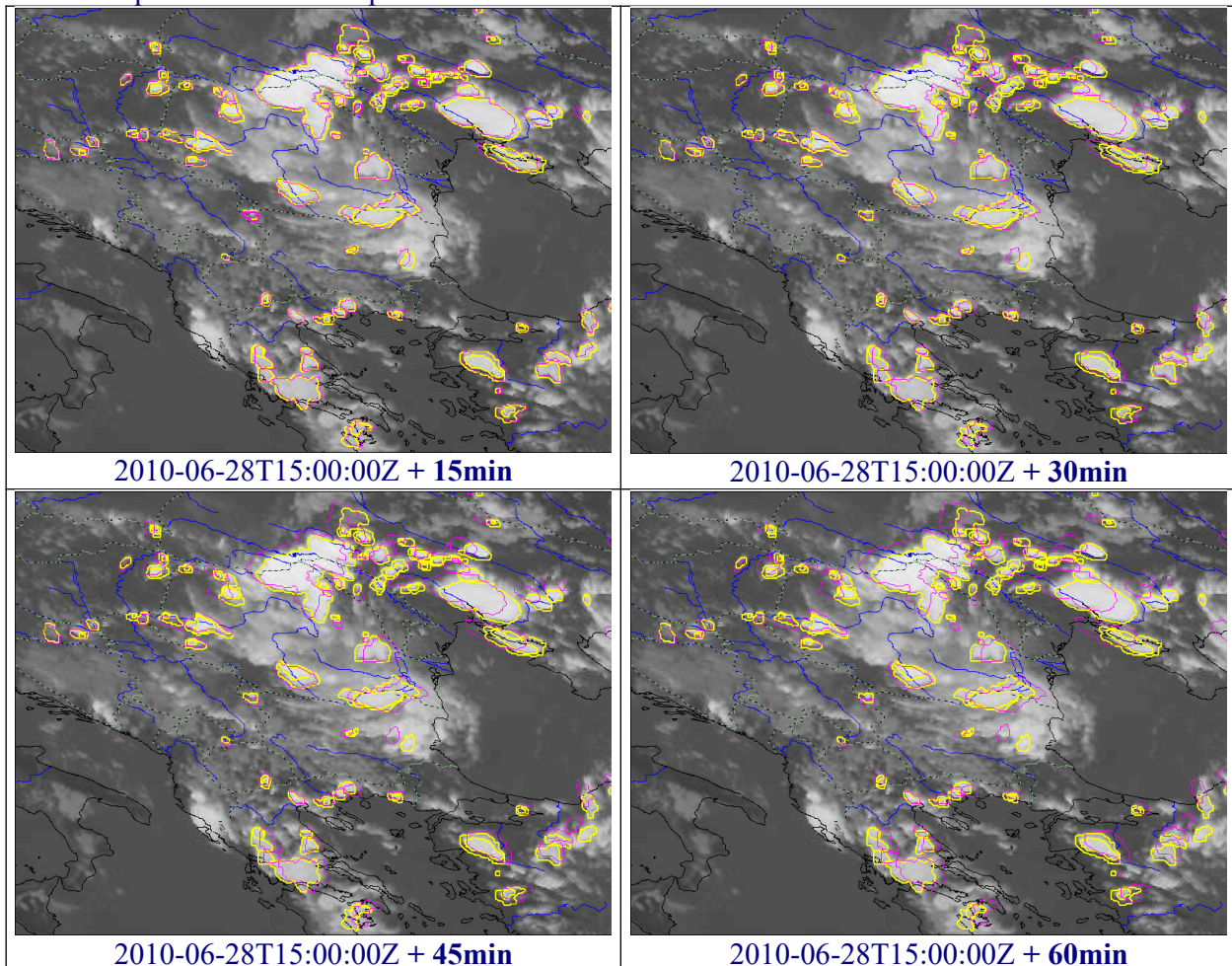
Figure 48: RDT-CW v2016 advection products (forecast contours in Magenta) from slot 2010-06-28T15:00:00Z (Observed contours in yellow).

Figure above displays forecast products (magenta contours) issued from a given slot 1500Z on the 28-06-2010. Yellow contours are observed at 1500Z. One can note very few overlap between forecast cells at various ranges, which assess a good spatial coherence of the movement speed of each cloud system.

The cyclonic movement field is very well taken into account with the forecast cloud systems.

This forecast set has been produced without smoothing neither dilation of the contours.
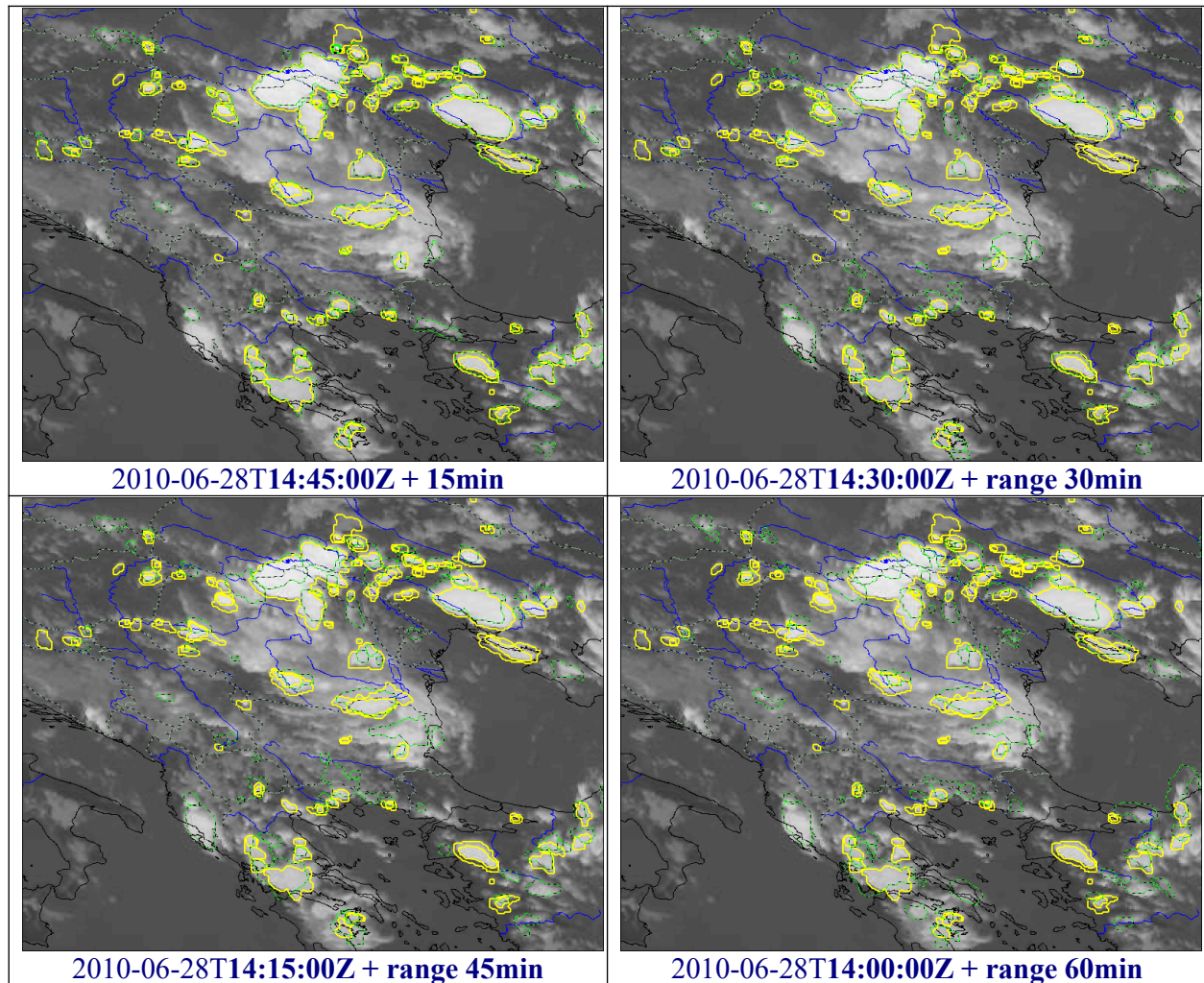

Figure 49: RDT-CW v2016 advection products (green forecast contours) from previous slots valid for slot 2010-06-28T15:00:00Z (yellow observed contours).

Figure above displays forecast products (green contours) valid for a same slot 1500Z on the 28-06-2010. Yellow contours are observed at 1500Z. The first ranges (15 and 30min) show a pretty good correspondence between previous forecast and current observation. For larger ranges, mainly large cloud systems, with longer duration, find a correspondence in the observed set.

One can note that some forecast green contours are no more valid in the analysed set, probably after declassification of the cloud system (no observed yellow contour).

On the other hand, some new cloud cells (yellow contours) appear only on the analysed set, and can not be anticipated from the previous slots.

3.7 END-USERS FEEDBACKS

RDT, a very satisfying product widely used for Research and Operations, by Météo-France and its partners.

The use of RDT concerns for example

- Forecasters of Météo-France, in France and overseas territories (La Réunion, Antilles, Polynésie, Wallis et Futuna, Nouvelle Calédonie). RDT provides a significant help for regions not covered by radars.

- HAIC project and successive experiments (2014 and 2016 Australia, 2015 French Guyana, 2016 Darwin/La Réunion)

- RDT provided to airlines through EFB (Electronic Flight Bag)

- Collaboration has previously concerned

- SESAR project and TOPMET/TOPLINK experiments

- Hymex project

- 2006 AMMA experiments (http://aoc.amma-international.org/observation/mcstracking/)

- European FlySafe Project with RDT software adapted to radar data

- NOAA for a RDT GOES (Operation + Research)

- From 2008, 2010 and 2015 Surveys distributed to SAF/NWC users, it appeared that RDT is mainly used for Research activities and operations for forecasting. The judgement of overall quality of RDT product is very satisfying (rate of 6.7/10 in 2015 survey).

RDT produced by Météo-France is downloaded by ACMAD and some African countries can visualize real-time RDT through ACMAD website (www.acmad.net).

RDT is operated by SAWS, hereafter two feedbacks

• De Coning, E., Strydom, J., Powell, C. , Gijben, M., de Beer, A., 2016, Nowcasting for aviation purposes in South Africa – a case study: Part1 – Satellite and radar based tools, WSN16 Hong-Kong, 25-29/7/2016

"[the RDT] has provided good validation against lightning occurrence and radar reflectivities of more than 35 dBZ over South Africa"

• De Coning, E.,Gijben, M., 2017, Using Satellite and Lightning Data to Track Rapidly Developing Thunderstorms in Data Sparse Regions, Atmosphere 2017, 8 (4), 67;doi:10.3390/atmos8040067

"The outcomes of this study are very encouraging for other countries in Africa where convection and severe convection often occur and sophisticated data sources are absent. Initial studies over East Africa indicate that the RDT product can benefit operational practices for the nowcasting of severe convection events."

RDT produced by SAWS is now available on WMO RSMC (*Regional Specialised Meteorological Center*) of Pretoria. Sixteen African countries of SWDFP (*Severe Weather Forecasting Demonstration Project*) from southern part of African continent can visualize real-time RDT.

3.8 CONCLUSION AND COMPLIANCE REQUIREMENTS

From a subjective point of view, the use of NWP data with RDT has allowed an improving gap of the discrimination efficiency. False alarms are lowered thanks to a "NWP convective mask" used as a guidance for the diagnosis, and precocity is increased with early diagnosis in warmest categories, thanks to a new tuning with NWP data and mask. The objective validation of GEN scheme over a wide region thanks to EUCLID data detailed in a previous report has confirmed this first analysis. It had been undertaken through various approaches from time step cell to the full life cycle of a cloud system, and taking into account the limitations of the ground truth. With a moderate ground truth

(defined by 5 flash impacts at least during a trajectory) and non convective trajectories defined by being away from flashes of more than about 200km, satisfying skills are reached for full-trajectory approach: POD of 74% together with 2% POFD, FAR 22% and a TS of 61%. Scores are even better when considering sections of trajectories or cloud cells individually. RDT keeps good performances when taking into account intermediate season period (Spring, Autumn). Of course RDT scores are better for summer. Moreover, the skills obtained with EUCLID data over Europe are better in all configurations and for all approaches than for the previous validation. This improvement does not appear so clearly concerning the precocity of RDT GEN discrimination. It is limited to systems which are able to be early discriminated, i.e. with isolated convective system depicted from low levels. **Finally, those results fulfil the target accuracy requirements (see 1.2) over a large domain and for an extended period, i.e. 70% of detection and 25% of convective systems diagnosed before lightning activity.**

RDT-CW Calibration discrimination scheme implemented for several geostationary satellites has been tuned again for this release, over several months, taking also advantage of a stricter filter thanks to NWP data, and availability of reliable lightning data for operational runs (Meteorage network, but also GLD360 and GLM-GOES16).

Subjective cases studies have illustrated the improvement, and objective scoreshave met those from generic scheme. Moreover, those improvements are applicable to various geographical regions, and most of geostationary satellites.

We consider nevertheless that there is still room to improve the false alarms, the number of miss cases and the early diagnosis. Also, improvements remain necessary over oceanic regions, where the signature of convective systems differ from continental regions.

RDT provides an accurate depiction of convective phenomena, from triggering phase to mature stage. The RDT object allows pointing out some areas of interest of a satellite image. It provides relevant information on triggering and development clouds and on mature systems. Even if the precocity on the first lightning occurrence remains to be improved, the subjective evaluation confirmed the precocity usefulness on moderate lightning activity.

Those good results consolidate the status of RDT which had been set up to "operational" by EUMETSAT (since v2011).

Subjective validation exhibits very good results of the algorithm concerning OTD. It is a major point to improve RDT by focusing on the areas of the most severe and intense convection. This is not fully confirmed by an objective validation versus the CMHI overshooting top database made available by a Convection Working Group. Against this expert-based approach in a super rapid-scan context (2,5 minutes), RDT-CW OTD reveals as expected numerous misses, but also lower detection efficiency, and significant false alarms. A deeper use of high resolution will be necessary to improve this point.

Despite this, depending on cloud system morphology, RDT is able to present a kind of multidimensional description of convective systems.

The lightning jump algorithm, implemented in previous release, has been compared to hail and hazards detection systems on some situations. A good correlation between those events and lightning jump detection has been observed. An operational monitoring of this attribute has then led to a tightening of the criteria to limit an excessive number of diagnosis. An objective assessment is foreseen in a next scientific report.

It completes the data fusion approach with other products of NWCSAF.

4   CONCLUSION

CI and RDT-CW help to follow convection in different stage. There are compliant with MSG, GOES- 16/17 and Himawari-8/9 satellites.

This release of CI offers various improvements of the product: more precise use of Cloud products (cloud type and cloud top micro-physics), stricter definition of the areas of pixels of interest, day-time and night-time tunings. This improvement reflects on subjective and the objective validation exhibit scores at the boundary of the requirements. The use of CI in conjunction with RDT-CW offer to end-users reliable tools to assess various phases of convection.

RDT-CW is a mature product with several years of continuous development and improvement and several version operational. A new tuning of the calibrated discrimination scheme confirms the operational capability of RDT-CW to generate convection warning product from geostationary satellite data whatever the geographical region. Additionally the forecast scheme takes benefit from slight improvements in the movement estimation of cloud cells such that the nowcasting capabilities of the RDT-CW up to +1 hour can be made possible.