Characterization of convective systems dynamics and microphysics using time-delayed tandem microwave radiometers

Thomas Lefebvre, Hélène Brogniez, Ilhem Gharbi, Laura Hermozo, Dominique Bouniol, Florent Dralet, Rémy Roca

Characterization of convective systems dynamics and microphysics using time-delayed tandem microwave radiometers

Thomas Lefebvre*[1], Hélène Brogniez[1,2], Ilhem Gharbi[3], Laura Hermozo[4], Dominique Bouniol[5], Florent Dralet[6], and Rémy Roca[7]

[1]LATMOS, Université Paris-Saclay/UVSQ/CNRS/IPSL, Guyancourt, France
[2]LSCE, Université Paris-Saclay/UVSQ/CEA/CNRS/IPSL, Gif-sur-Yvette, France
[3]GIPSA-Lab, Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, Grenoble, France
[4]CNES, Toulouse, France
[5]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[6]Université Paris-Saclay, ENS Paris-Saclay, DER de Physique, Gif-sur-Yvette, France
[7]Université de Toulouse, Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (CNRS/CNES/IRD/UPS), Toulouse, France

Abstract: Deep convective cloud systems are central to the global water and energy cycle, and yet their representation in climate models remains challenging. This study explores the potential of machine learning to classify and characterize cloud structures inside cloud systems using radiometric measurements from the C²OMODO (Convective Core Observation through MicrOwave Derivative in the trOpics) mission. The gradient boosting algorithm is used to classify clouds into four types: anvil, stratiform, convective, and deep convective, and achieves high performances. Furthermore, retrievals of both dynamical and microphysical quantities are shown to perform well. The classification method has also been shown to significantly improve the performance of geophysical variable retrieval. This study highlights the potential of the forthcoming C²OMODO mission in advancing our understanding of convective systems.

Keywords: Classification and regression models, Cloud detection, Convective systems, Deep convection, Gradient Boosting Decision Tree (GBDT), Microwave radiometers, Time-derivative measurements.

1 Introduction

Deep convection transports mass, water and momentum from the lower troposphere to the tropopause, generating cumulonimbus clouds that are at the heart of the water and energy cycle [1]. These small clouds can be further organized in a much larger deep-cloud system (DCS). DCSs extend for hundreds of kilometers horizontally and persist for several hours, producing more than 50% of precipitation in the tropics [2–6]. The largest and longest-lived of these systems are also strongly related to extreme precipitation events [7]. Their importance is not only restricted to the water cycle, as their extended ice cloud decks strongly interact with the Earth radiation budget [8].

These organized convective systems are not well represented in climate models, where convective parametrization traditionally fails to incorporate such scales. Emerging km-scale models, used both for climate and meteorology purposes, while successfully able to resolve such mesoscale features, still suffer from strong biases in terms of morphology and duration [9–11]. This could be linked to some known biases in the representation of deep convection

*Corresponding author: thomas.lefebvre@latmos.ipsl.fr

for which strong convective updrafts intensity is generally underestimated [12–16].

To better constrain these models, relevant observations at both the scale of the overall system and the finer scale of deep convective features are needed. Furthermore, the wide diversity of DCS encountered across the globe [17] requires a global perspective. Dedicated satellite observations have demonstrated their ability to support cloud and precipitation research [18]. More specifically space borne radars and passive microwave radiometers in combination with geostationary observations have been extensively used to explore the physics of DCS [19–23]. Yet the dynamical features of deep convection remain difficult to access from space. The recent advent of nadir looking Doppler radar measurements from the EarthCARE mission [24], provide a first exploration into convective dynamics that could be completed in the future by a more systematic measure of in-cloud winds [25]. Convoy of radars are also being considered as a way to access the short-term deep convection dynamics as anticipated by the INCUS (Investigation of Convective Updrafts) mission [26, 27].

In the same philosophy, the C²OMODO (Convective Core Observation through MicrOwave Derivative in the trOpics) mission aims to investigate dynamical properties of deep convection but with a larger swath than INCUS, to observe the horizontal extension of convective systems [28]. Proposed by CNES as a contribution to the NASA AOS program scheduled for 2029, the C²OMODO mission focuses on the vertical updraft of deep convective cells. It involves the use of two identical passive microwave radiometers positioned less than 3 minutes apart to observe the temporal evolution of cloud systems over a short period of time.

A study by Auguste and Chaboureau (2022) [29] demonstrates the informational content of C²OMODO measurements. The objective of the present study is to highlight the capacity of the C²OMODO tandem to detect, attribute, and differentiate various parts of convective systems. Classifications based on data from a single radiometer are performed, as well as classifications that consider time-derivative measurements provided by the tandem. Moreover, integrating the characterization of different cloud components into the retrieval processes of geophysical variables (GVs) allows the regression models to be better adapted to these cloud classes. Indeed, cloud properties, particularly vertical motions within updrafts, vary significantly within a deep convective system [30, 31]. It is expected that the relationships between microphysical properties (ice content, production and dissipation) and dynamical properties (ice mass flux) will vary based on the cloud component.

This work is organized as follows. The first part deals with the data used, describing the characteristics of C²OMODO radiometers, the classification criteria for different parts of convective systems, the GVs, and the numerical simulations using the radiative transfer model and the cloud-resolving model. The second part presents the machine learning methods used for the classification and the regression of a subset of GVs targeted. It is not intended here to perform an exhaustive overview of the GVs that are under development for C²OMODO, but rather to use some of them as tools to highlight the added value of the short-time delayed measurements for the detection and classification of convection classification. The third part of the paper is devoted to results and discussion, including an analysis of the performance of both classification and retrieval models. The paper concludes with a summary of the results obtained and the foreseen works on this topic.

2 DATA

2.1 Characteristics of the C²OMODO tandem

The C²OMODO mission uses a temporal differentiation method with twin passive microwave radiometers, separated by a short time delay (dt) of approximately 1-3 minutes. This delay, defined based on targeted dynamical properties, aims to minimize contributions from horizontal advection between observations [28]. Even though few dt values were tested, the results in this study are presented only for a dt of 180 seconds due to its minimal impact on the outcomes. The underlying idea behind C²OMODO is to interpret short-term brightness temperature (Tb) variations from the radiometers as changes in geophysical variables (GVs) tied to dynamic information.

Each radiometer of the C²OMODO tandem, which are strictly identical, inherits from the SAPHIR humidity sounder on-board the Megha-Tropiques mission [32]. They benefit from enhancements to improve radiometric performance and adapt the spectral range of observations. This upgraded instrument will sample water vapor absorption lines at 183.31 GHz and 325.15 GHz. A window channel centered around 89 GHz, that benefits

from the ESA AWS (Arctic Weather Satellite) EPS (EUMETSAT Polar System) Sterna radiometer, and the ESA/EUMETSAT MetOp-SG MWS (MicroWave Sounder), is also present to better constrain the information on precipitation [3, 33–35].

Table 1 summarizes the spectral configuration of the microwave radiometer (hereafter MWR) used in the present study. This configuration closely approximates the final design of the microwave radiometer, which had not yet been finalized at the time of the study. This configuration is based on SAPHIR, with six double sideband channels at 183 GHz and similar sampling at 325 GHz. Gaussian noises have been added to each channel based on the radiometric sensitivity values provided in Table 1 defined by preliminary studies by CNES and ADS on the instrument.

Table 1: Description of spectral configuration of the microwave radiometer (MWR) used in this study.
NEDT: Noise Equivalent Delta Temperature;
IFOV: Instantaneous Field of View.

Central frequency (GHz)	Channel (GHz)	Bandwidth (MHz)	NEDT (K) 183 | 325	Targeted IFOV (km)
183	±0.2	200	2.73 | 5.9	≤ 10
	±1.1	350	2.06 | 4.4	5 at nadir
	±2.8	500	1.73 | 3.9	
325	±4.2	700	1.46 | 3.2	≤ 6
	±6.8	1200	1.11 | 2.5	3 at nadir
	±11.0	2000	0.86 | 1.9	
89	89	4000	0.5	≤ 20
				10 at nadir

The 183 GHz and 325 GHz measurements are sensitive to the ice hydrometeors present in clouds primarily through scattering processes [36–39]. Sensitivity to various ice crystal sizes is achieved with both bands: 183 GHz responds to larger icy hydrometeors while 325 GHz is more sensitive to smaller crystals. Hence, measurements at 183 GHz are commonly employed to detect deep convection as well as associated overshooting phenomena [40, 41]. Moreover, these frequencies capture radiation from different altitudes, and multiple observation channels per band provides information on the vertical distribution of hydrometeors [42–44].

2.2 Cloud resolving simulations

Cloud resolving numerical simulations of DCSs are performed using the Meso-NH model [45]. This numerical model is extensively used to investigate cloud and precipitation characteristics leveraging satellite observations [40, 46, 47].

The Meso-NH configuration used here is the same as in Auguste and Chaboureau (2022) [29] and Brogniez et al (2022) [28], representing the deep convective cloud called "Hector" (hereafter HEC) that develops over the Tiwi Islands, north of Darwin, Australia [48]. The simulation spans six hours with a 30 seconds timestep, covering a 256×200 km^2 area at 1 km horizontal resolution, extending to 30 km altitude with 146 vertical levels from 60 to 200 m below 22 km.

The initial conditions for these simulations are derived from Darwin radiosonde data at 00:00 UTC on November 30, 2005, as in Dauhut et al. (2015) [48]. Four additional simulations have been generated by modifying the initialization conditions to expand the available dataset. These adjustments involved increasing humidity by +10% and +20% under two scenarios: one with horizontal component of wind and one without, in order to boost or attenuate the strength of the convection.

2.3 Radiative transfer model for synthetic measurements

The COMODO-R measurements, expressed in Tb, are computed for each time step of the Meso-NH HEC simulations using the radiative transfer model RTTOV (Radiative Transfer for the Television and Infrared Observation Satellite - TIROS Operational Vertical Sounder), version 13 [49]. The size and shape distribution of hydrometeors are set as default as described in Geer et al (2021) [50], which could be questionable, but sensitivity studies on assumptions on the particle size distribution and shape of the hydrometeors are off the scope of the present study.

In the present experiments, the Tb are idealized and do not account for observation geometry (pixel resolution of each frequency, number of observations per scan, viewing angle, pixel distortion according to these angles, and parallax effects). They are considered at nadir at the top of the atmosphere for each column of the Meso-NH simulation, with the same resolution as the cloud simulations. A future sensitivity study will address the impact of observation geometry and instrumental effects on inversion methods.

2.4 Classification of the structure of convective systems

Following Feng et al (2011) [51] and Marinescu et al (2016) [30], the atmospheric columns of Meso-NH simulations are categorized into classes using a physical threshold partitioning whose criteria are given in Table 2: clear-sky (CS), anvil (ANVL), stratiform (STRAT), convective (CONV), and deep convective (DC). The partitioning method considers: the precipitation rates (PR), the vertical velocity of air (W_{air}), and the cloud top altitude (Z_{top}) defined as the highest altitude where the mixing ratio of ice (r_{ice}) exceeds 0.1 g.kg^{-1}. An IR window channel, here simulated at 11.5 µm based on the MVIRI sensor from METEOSAT-7, is used to detect ANVL regions. Data from the geostationary satellites on operation at the time of the C²OMODO mission will provide such information. For CONV and DC classes, criteria 2 and 3 in Table 2 require that W_{air} exceed thresholds at any altitude above freezing (criterion 2) or below freezing (criterion 3).

Table 2: Classification criterion [30] for individual atmospheric column. PR: Total Precipitation rate; W_{air} (z): Vertical velocity of air; T: Air temperature; $Tb_{11.5\mu m}$: Tb in the $11\mu m$ window; Z_{top}: Altitude of cloud tops.

Classes	Criteria
CONV & DC	• At least one of the following: 1. PR $> 25\ mm.h^{-1}$ 2. $\lvert W_{\text{air}}(z)\rvert$ (T > 0°C) $> 3\ m.s^{-1}$ 3. $\lvert W_{\text{air}}(z)\rvert$ (T ≤ 0°C) $> 5\ m.s^{-1}$ --- • CONV: $Z_{\text{top}} > 6\ km$ • DC: $Z_{\text{top}} > 12\ km$
STRAT	• PR $> 0.1\ mm.h^{-1}$ • $Z_{\text{top}} > 6\ km$
ANVL	• $T_{\text{b,11.5 µm}} < 270\ K$ • $Z_{\text{top}} > 6\ km$
CS	None of the previous criteria are met

Figure 1 illustrates the application of the classification criteria on a longitudinal transect of an HEC cloud

simulation, (as described in Section II.B). This figure also displays the Tb and dTb/dt of C²OMODO along the transect, in addition to the three GVs presented in Section II.E. This example demonstrates the relationship between C²OMODO radiometric measurements (Figure 1(d)) and cloud structure (Figure 1(a), (b) and (c)). The cloudy regions show a strong depression in Tb, especially in the convective classes (CONV and DC), due to the attenuation of radiation from ice scattering.

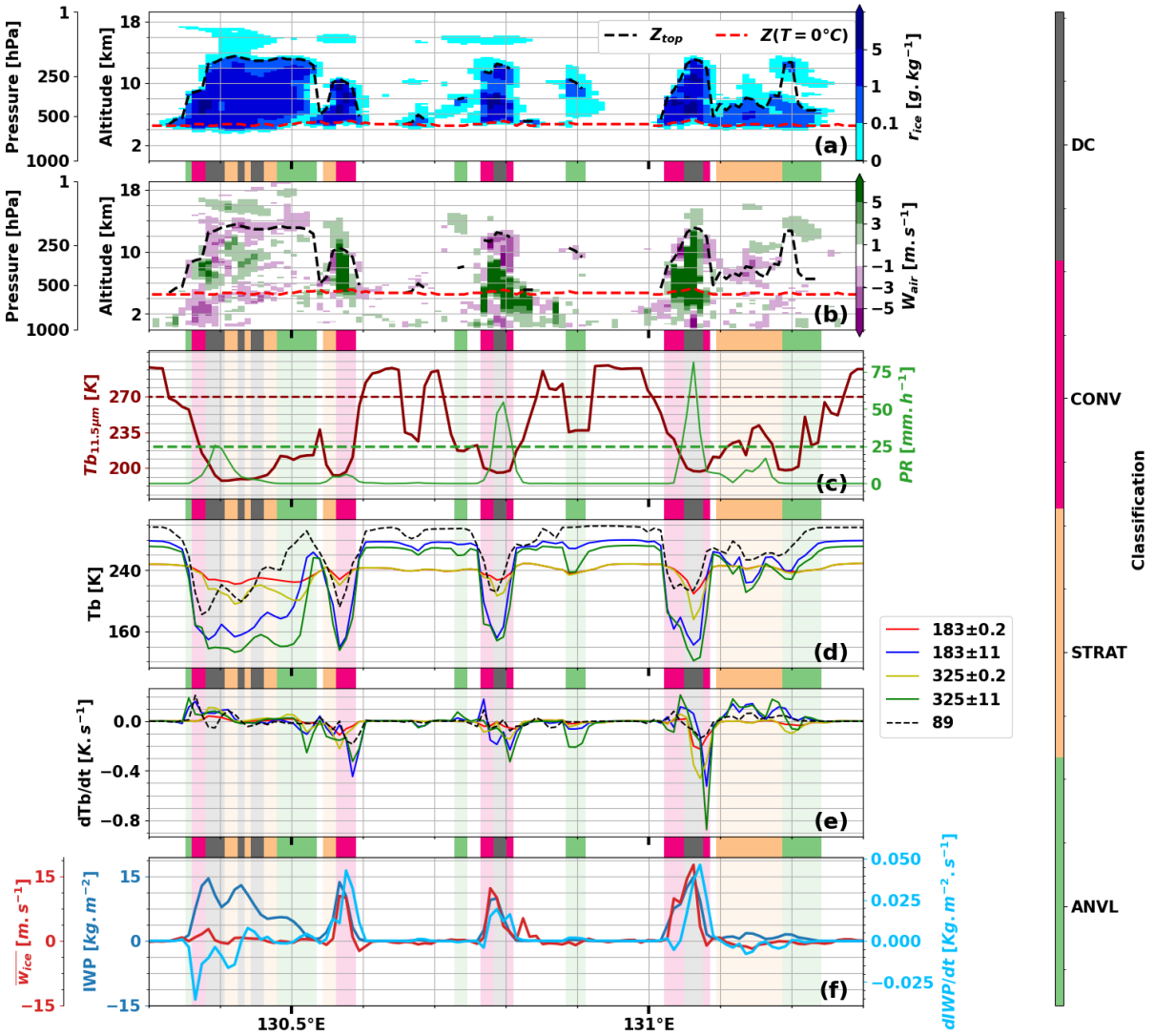

Figure 1: Application of classification criteria (color bar on the right-hand side) on a longitudinal transect of the HEC cloud simulation. (a): Mixing ratio of ice, r_{ice}; (b) Vertical wind velocity, W_{air}; (c): Infrared Tb at 11.5µm in red and Precipitation Rate, PR, in green; (d) Tb of 5 C²OMODO channels; (e): dTb/dt of the same 5 channels; (f): IWP in blue, $\overline{w_{\text{ice}}}$ in red, dIWP/dt in cyan (GVs defined in Section II.E).

Tb do not vary in the same manner depending on the observation channel. Comparing Tb at 183 GHz and 325 GHz in clear-sky and cloudy conditions, we see that Tb decrease slightly near the center of the absorption line (± 0.2 GHz), whereas Tb decrease significantly in the presence of ice particles at the wing of the line (± 11 GHz). This is due to the wings of the 183/325 GHz channels being associated with lower atmospheric opacity related to water vapor, making these channels more sensitive than those at the line centers. The 325 GHz tend to shown an even greater Tb decrease than 183 GHz, which is particularly apparent in the section of ANVL at 130.9°E. Figure 1 demonstrates the good agreement of the present classification with this former study.

Figure 2 shows, for the entire set of HEC, the mean profiles and standard deviations of condensed water phases (liquid and ice) and the vertical component of atmospheric velocity for the different identified classes. ANVL contains a small amount of condensed water, almost entirely in solid form, with vertical velocities close to zero. This is consistent with high-altitude clouds maintained aloft by minimal vertical motion, which explains their buoyancy [31].

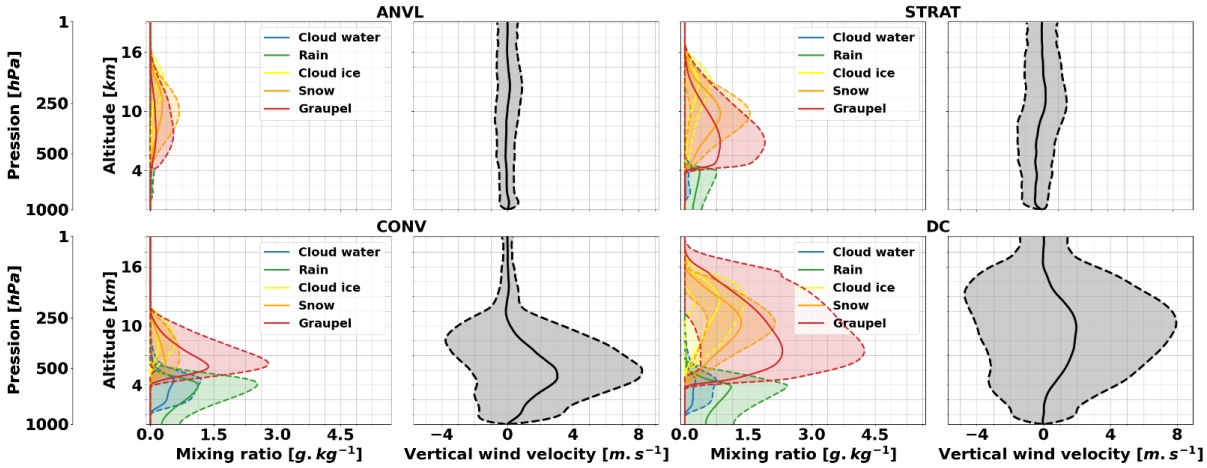

Figure 2: Mean vertical profiles and standard deviations of mixing ratios of water species and vertical wind velocities for each cloudy class in the entire set of HEC simulations.

The STRAT class shows a higher amount of ice, a small precipitation rate, and small vertical velocities. In contrast, CONV and DC classes exhibit much stronger vertical velocities and intense precipitation. The CONV class water content extends to lower altitudes (below 12km) compared to the DC class, probably corresponding to convective cells in the growth phase or convective cells with lower tops. On the other hand, the DC class ice content shows a much greater vertical extension, indicative of more fully developed convective regions. Both classes show near cloud top higher amount of graupel with respect to other species.

These profiles agree well with the common definitions found in the literature [30, 31, 52–57]: the anvil is characterized by air detrained from the upper part of the convective regions, which remains at high altitude due to its neutral buoyancy. Stratiform regions are defined as cloudy regions containing both liquid and frozen water, with low precipitation rates and slow cloud rise. Convective regions of an individual thunderstorm contain precipitation and vertical velocities, positively or negatively, much stronger than those in the surrounding regions such as the anvil and stratiform regions [30, 31].

2.5 Geophysical variables related to the production of ice

Ice Water Path (IWP, in kg.m^{-2}) is defined as:

$$\mathrm{IWP} = \int_z \rho(z)\, r_{\mathrm{ice}}(z)\, dz \tag{1}$$

with ρ being the density of dry air (kg.m^{-3}), r_{ice} the mixing ratio of total ice water (kg.kg^{-1}), and z the altitude (m).

The short time delay between successive measurements of C²OMODO can be used to infer the rate of change of IWP over time, dIWP/dt (in kg.m^{-2}.s^{-1}), expressed as:

$$\frac{d\mathrm{IWP}}{dt} = \frac{\mathrm{IWP}_{t_f} - \mathrm{IWP}_{t_i}}{t_f - t_i} \tag{2}$$

where t_i is the time of measurement of the leading radiometer and t_f the time of the trailing radiometer.

Moreover, the in-cloud vertical velocity of ice is estimated following Auguste and Chaboureau (2022) [29], who proposed an integrated variable $\overline{w_{\mathrm{ice}}}$ that can be related to the temporal variability of Tb, and defined as:

$$\overline{w_{\mathrm{ice}}} = \frac{\int_z W_{air}(z)\, \rho(z)\, r_{\mathrm{ice}}(z)\, dz}{\int_z \rho(z)\, r_{\mathrm{ice}}(z)\, dz} \tag{3}$$

where $W_{air}(z)$ is the vertical wind velocity (m.s^{-1}). Thus, $\overline{w_{\mathrm{ice}}}$ represents the averaged vertical velocity of ice particles within the cloudy atmospheric columns.

These three variables serve as benchmark to assess the performance of inversion methods for specific scene properties. Figure 1(d), (e) and (f) illustrate the relationship between Tb and IWP on one hand, and the consistency between dTb/dt, dIWP/dt, and $\overline{w_{\text{ice}}}$ on the other hand. As previously mentioned, Tb are lowest when IWP is high and increase as IWP decreases. Peaks in dIWP/dt correspond to opposite-sign peaks in dTb/dt, showing strong relationship between these variables. High values of $\overline{w_{\text{ice}}}$ appear to coincide with low Tb and non-zero dTb/dt regions. Comparing Figure 1(a) and Figure 1(b) with Figure 1(e), it is shows that high $\overline{w_{\text{ice}}}$ values correspond to regions where both the IWP and the vertical wind velocities are both significant.

While Figure 1 shows the variation of GVs accordingly to Tb and dTb/dt along a transect of one HEC simulation, Figure 3 generalizes these relationships for an entire HEC simulations (spanning 6 hours with a 30-second temporal resolution over a 256 × 200 km2 domain). Figure 3(a) shows the relationships between Tb, at the center and wings of the 183 GHz line and IWP across cloud classes. Tb decrease almost linearly as IWP increases, with a stabilization beyond 45 kg.m^{-2} for both channels due to a saturation effect. All classes exhibit a similar relationship, except CONV at the center of the 183 GHz line, which shows a less pronounced decrease in Tb at the center of the 183 GHz line but not at the 183 GHz wing.

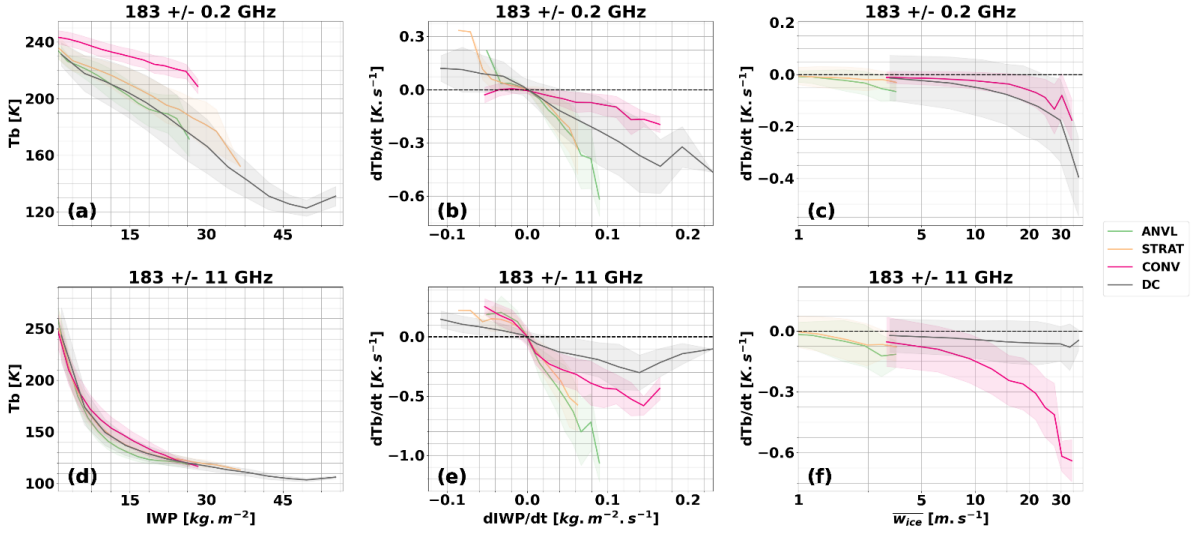

Figure 3: Relationships between Tb and IWP (a)-(d), dTb/dt and dIWP/dt (b)-(e), dTb/dt and $\overline{w_{\text{ice}}}$ (c)-(f) for channels at the center (a,b,c) and wings (d)-(e)-(f) of the 183 GHz line for each of the cloud classes.

The distinct Tb-IWP behavior observed in the CONV class can be attributed to the channel's sensitivity to higher altitudes due to the water vapor absorption band. The central channel of the 183 GHz line is opaque to lower altitudes, as radiation is absorbed by water vapor. As shown in Figure 2, the CONV class includes areas of intense precipitation and significant ice content below 12 km, resulting in a higher ice concentration at lower cloud tops, even for thicker ice layers. Consequently, a large part of the ice layer is positioned at altitudes to which the channel is less sensitive, resulting in reduced radiation attenuation by ice scattering and thus higher Tb values, even at relatively high IWP, compared to other cloud classes.

This is further detailed in Supplementary figure, which provides a deeper analysis of Figure 3(a) and Figure 3(d), with colors representing the cloud top altitude. In this figure, it is shown that the CONV class corresponds to lower cloud top altitudes, even for the highest IWP values.

As depicted in Figure 3(b), (c), (e), and (f), the relationship between GVs (dIWP/dt and $\overline{w_{\text{ice}}}$) and dTb/dt varies across cloud classes and channels of the 183 GHz absorption line. Firstly, the ANVL and STRAT classes exhibit strongly negative dTb/dt values for low dIWP/dt, likely due to significant horizontal detrainment, while they generally show low $\overline{w_{\text{ice}}}$ with dTb/dt close to zero. In contrast, CONV and DC classes show higher $\overline{w_{\text{ice}}}$ values associated with significantly negative dTb/dt, reflecting dynamical processes or cloud top elevation.

The relationships between GVs and dTb/dt also differ between the central and wing channels of the 183 GHz line. At ±0.2 GHz, CONV and DC show similar behaviour, with $\overline{w_{\text{ice}}}$ corresponding to decreasing dTb/dt. However, at ±11 GHz, while CONV maintains this pattern, DC shows very high $\overline{w_{\text{ice}}}$ with near-zero dTb/dt.

This can be explained by a lower increase in hydrometeor concentration despite the presence of very high vertical velocities in the DC class for mature, quasi-stationary cases. It could also be due to a saturation effect when the ice water content becomes particularly high.

These varied relationships suggest that a tandem like C²OMODO will be efficient in providing the required measurements to retrieve the targeted variables, with a potential improvement if an initial step of cloud structure detection (CS, ANVL, STRAT, CONV, DC) is performed.

3 Classification and retrieval methods

In this study, the Gradient Boosting Decision Tree (GBDT) is used for both the cloud classification and the GV retrieval. This method belongs to the family of supervised learning methods. The general concept of supervised learning is briefly introduced, followed by a description of the GBDT method. The metrics used to assess the performance of the statistical models are outlined and presented in more detail in Appendix B.C.

3.1 Supervised learning

Supervised machine learning algorithms build predictive statistical models from labelled datasets by learning from input-output pairs, where inputs are features and outputs are target values. Through iterative adjustments of model parameters to minimize prediction errors, the algorithm refines its accuracy based on a predefined loss [58]. Supervised learning encompasses classification, where input data is assigned to predefined categories by identifying patterns that enable discrete class predictions, and regression, which predicts continuous values by modelling relationships between input features and target variables, as seen in linear or polynomial regression models.

3.2 The Gradient Boosting Decision Tree algorithm

The GBDT algorithm, an iterative learning technique based on decision trees, is widely used for its strong generalization abilities, making it popular across diverse machine learning applications [59–62]. It builds a sequence of decision trees, each correcting the errors of the previous one, gradually improving accuracy by predicting residuals. Although more computationally intensive, deep learning techniques like Convolutional Neural Networks (CNNs) have also proven effective for classification tasks. Liu et al. (2018) [63] uses CNNs for cloud classification, while Liu et al. (2020) [64] develops a Multi-Evidence and Multi-Modal Fusion Network. Similarly, Zhang et al. (2020) [65] combines CNNs with Recurrent Neural Networks (RNNs) to enhance classification performance. While gradient boosting is simpler and less resource-intensive than CNNs but does not achieve the same level of performance, it remains effective and is ideal for preliminary studies like this one.

This study employs Light Gradient Boosting Machine (LGBM) for robust classification and regression, leveraging its optimized implementation of GBDT [62]. LGBM's efficiency with large-scale datasets and high-dimensional features makes it ideal for cloud class prediction (Section II.D) and GVs retrieval (Section II.E). A key feature of GBDT-based classifiers is their probabilistic output, which provides valuable insights into prediction confidence across classes.

GBDT, and by extension LGBM, relies on several critical hyperparameters to control statistical model training, balancing complexity, preventing overfitting, and enhancing accuracy. This study emphasizes potential C²OMODO observation applications with basic tests rather than full parameter optimization. The adjusted hyperparameters, defined using performance metrics are provided in Appendix A.

The classification algorithm's performance is evaluated using Precision (P), Recall (R) also known as Hit Rate, and F1-score, calculated based on True and False predictions (see Appendix B). F1-score serves as the primary metric for comparing classification model performance, while Precision and Recall aid in fine-tuning LGBM hyperparameters, especially class weight (Appendix A). For regression analysis, key evaluation metrics, Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Bias Error (MBE), provide essential insights and are detailed in Appendix C.

4 Results and discussion

4.1 Setting of the statistical models

Various spectral configurations of the C²OMODO microwave radiometers (from Table 1) are tested as inputs to the statistical models to evaluate their impact in both the classification and the GVs retrieval. To avoid any misunderstandings in the following of the study, the tandem concept is named C²OMODO while the spectral configurations are named below for the remainder of the study:

- MWR-C1: 183 GHz channels (6 channels)
- MWR-C2: 325 GHz channels (6 channels)
- MWR-C3: 183 & 325 GHz channels (12 channels)
- MWR-C4: 183 & 325 & 89 GHz channels (13 channels)

Different delays (dt) between the head and tail radiometers have been tested, but not shown here, because there is a small impact of dt on the results. In this study, results are presented for dt=180s, except for instantaneous variables (classification and IWP) where dt=0s.

As outlined in Section II.B, five independent HEC simulations have been performed, each initialized under different conditions. The training dataset is derived from two simulations, with 10% of these points set aside for validation. The test dataset is created from two other simulations, while the fifth simulation provided data for the map and dispersion diagrams. Classification models are trained based on criteria applied to HEC simulations, serving as the reference, while regression models aimed to estimate GVs from Tb and dTb/dt. Both global and specific regression models (Section IV.D) are trained to evaluate the impact of classification on retrievals: global models aggregate data across all cloud classes, whereas specific models are trained per cloud class and combined in a second step.

The regression models take Tb and dTb/dt as inputs, with dTb/dt restricted to dt =180s. Figure 4 (left panel) displays the number of points per class in the test dataset at dt=0 s, showing similar distributions of each class in the training dataset, except for the CONV class which is underrepresented. Class weights (see Appendix A) are applied in Section IV.B.a to overcome these disparities. Figure 4 (right panel) illustrates the class-specific reduction of the number of points when dt > 0 s. Only points where initial and final classes (t_i and t_f) match are included in the analysis, avoiding rapid class transitions and potentially contributing to performance differences at various dt values in Section IV.D.

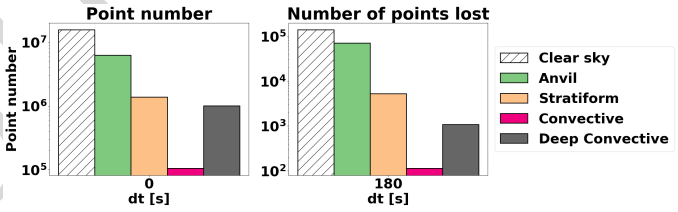

Figure 4: Distribution and reduction of training data Points. Left: Number of points in the training dataset for each class; Right: Reduction in the number of points for each class based on dt=180s.

4.2 Cloud classification using a single radiometer

4.2.1 Selection of class weight configuration

This section focuses on the selection of the appropriate class weight configuration. As explained in Section III.B, LGBM assigns different weights to classes to handle imbalanced datasets. As the dataset used is highly imbalanced (Figure 4), three tests were conducted to assess how different weight configurations impact classification performance.

Table 3 presents scores for classification models trained with different class weights. While F1-scores are similar, Precision and Recall vary with respect to class weights. The unweighted model has higher Precision but lower Recall, leading to poor identification of the underrepresented class. The balanced model achieves higher Recall but lower Precision, while the compromise weight configuration strikes a balance between False Positive (FP) and True Positive (TP) (see Appendix B).

Table 3: Precision (P), Recall (R), F1-score (F1) for each class prediction over 3 class-weight hyperparameter configurations. Results are expressed as percentage.

Class-weight options	Unweighted			Balanced			Compromise		
Score	P	R	F1	P	R	F1	P	R	F1
CS	92	97	94	94	94	94	92	97	95
ANVL	91	87	89	94	79	86	94	84	89
STRAT	53	54	53	51	50	51	48	64	55
CONV	49	52	50	61	47	53	36	70	48
DC	72	69	70	66	71	69	70	70	70

The choice of class weighting depends on the objective. If identifying the CONV class whenever it occurs is prioritized, the balanced model is preferred. Conversely, if ensuring accuracy when the CONV class is detected is the objective, especially as it is underrepresented, an unweighted model is more suitable. For this study, which aims to assess the feasibility and information content of C²OMODO tandem, the compromise model is selected.

4.2.2 Results of the compromise classification model

Table 4 shows the Precision (P), Recall (R) and F1-Score of the classification for each class across different spectral configurations when applying the compromise class weighting. Performance for MWR-C4 varies according to the class: CS and ANVL are consistently predicted with F1-score of 97% and 84%, respectively. The DC class is moderately well-predicted with a 70% F1-score, while STRAT and convective classes show a slightly higher prediction uncertainty. These trends remain for all spectral configurations, with slight variations in F1-score.

Table 4: Precision (P), Recall (R), F1-score (F1) for each class prediction over 4 spectral configurations. Results are expressed as percentages.

Config	MWR-C1			MWR-C2			MWR-C3			MWR-C4		
Score	P	R	F1	P	R	F1	P	R	F1	P	R	F1
CS	88	91	90	92	96	94	92	97	95	92	95	95
ANVL	86	80	83	88	82	85	91	84	87	94	89	89
STRAT	43	45	44	39	38	38	46	44	45	48	55	50
CONV	30	62	40	25	68	37	27	42	32	30	36	33
DC	66	69	68	63	68	66	70	69	70	70	70	70

Examining Precision and Recall separately reveals a clear disparity, with Precision generally lower than Recall across scores. For instance, Precision for the CONV class in the MWR-C4 is 36% compared to Recall of 70%, indicating a high rate of FP. While the DC class predictions are relatively accurate with few FP, the CONV class, though frequently predicted, suffers from more FP.

Spectral configurations impact performance, with the 183 GHz channel (MWR-C1) outperforming the 325 GHz (MWR-C2) for STRAT, CONV, and DC classes, whereas MWR-C2 performs better for CS and ANVL. Combining the 183 and 325 GHz channels (MWR-C3) slightly improves predictions, with optimal results achieved by adding the 89 GHz channel, particularly enhancing predictions for STRAT, CONV, and DC. In the remainder of this

section, we will take a closer look at the classification model of the spectral configuration using all the channels (MWR-C4).

The confusion matrix in Figure 5 details the classification model performance for MWR-C4, comparing true and predicted class labels, where each row represents the true class and each column the predicted class. CS and ANVL classes are generally well predicted (over 80% accuracy), whereas STRAT is often misclassified as ANVL (15%) or DC (17%). CONV is sometimes misclassified as ANVL (9%) or STRAT (13%), while DC is frequently misclassified as STRAT (22%). Misclassifications may arise from several sources. First, the dataset itself is imbalanced, skewing class weights during training. Second, radiometric measurements can be similar for classes with similar microphysical properties (see Figure 2 and Figure 3), complicating classification. In contrast, distinct classes like CS and ANVL exhibit fewer errors due to their clearly differentiated measurements.

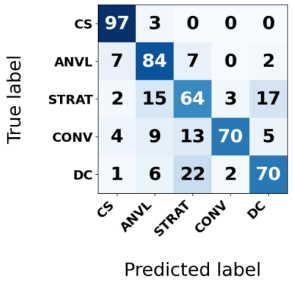

Figure 5: Confusion matrix of compromise classification model, expressed as percent. Diagonal are the correct prediction (TP), off-diagonal are the misclassifications.

Figure 6 shows a classification map comparing true (left panel) and predicted classes (right panel) for MWR-C4 at a specific simulation moment when the cloud system is well developed, with Z_{top} between 6 and 18 km, a maximum IWP of 50 kg.m^{-2}, and a horizontal extent exceeding 100 km. The map shows that the overall structure of the HEC cloud system, including ANVL, STRAT, and DC classes, is well captured, with isolated CONV cells also accurately predicted. However, distinguishing precisely between STRAT, CONV, and DC classes remains challenging.

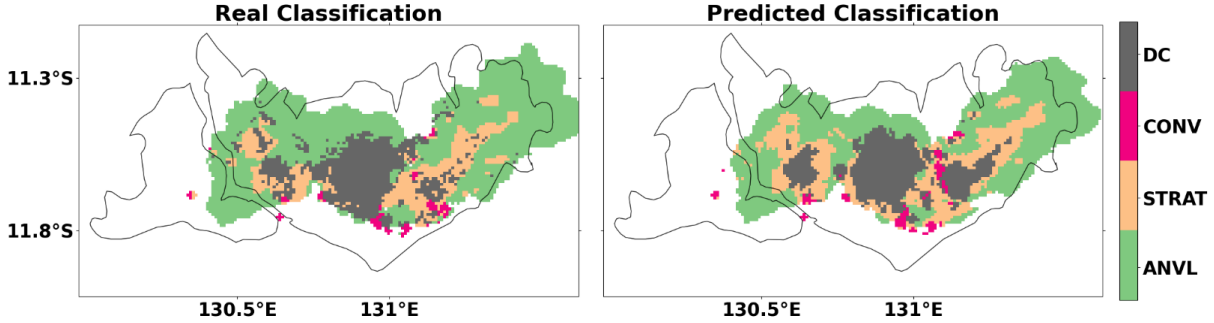

Figure 6: Classification map comparison between reference and predicted class for the MWR-C4 spectral configuration.

The partitioning method, partially based on dynamical information (see Table 2), results in classes with distinct vertical structures and dynamical properties (see Section II.D). Their temporal evolution depends on these characteristics. C²OMODO tandem measurements are expected to help retrieve both the classes evolution and their associated dynamics.

4.2.3 Analysis of class prediction probabilities

To investigate retrieval errors and uncertainties, it is important to examine prediction probabilities for each class, as shown in Figure 7. The classification model assigns probabilities to each grid point in the HEC simulation, selecting the class with the highest probability. Grid points with probabilities of 50% or higher are necessarily assigned to that class, although this does not always indicate high confidence in the classification.

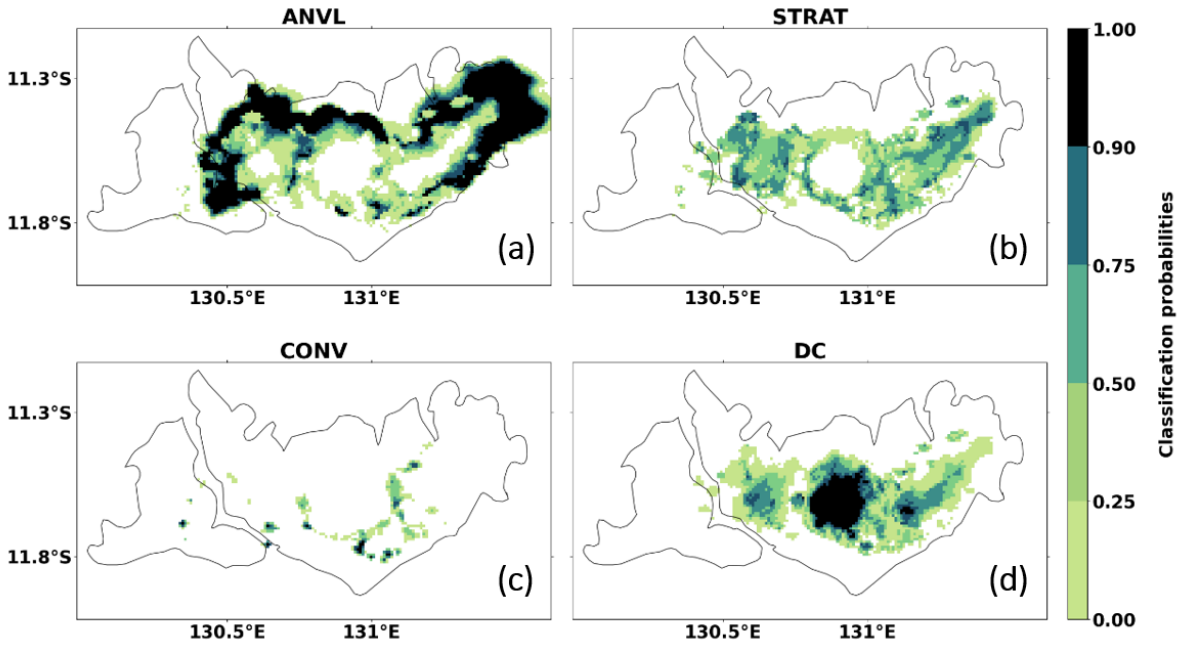

Figure 7: Map of classification prediction probability of (a) ANVL, (b) STRAT, (c) CONV, (d) DC.

Figure 7 displays a probability map for each class, highlighting variability, particularly for the STRAT class, where probabilities often fall below 80%, indicating uncertainty. In contrast, ANVL is predicted with high confidence, except near boundaries with other classes (see Figure 5 and Figure 6). Some grid points display probabilities above 25% for multiple classes, indicating the similar microphysical and radiometric characteristics mentioned in the previous section.

To better understand the classification model's errors, confusion matrices were computed across specific probability ranges, treating each cloud class separately. This approach enables to identify point that appear in multiple classes within the same probability range. Figure 8 illustrates this, where rows represent true classes and columns represent predicted classes, displaying the percentage of points within each probability range that belong to the true class. For example, 83% of the points predicted as CONV with a probability between 0.9 and 1 are correctly classified. The right panel shows the distribution of points for each class across probability ranges within the entire HEC simulation.

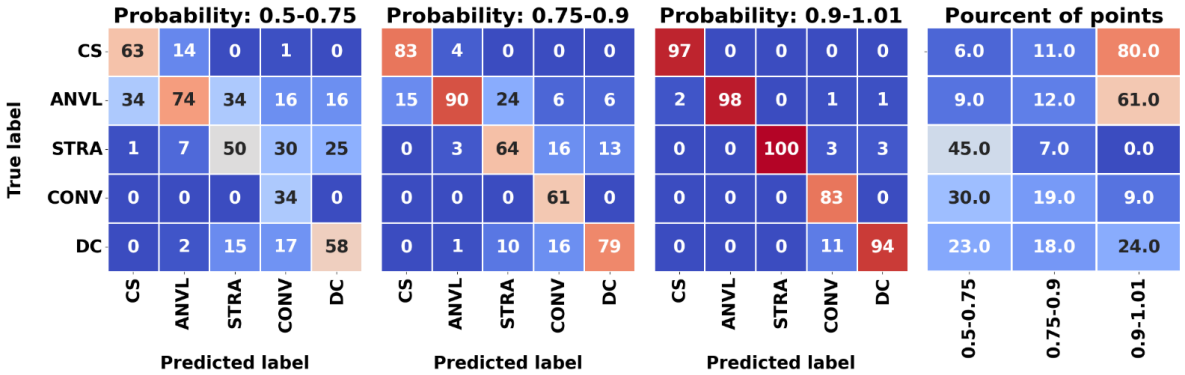

Figure 8: Confusion matrix for prediction probability bins (Three left panels). Percent of points of each true class within a probability bin (Right panel).

Confusion is notable in the 0.5-0.75 range, especially between CS and ANVL, and among STRAT, CONV, and DC classes. STRAT often overlaps with CONV and DC, and DC is similarly mistaken for STRAT and CONV. Errors decrease significantly in the 0.75-0.9 range. At the highest confidence level (0.9-1), errors are minimal, though 3% of points predicted as CONV actually belong to DC. CS and ANVL are confidently predicted, with

over 60% of their points in the 0.9-1 range, while DC has 24% and CONV only 9% at this confidence level. STRAT is rarely predicted with high confidence, having almost no points in the 0.9-1 range and only 7% in the 0.75-0.9 range, indicating limited confidence in STRAT predictions.

These promising results indicate that the MWR-C4 spectral configuration contains the necessary information to infer the physical processes defining each class. Furthermore, the dTb/dt provided by C²OMODO tandem enables short-term temporal cloud evolution analysis, which could improve the classification, especially near boundaries between different classes.

4.3 Ice variation phase classification using C²OMODO tandem

After evaluating the potential of cloud classification using single-radiometer data, C²OMODO tandem measurements, using MWR-C4 spectral configuration and dt=180s, are introduced and assessed for their classification capabilities. This approach, among other possible methods, extends classification to include both cloud classes and phases of ice variation. The model now categorizes 13 classes: CS, and ANVL, STRA, CONV, DC, each of the cloud classes subdivided into three ice variation phases: neutral (dIWP/dt = 0 $kg.m^{-2}.s^{-1}$), production (dIWP/dt > 0 $kg.m^{-2}.s^{-1}$), or dissipation (dIWP/dt < 0 $kg.m^{-2}.s^{-1}$). Table 5 presents F1-scores for MWR-C4 for dt=180s, with the first row is for each cloud class and phase, and the second row is for all ice variation phases, and across all cloud classes.

Table 5: F1-scores of the tandem classification model, for neutral, production and dissipation phases for each cloud class. The scores show both individual class performance and overall phase scores.

Phase	Neutral					Production				Dissipation			
Class	CS	ANVL	STRAT	CONV	DC	ANVL	STRAT	CONV	DC	ANVL	STRAT	CONV	DC
F1 per class	97	92	45	19	23	76	37	65	76	71	67	41	73
F1 per phase	98					89				90			

Table 5 shows that both ice variation phases and cloud classes are generally well predicted. F1-scores are high for neutral phases in CS and ANVL, which predominantly exhibit this phase. The STRAT class, with lower dIWP /dt values, achieves higher scores than CONV and DC, despite having fewer points in the neutral phase. Scenes identified in the production phase align closely with single-radiometer classifications, showing slightly lower scores for ANVL (≈70%) but higher scores for convective classes, particularly DC (>70%). Finally, scenes belonging to dissipation phase are also well predicted, with a large improvement of performance for the STRAT class (61% to 67% depending on the dt), although the identification of CONV cases slightly declines, as this class often corresponds to growing convective cells, hence better detection during production phases.

When focusing on ice variation phases alone (fourth line Table 5), the classification model achieves high F1-scores: 98% for neutral, 89% for dissipation, and 90% for production phases. This indicates that the C²OMODO tandem can effectively detects ice production phases, though some cloud class identification errors persist. Lower F1-scores for specific classes and phases result from combined errors in cloud class and ice phase classification.

Assessing the model's ability to detect cloud classes independently of ice production phases, Figure 9 shows significant improvements over single-radiometer classification (Section IV.B). TP increase for ANVL (from 84% to 89%), STRAT (from 64% to 71%), and DC (from 70% to 81%), with a notable reduction in FP and improved Precision. Specifically, STRAT is less frequently confused with DC (from 17% to 11%) and vice versa (from 22% to 13%).

4.4 Evaluation of regression model

Retrieval performances of three GVs: IWP, dIWP/dt and $\overline{w_{\text{ice}}}$ are assessed across the spectral configurations described in Section IV.A. Dynamic variables dIWP/dt and $\overline{w_{\text{ice}}}$ are evaluated at dt=180s, while IWP, as an instantaneous variable, is assessed at dt=0s.

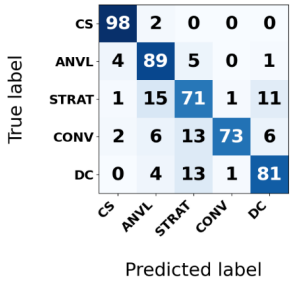

Figure 9: Confusion matrix of classification model of ice variation phase focusing only on cloud class detection, express as percent. Diagonal are the correct prediction (TP), off-diagonal are the misclassifications. For a given class, FP are the misclassifications on its column.

Table 6 presents retrieval scores for each GV across different spectral configurations. IWP shows a high R^2, indicating that the regression model is able to explain a large part of the IWP variability, alongside a low RMSE and negligible MBE, reflecting low error dispersion and minimal systematic error. The MWR-C2 offers no improvement. In contrast, the MWR-C4, the combination of 183, 325 and 89 GHz, yields higher R^2 and reduced RMSE, enhancing retrieval accuracy.

Table 6: R^2, RMSE, and MBE of retrieved GVs across 4 spectral configurations for dt=180s.

Variable	IWP ($kg.m^{-2}$)			dIWP/dt ($kg.m^{-2}.s^{-1}$)			$\overline{w_{ice}}$ ($m.s^{-1}$)		
Metric	R^2	RMSE	MBE	R^2	RMSE	MBE	R^2	RMSE	MBE
MWR-C1	0.95	0.84	7.3e-03	0.70	2.6e-3	6.96e-6	0.40	1.12	3.56e-4
MWR-C2	0.85	1.43	4.5e-03	0.64	3.66e-3	1.70e-5	0.28	1.22	7.99e-4
MWR-C3	0.95	0.81	8.2e-03	0.71	2.57e-3	4.81e-6	0.44	1.08	5.38e-5
MWR-C4	0.99	0.39	1.7e-03	0.91	1.46e-3	3.79e-6	0.56	0.97	-2.58e-3

For MWR-C1, the dIWP/dt regression model has R^2 of 0.70, significantly lower than that of IWP, with an RMSE of 2.62e-3 $kg.m^{-2}.s^{-1}$, indicating small errors, and a negligible MBE, showing no bias. MWR-C2 shows lower performance, and MWR-C3 does not improve results. However, MWR-C4 achieves a much higher R^2 of 0.91. The retrieval of $\overline{w_{ice}}$ also depends on the spectral configuration but with lower R^2, between 0.28 and 0.55, and an RMSE around 1 $m.s^{-1}$, suggesting higher errors, while the low MBE indicates no bias. MWR-C4 significantly improves the retrieval performances. Overall, 183 GHz shows better performance as it saturates less quickly with increasing ice content compared to 325 GHz.

The combination of 183 and 325 GHz channels (MWR-C3) provides a slight improvement in performance compared to MWR-C1, with an R^2 of 0.44 instead of 0.40 and a reduced RMSE. The inclusion of 325 GHz channels has a noticeable impact on the $\overline{w_{ice}}$ retrievals. Including the 89 GHz channel further enhances retrieval, as its sensitivity to precipitation helps identify regions with stronger positive (updrafts) and negative (downdrafts) vertical wind velocities.

Testing different time steps (not shown) reveals that performance slightly improves as dt increases, possibly due to the averaging effect, as extreme, short-lived events are less captured at larger dt. To explore whether classification can improve dIWP/dt and $\overline{w_{ice}}$ predictions, retrievals were performed with and without classification. Real classes were used to avoid propagation errors, which are not discussed here.

Table 7 compares retrieval scores for three GVs across two regression models: specific models trained on points from specific classes, and global model that ignores class distinctions. Specific models show that dIWP/dt predictions are relatively consistent across classes, though slightly lower for the DC class, likely due to a saturation effect in Tb at high ice content [28]. For $\overline{w_{ice}}$, there is a notable improvement in the CONV and DC classes, with R^2 of 0.67 and 0.80, compared to 0.17 and 0.21 in the ANVL and STRAT classes, confirming that retrievals of this dynamic variable are more accurate in convective zones with strong vertical wind velocities (updrafts or

downdrafts).

Table 7: R^2 metrics of regression models for 3 variables (IWP, $dIWP/dt$, $\overline{w_{\text{ice}}}$) across 4 cloudy classes for global model (trained for all point) or specific model (trained for each class).

Variable	IWP ($kg.m^{-2}$)		dIWP/dt ($kg.m^{-2}.s^{-1}$)		$\overline{w_{\text{ice}}}$ ($m.s^{-1}$)	
Model	Global	Specific	Global	Specific	Global	Specific
ANVL	0.99	0.99	0.92	0.93	0.05	0.17
STRAT	0.99	0.99	0.93	0.94	0.06	0.21
CONV	0.97	0.98	0.92	0.94	0.67	0.80
DC	0.97	0.98	0.89	0.90	0.56	0.59

Another key aspect is the comparison between specific and global models. Table 7 highlights improved performance with specific models, particularly for $\overline{w_{\text{ice}}}$, with R^2 increases of 0.12, 0.15, 0.13, and 0.03 for the ANVL, STRAT, CONV, and DC classes, respectively. Given that GBDT is a non-linear inversion model based on decision trees, it could be assumed that the global model might capture the relationships between Tb and GVs across classes. However, specific models perform better, likely because each class has partially different Tb-GVs relationships, reducing retrieval uncertainty when points are class-separated.

For MWR-C4, R^2 for the global model is around 0.55 (see Table 6). Table 7 shows that both global and specific models explain $\overline{w_{\text{ice}}}$ variability better in the CONV class, with R^2 of 0.67 (global) and 0.80 (specific). The DC class has R^2 of 0.59, while ANVL and STRAT classes fall below 0.20, suggesting challenges in retrieving accurate values for these classes, whose values are close to zero (see Figure 2), resulting in small retrieval errors that significantly affect performance.

Figure 10 displays scatter plots of retrieved as a function of reference (true) GVs for global and specific models. Both models accurately reproduce IWP (Figure 10(a) and (d)) and dIWP/dt (Figure 10(b) and (e)), though the global model misses higher IWP values (>40 $kg.m^{-2}$), primarily in the DC class, which the specific model reproduces well without the observed saturation effect. This saturation in the global model likely stems from data imbalance, as fewer data points with high IWP values limit the regression model's ability to fully express this range of IWP. Additionally, specific models help differentiate points with similar Tb-GVs relationships, reducing retrieval uncertainties. Points with lower IWP values (<40 $kg.m^{-2}$) are more abundant, giving them more weight in training, which leads to better retrieval of lower values at the expense of higher ones. The prior classification enables the models to specialize for specific value ranges, enhancing their performance in distinct regimes.

Figure 10(c) and (f) show that the scatter plot points are well centered around the diagonal line where Prediction equals Reference. However, there is noticeable dispersion, with a denser cluster of points around values close to zero. Additionally, while the difference between global and specific models is subtle, it becomes apparent when comparing the R^2 values of 0.61 and 0.65, showing a slight improvement in accuracy for the specific model. Overall, the regression models provide a good agreement in reproducing $\overline{w_{\text{ice}}}$.

5 Conclusion and perspective

Deep convection, which generates clouds central to the water and energy cycles, has a significant impact on Earth's meteorological systems but remains poorly understood due to limited observations. This lack of understanding affects its representation in weather and climate models. The C^2OMODO mission, a CNES contribution to the AOS mission, aims to address this observational gap by studying storm dynamics through innovative radiometric observations. This paper evaluates the potential and information content of the future C^2OMODO observations.

We propose a cloud classification method and a geophysical variable regression method based on a gradient boosting algorithm. These methods uses four distinct spectral configurations, incorporating radiometric channels at 183 GHz, 325 GHz, 183+325 GHz, and 183+325+89 GHz. The final configuration, 183+325+89 GHz, closely

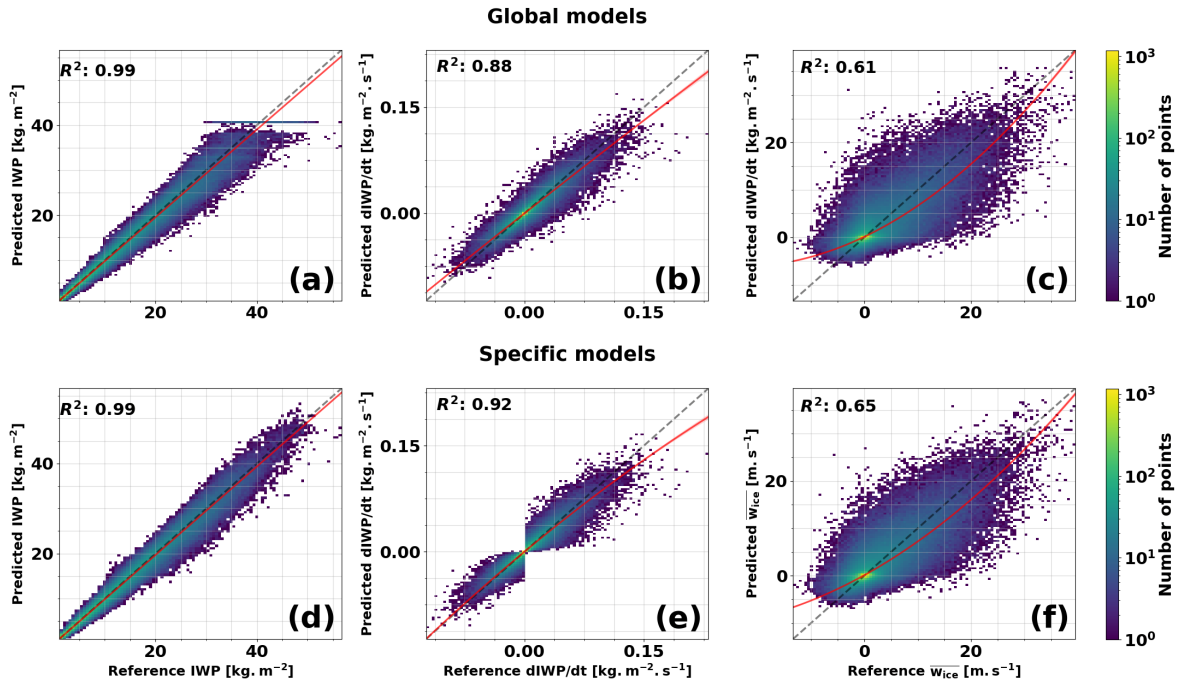

Figure 10: Scatter plots of retrieved IWP (a)-(d), dIWP/dt (b)-(e) and $\overline{w_{\text{ice}}}$ (c)-(f) as a function of true values for global (a)-(b)-(c) and specific (d)-(e)-(f) regression models. Predictions are obtained for MWR-C4.

approximates the finalized spectral configuration fixed for C²OMODO mission. Reality-like cloud simulations were performed using MESO-NH, and synthetic radiometric measurements were generated with RTTOV. The classification method efficiently categorizes convective system structure into four categories: anvil, stratiform, convective, and deep convective; using brightness temperature data from a single radiometer. The model achieves high performance, particularly for clear sky and anvil classes, and maintains good accuracy for convective and deep convective classes with slightly lower performance for the stratiform class. Misclassifications can be explained by the similarities in microphysical and radiometric properties between stratiform, convective and deep convective classes.

The use of time-derivative brightness temperatures allows for highly effective detection ice variation phases with high performances. This approach significantly improves the classification of the four cloud system structure classes compared to classification using a single radiometer. The regression method demonstrates excellent performance, with 99% recovery for IWP, up to 94% for dIWP/dt and up to 80% for $\overline{w_{\text{ice}}}$; providing vertically integrated insights on ice water content and vertical velocities within convective systems. Prior cloud classification enhances retrieval accuracy by better constraining relationships between Tb and different geophysical variable ranges. For extreme IWP values, particularly in deep convective clouds, retrieval performance improves by overcoming saturation effects. Additionally, dIWP/dt retrievals are refined near-zero values across all classes, while $\overline{w_{\text{ice}}}$ retrievals improve most notably for convective and deep convective classes, which exhibit strong vertical dynamics.

Overall, the MWR-C4 spectral configuration using 183, 325, and 89 GHz channels shows higher performance. The 183 GHz channels alone outperform the 325 GHz channel, and combining 183 and 325 GHz without the 89 GHz channel does not offer significant performance gains. The limited improvement from the 325 GHz channel can be attributed to the retrieval of integrated variables, which does not fully exploit its capability to provide vertically resolved information, similar to 183 GHz, using atmospheric sounding principle. Combining these channels could be valuable for retrieving vertical profiles of ice water content and vertical wind velocities, which may require alternative methods such as a variational approach (1D-VAR). This will be explored in future work, with preliminary results already showing potential.

These findings highlight the potential of C²OMODO for both structural classification and microphysical and dynamical properties retrievals of DCSs. They enable enhanced characterization of convective updraft intensity at the scale of deep convective features, with a global perspective along the tropics. Such results offer promising

insights for meteorology and climatology, contributing both to forecast improvements through data assimilation and to a more profound understanding of convective processes. However, dataset limitations must be considered in evaluating model performance. Statistical model errors and uncertainties may stem from under-representation of certain cloud classes or high IWP ranges. The MESO-NH simulations of Hector represent strong deep convective thunderstorms, which may not fully capture the diversity of all convective systems in various environments. Results are also shaped by MESO-NH physical parametrizations and its one-moment microphysical scheme. Assumptions in RTTOV, such as particle size and shape distributions, introduce further uncertainties in the simulated measurements.

There is still room for improvement. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), could enhance identification and characterization of convective system structure by leveraging the spatial coherence and horizontal structure of convective systems, albeit at the cost of increased computational and temporal resources. Future work could also explore tandem classification approaches, such as tracking cloud structure transitions between times t_i and t_f using dTb/dt to detect these shifts. This would require careful parametrization and consideration of numerous classes. Focusing on realistic phase transitions is crucial; for instance, a direct shift from clear-sky to deep convection in minutes seems improbable unless horizontal advection within the numerical model grid is accounted for. Another, more complex, approach would involve identifying regions of cloud growth, stability, or decay based on a study of ice content evolution, its vertical profile, and vertical wind velocities within convective systems.

Finally, a dedicated study will assess the impact of spatial observation geometry and tandem parameters on classification and regression retrieval performance, using an end-to-end C²OMODO simulator named RadioSpy (Radiometer Simulator in Python) developed by CNES. While the current study uses idealized C²OMODO measurements, the simulator will provide more realistic measurements, enabling a closer examination of the potential and informational content that C²OMODO could offer in an operational context following tandem launch.

Acknowledgments

This work was supported by Centre National d'Études Spatiales (CNES) and Airbus Defense and Space (ADS) under the AOS/C²OMODO program. The authors acknowledge the CNES and Airbus Defense and Space for funding support under the AOS/C²OMODO program. Jean-Pierre Chaboureau is acknowledged for providing the MESO-NH assistance. The authors thank also Thomas Fiolleau and Franck Auguste for insightful discussions for cloud structure classification criteria thresholds. The authors also acknowledge the NWPSAF team, and in particular the helpdesk which provided support for RTTOV. This study benefited from the IPSL mesocenter ESPRI facility for data and computing resources.

A LightGBM Model Parameterization

For each of the parameters shown below, the configured value is shown in square brackets.

- Learning Rate [0.01]: Controls the step size during boosting. A lower learning rate reduces each tree's impact, requiring more trees but potentially improving accuracy and generalization by making training more gradual and controlled.
- Subsample for Bins [200000]: Determines the sample size for constructing histograms to find optimal splits. Reducing the subsample can save memory and computation time but might lead to less precise binning, potentially affecting model performance.
- Number of Leaves [400]: Sets the maximum number of leaves per tree. More leaves allow the model to capture complex patterns, but an excessive number may lead to overfitting.
- Number of Estimators [500]: Specifies the total number of trees in the model. Increasing the number reduces bias but may increase overfitting and computational cost if not properly regularized.

- Class Weight (classification only): Addresses class imbalance by assigning weights, focusing on minority classes to improve performance. Three configurations are tested in this study:
  - Unweighted: Equal treatment of all classes, often underperforming on minority classes.
  - Balanced Weighted Class: Weights computed as:

$$\text{weight}_i = \frac{N}{n_i}$$

    Where N is the total sample count and n_i is the sample count for class i. This approach assigns higher weights to less frequent classes.
  - Intermediate Compromise: Empirically defined weights to balance emphasis, giving distinct importance to certain classes without overemphasizing the least represented. Assigned weights: CS=1, ANVL=1, STRAT=2, CONV=4, DC=2.

B Metrics for the Evaluation of the Classifications

The performance of the classification algorithm relies on the algorithm's True and False predictions. For a given class, True Positives (TP) represent correctly classified instances. False Positives (FP) are instances incorrectly predicted to belong to the class, while False Negatives (FN) are instances from the class that were misclassified. True Negatives (TN) represent instances accurately identified as not belonging to the class.

Then the performance criterions include Precision, Recall, and F1-score. Precision (P) measures the proportion of correct predictions for a specific class among all predictions made for that class, highlighting the rate of false detections. Recall (R), also known as Hit Rate, indicates the ratio of correct predictions for a class relative to all actual instances of that class. The F1-score (F1) provides a balanced measure by taking the harmonic mean of Precision and Recall, effectively addressing both false positives and false negatives. These metrics are calculated as follows:

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall (R, also called Hit Rate)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1-score} = \frac{2 \times P \times R}{P + R}$$

C Metrics for the Evaluation of the Retrievals

$\mathbf{R^2}$ measures the goodness of fit, showing how well it explains variance in the target variable, with values close to 1 indicating strong explanatory power. RMSE assesses prediction accuracy by measuring the average deviation between predicted and actual values, expressed in the same units as the target variable. Finally, MBE evaluates prediction bias, identifying whether the statistical model systematically overestimates or underestimates predictions. Together, these metrics allow for a comprehensive evaluation of performances, and are defined as follows:

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$$

Where y_i is the observed value of sample i, $\hat{y}_i$ is the predicted value for this sample, $\bar{y}$ is the mean of observed values, and n is the total number of observations.

References

[1] G. Stephens, J. Polcher, X. Zeng, P. v. Oevelen, G. Poveda, M. Bosilovich, M.-H. Ahn, G. Balsamo, Q. Duan, G. Hegerl, C. Jakob, B. Lamptey, R. Leung, M. Piles, Z. Su, P. Dirmeyer, K. L. Findell, A. Verhoef, M. Ek, T. L'Ecuyer, R. Roca, A. Nazemi, F. Dominguez, D. Klocke, and S. Bony, "The First 30 Years of GEWEX," *Bulletin of the American Meteorological Society*, vol. 104, no. 1, pp. E126–E157, Jan. 2023, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/104/1/BAMS-D-22-0061.1.xml

[2] Z. Feng, L. R. Leung, N. Liu, J. Wang, R. A. Houze Jr, J. Li, J. C. Hardin, D. Chen, and J. Guo, "A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking," *Journal of Geophysical Research: Atmospheres*, vol. 126, no. 8, p. e2020JD034202, 2021, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD034202. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2020JD034202

[3] C. Kidd, G. Huffman, V. Maggioni, P. Chambon, and R. Oki, "The Global Satellite Precipitation Constellation: Current Status and Future Requirements," *Bulletin of the American Meteorological Society*, vol. 102, pp. 1–47, May 2021.

[4] S. W. Nesbitt, R. Cifelli, and S. A. Rutledge, "Storm morphology and rainfall characteristics of TRMM precipitation features," *Monthly Weather Review*, vol. 134, no. 10, pp. 2702–2721, Oct. 2006. [Online]. Available: http://www.scopus.com/inward/record.url?scp=33748376450&partnerID=8YFLogxK

[5] R. Roca, J. Aublanc, P. Chambon, T. Fiolleau, and N. Viltard, "Robust Observational Quantification of the Contribution of Mesoscale Convective Systems to Rainfall in the Tropics," *Journal of Climate*, vol. 27, no. 13, pp. 4952–4958, Jul. 2014, publisher: American Meteorological Society Section: Journal of Climate. [Online]. Available: https://journals.ametsoc.org/view/journals/clim/27/13/jcli-d-13-00628.1.xml

[6] R. S. Schumacher and K. L. Rasmussen, "The formation, character and changing nature of mesoscale convective systems," *Nature Reviews Earth & Environment*, vol. 1, no. 6, pp. 300–314, Jun. 2020, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s43017-020-0057-7

[7] R. Roca and T. Fiolleau, "Extreme precipitation in the tropics is closely associated with long-lived convective systems," *Communications Earth & Environment*, vol. 1, no. 1, pp. 1–6, Sep. 2020, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s43247-020-00015-4

[8] G. L. Stephens, K. A. Shiro, M. Z. Hakuba, H. Takahashi, J. A. Pilewskie, T. Andrews, C. J. Stubenrauch, and L. Wu, "Tropical Deep Convection, Cloud Feedbacks and Climate Sensitivity," *Surveys in Geophysics*, May 2024. [Online]. Available: https://doi.org/10.1007/s10712-024-09831-1

[9] Z. Feng, L. R. Leung, J. Hardin, C. R. Terai, F. Song, and P. Caldwell, "Mesoscale Convective Systems in DYAMOND Global Convection-Permitting Simulations," *Geophysical Research Letters*, vol. 50, no. 4, p. e2022GL102603, 2023, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL102603. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2022GL102603

[10] A. F. Prein, Z. Feng, T. Fiolleau, Z. L. Moon, K. M. Núñez Ocasio, J. Kukulies, R. Roca, A. C. Varble, A. Rehbein, C. Liu, K. Ikeda, Y. Mu, and R. M. Rasmussen, "Km-Scale Simulations of Mesoscale Convective Systems Over South America—A Feature Tracker Intercomparison," *Journal of Geophysical Research: Atmospheres*, vol. 129, no. 8, p. e2023JD040254, 2024, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023JD040254. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2023JD040254

[11] Z. Feng, A. F. Prein, J. Kukulies, T. Fiolleau, W. K. Jones, B. Maybee, Z. Moon, K. M. N. Ocasio, W. Dong, M. J. Molina, M. G. Albright, R. Feng, J. Song, F. Song, L. R. Leung, A. Varble, C. Klein, and R. Roca, "Mesoscale Convective Systems tracking Method Intercomparison

(MCSMIP): Application to DYAMOND Global km-scale Simulations," 2024. [Online]. Available: https://doi.org/10.22541/essoar.172405876.67413040/v1

[12] L. J. Donner, T. A. O'Brien, D. Rieger, B. Vogel, and W. F. Cooke, "Are atmospheric updrafts a key to unlocking climate forcing and sensitivity?" *Atmospheric Chemistry and Physics*, vol. 16, no. 20, pp. 12 983–12 992, Oct. 2016, publisher: Copernicus GmbH. [Online]. Available: https://acp.copernicus.org/articles/16/12983/2016/

[13] J. Fan, B. Han, A. Varble, H. Morrison, K. North, P. Kollias, B. Chen, X. Dong, S. E. Giangrande, A. Khain, Y. Lin, E. Mansell, J. A. Milbrandt, R. Stenz, G. Thompson, and Y. Wang, "Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts," *Journal of Geophysical Research: Atmospheres*, vol. 122, no. 17, pp. 9351–9378, 2017, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017JD026622. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/2017JD026622

[14] A. Varble, E. J. Zipser, A. M. Fridlind, P. Zhu, A. S. Ackerman, J.-P. Chaboureau, S. Collis, J. Fan, A. Hill, and B. Shipway, "Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft properties," *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 24, pp. 13,891–13,918, 2014, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD021371. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021371

[15] J. Wu, A. D. Del Genio, M.-S. Yao, and A. B. Wolf, "WRF and GISS SCM simulations of convective updraft properties during TWP-ICE," *Journal of Geophysical Research: Atmospheres*, vol. 114, no. D4, 2009, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD010851. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010851

[16] D. Wang, S. E. Giangrande, Z. Feng, J. C. Hardin, and A. F. Prein, "Updraft and Downdraft Core Size and Intensity as Revealed by Radar Wind Profilers: MCS Observations and Idealized Model Comparisons," *Journal of Geophysical Research: Atmospheres*, vol. 125, no. 11, p. e2019JD031774, 2020, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD031774. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031774

[17] T. Fiolleau and R. Roca, "A database of deep convective systems derived from the intercalibrated meteorological geostationary satellite fleet and the TOOCAN algorithm (2012–2020)," *Earth System Science Data*, vol. 16, no. 9, pp. 4021–4050, Sep. 2024, publisher: Copernicus GmbH. [Online]. Available: https://essd.copernicus.org/articles/16/4021/2024/

[18] M. Kikuchi, S. Braun, K. Suzuki, G. Liu, and A. Battaglia, "Satellite Precipitation Measurements: What Have We Learnt About Cloud-Precipitation Processes From Space?" Dec. 2023, pp. 303–324.

[19] J. M. Futyan and A. D. D. Genio, "Deep Convective System Evolution over Africa and the Tropical Atlantic," *Journal of Climate*, vol. 20, no. 20, pp. 5041–5060, Oct. 2007, publisher: American Meteorological Society Section: Journal of Climate. [Online]. Available: https://journals.ametsoc.org/view/journals/clim/20/20/jcli4297.1.xml

[20] T. Fiolleau and R. Roca, "Composite life cycle of tropical mesoscale convective systems from geostationary and low Earth orbit satellite observations: method and sampling considerations," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 673, pp. 941–953, 2013, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2174. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2174

[21] D. Bouniol, R. Roca, T. Fiolleau, and D. E. Poan, "Macrophysical, Microphysical, and Radiative Properties of Tropical Mesoscale Convective Systems over Their Life Cycle," *Journal of Climate*, vol. 29, no. 9, pp.

3353–3371, May 2016, publisher: American Meteorological Society Section: Journal of Climate. [Online]. Available: https://journals.ametsoc.org/view/journals/clim/29/9/jcli-d-15-0551.1.xml

[22] G. S. Elsaesser, R. Roca, T. Fiolleau, A. D. Del Genio, and J. Wu, "A Simple Model for Tropical Convective Cloud Shield Area Growth and Decay Rates Informed by Geostationary IR, GPM, and Aqua/AIRS Satellite Data," *Journal of Geophysical Research: Atmospheres*, vol. 127, no. 10, p. e2021JD035599, 2022, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JD035599. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2021JD035599

[23] K. Yang, Z. Wang, M. Deng, and B. Dettmann, "Combining CloudSat/CALIPSO and MODIS measurements to reconstruct tropical convective cloud structure," *Remote Sensing of Environment*, vol. 287, p. 113478, Mar. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425723000299

[24] A. Illingworth, H. Barker, A. Beljaars, M. Ceccaldi, H. Chepfer, N. Clerbaux, J. Cole, J. Delanoë, C. Domenech, D. Donovan, S. Fukuda, M. Kikuchi, R. Hogan, A. Hünerbein, P. Kollias, T. Kubota, T. Nakajima, T. Nakajima, T. Nishizawa, and G.-J. Zadelhoff, "THE EARTHCARE SATELLITE: The next step forward in global measurements of clouds, aerosols, precipitation and radiation." *Bulletin of the American Meteorological Society*, vol. 96, pp. 1311–1332, Aug. 2015.

[25] A. J. Illingworth, A. Battaglia, J. Bradford, M. Forsythe, P. Joe, P. Kollias, K. Lean, M. Lori, J.-F. Mahfouf, S. Melo, R. Midthassel, Y. Munro, J. Nicol, R. Potthast, M. Rennie, T. H. M. Stein, S. Tanelli, F. Tridon, C. J. Walden, and M. Wolde, "WIVERN: A New Satellite Concept to Provide Global In-Cloud Winds, Precipitation, and Cloud Properties," *Bulletin of the American Meteorological Society*, vol. 99, no. 8, pp. 1669–1687, Aug. 2018, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/99/8/bams-d-16-0047.1.xml

[26] G. Stephens, S. van den Heever, Z. Haddad, D. Posselt, R. Storer, L. Grant, O. Sy, T. Rao Fna Fasc, S. Tanelli, and E. Peral, "A Distributed Small Satellite Approach for Measuring Convective Transports in the Earth's Atmosphere," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, pp. 1–10, Oct. 2019.

[27] Z. Haddad, R. Sawaya, S. Prasanth, M. van den Heever, O. Sy, S. van den Heever, L. Grant, T. Narayana Rao, G. Stephens, S. Hristova-Veleva, D. Posselt, and R. Storer, "Observation strategy of the INCUS mission: retrieving vertical mass flux in convective storms from low-earth-orbit convoys of miniaturized microwave instruments," pp. EGU22–2124, May 2022, conference Name: EGU General Assembly Conference Abstracts ADS Bibcode: 2022EGUGA..24.2124H. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2022EGUGA..24.2124H

[28] H. Brogniez, R. Roca, F. Auguste, J.-P. Chaboureau, Z. Haddad, S. J. Munchak, X. Li, D. Bouniol, A. Dépée, T. Fiolleau, and P. Kollias, "Time-Delayed Tandem Microwave Observations of Tropical Deep Convection: Overview of the C2OMODO Mission," *Frontiers in Remote Sensing*, vol. 3, Apr. 2022, publisher: Frontiers. [Online]. Available: https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2022.854735/full

[29] F. Auguste and J.-P. Chaboureau, "Deep Convection as Inferred From the C2OMODO Concept of a Tandem of Microwave Radiometers," *Frontiers in Remote Sensing*, vol. 3, Apr. 2022, publisher: Frontiers. [Online]. Available: https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2022.852610/full

[30] P. J. Marinescu, S. C. van den Heever, S. M. Saleeby, and S. M. Kreidenweis, "The microphysical contributions to and evolution of latent heating profiles in two MC3E MCSs," *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 13, pp. 7913–7935, 2016, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD024762. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/2016JD024762

[31] R. Houze, "Clouds in Tropical Cyclones," *Monthly Weather Review*, vol. 138, pp. 293–344, Feb. 2010.

[32] R. Roca, H. Brogniez, P. Chambon, O. Chomette, S. Cloché, M. E. Gosset, J.-F. Mahfouf, P. Raberanto, and N. Viltard, "The Megha-Tropiques mission: a review after three years in orbit," *Frontiers in Earth Science*, vol. 3, May 2015, publisher: Frontiers. [Online]. Available: https://www.frontiersin.org/journals/earth-science/articles/10.3389/feart.2015.00017/full

[33] E. Di Tomaso, F. Romano, and V. Cuomo, "Rainfall estimation from satellite passive microwave observations in the range 89 GHz to 190 GHz," *Journal of Geophysical Research: Atmospheres*, vol. 114, no. D18, 2009, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD011746. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2009JD011746

[34] R. Eastman, M. Lebsock, and R. Wood, "Warm Rain Rates from AMSR-E 89-GHz Brightness Temperatures Trained Using CloudSat Rain-Rate Observations," *Journal of Atmospheric and Oceanic Technology*, vol. 36, no. 6, pp. 1033–1051, Jun. 2019, publisher: American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/36/6/jtech-d-18-0185.1.xml

[35] C. Kummerow, W. Olson, and L. Giglio, "A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1213–1232, Sep. 1996, conference Name: IEEE Transactions on Geoscience and Remote Sensing. [Online]. Available: http://ieeexplore.ieee.org/document/536538

[36] K. F. Evans and G. L. Stephens, "Microweve Radiative Transfer through Clouds Composed of Realistically Shaped Ice Crystals. Part II. Remote Sensing of Ice Clouds." *Journal of the Atmospheric Sciences*, vol. 52, pp. 2058–2072, Jun. 1995, aDS Bibcode: 1995JAtS...52.2058E. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1995JAtS...52.2058E

[37] K. F. Evans, S. J. Walter, A. J. Heymsfield, and M. N. Deeter, "Modeling of Submillimeter Passive Remote Sensing of Cirrus Clouds," *Journal of Applied Meteorology*, vol. 37, no. 2, pp. 184–205, Feb. 1998. [Online]. Available: http://journals.ametsoc.org/doi/10.1175/1520-0450(1998)037⟨0184:MOSPRS⟩2.0.CO;2

[38] K. F. Evans, J. R. Wang, P. E. Racette, G. Heymsfield, and L. Li, "Ice Cloud Retrievals and Analysis with the Compact Scanning Submillimeter Imaging Radiometer and the Cloud Radar System during CRYSTAL FACE," *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 839–859, Jun. 2005. [Online]. Available: https://journals.ametsoc.org/doi/10.1175/JAM2250.1

[39] S. A. Buehler, C. Jiménez, K. F. Evans, P. Eriksson, B. Rydberg, A. J. Heymsfield, C. J. Stubenrauch, U. Lohmann, C. Emde, V. O. John, T. R. Sreerekha, and C. P. Davis, "A concept for a satellite mission to measure cloud ice water path, ice particle size, and cloud altitude," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. S2, pp. 109–128, 2007, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.143. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.143

[40] J.-F. Rysman, C. Claud, J.-P. Chaboureau, J. Delanoë, and B. M. Funatsu, "Severe convection in the Mediterranean from microwave observations and a convection-permitting model," *Quarterly Journal of the Royal Meteorological Society*, vol. 142, no. S1, pp. 43–55, 2016, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2611. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2611

[41] G. Hong, G. Heygster, J. Miao, and K. Kunzi, "Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements," *Journal of Geophysical Research: Atmospheres*, vol. 110, no. D5, 2005, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004JD004949. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004949

[42] J. A. Gagliano, J. M. Schuchardt, T. T. Wilheit, and J. L. King, *Moisture sounding at millimeter wavelengths /94/183 GHz/ at high altitudes*, Jan. 1981, nTRS Author Affiliations: Georgia Inst. of Tech., Georgia

Institute of Technology, NASA Goddard Space Flight Center NTRS Meeting Information: Atmospheric effects on electro-optical, infrared, and millimeter wave systems performance; 1981-08-27 to 1981-08-28; undefined NTRS Document ID: 19830041339 NTRS Research Center: Legacy CDMS (CDMS). [Online]. Available: https://ntrs.nasa.gov/citations/19830041339

[43] K. F. Evans, J. R. Wang, D. O'C Starr, G. Heymsfield, L. Li, L. Tian, R. P. Lawson, A. J. Heymsfield, and A. Bansemer, "Ice hydrometeor profile retrieval algorithm for high-frequency microwave radiometers: application to the CoSSIR instrument during TC4," *Atmospheric Measurement Techniques*, vol. 5, no. 9, pp. 2277–2306, Sep. 2012, publisher: Copernicus GmbH. [Online]. Available: https://amt.copernicus.org/articles/5/2277/2012/amt-5-2277-2012.html

[44] C. Birman, J.-F. Mahfouf, M. Milz, J. Mendrok, S. A. Buehler, and M. Brath, "Information content on hydrometeors from millimeter and sub-millimeter wavelengths," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 69, no. 1, Jan. 2017. [Online]. Available: https://a.tellusjournals.se/articles/10.1080/16000870.2016.1271562

[45] C. Lac, J.-P. Chaboureau, V. Masson, J.-P. Pinty, P. Tulet, J. Escobar, M. Leriche, C. Barthe, B. Aouizerats, C. Augros, P. Aumond, F. Auguste, P. Bechtold, S. Berthet, S. Bielli, F. Bosseur, O. Caumont, J.-M. Cohard, J. Colin, F. Couvreux, J. Cuxart, G. Delautier, T. Dauhut, V. Ducrocq, J.-B. Filippi, D. Gazen, O. Geoffroy, F. Gheusi, R. Honnert, J.-P. Lafore, C. Lebeaupin Brossier, Q. Libois, T. Lunet, C. Mari, T. Maric, P. Mascart, M. Mogé, G. Molinié, O. Nuissier, F. Pantillon, P. Peyrillé, J. Pergaud, E. Perraud, J. Pianezze, J.-L. Redelsperger, D. Ricard, E. Richard, S. Riette, Q. Rodier, R. Schoetter, L. Seyfried, J. Stein, K. Suhre, M. Taufour, O. Thouron, S. Turner, A. Verrelle, B. Vié, F. Visentin, V. Vionnet, and P. Wautelet, "Overview of the Meso-NH model version 5.4 and its applications," *Geoscientific Model Development*, vol. 11, no. 5, pp. 1929–1969, May 2018, publisher: Copernicus GmbH. [Online]. Available: https://gmd.copernicus.org/articles/11/1929/2018/

[46] J.-P. Chaboureau, J.-P. Cammas, P. Mascart, J.-P. Pinty, C. Claud, R. Roca, and J.-J. Morcrette, "Evaluation of a cloud system life-cycle simulated by the Meso-NH model during FASTEX using METEOSAT radiances and TOVS-3I cloud retrievals," *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 566, pp. 1735–1750, 2000, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49712656609. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712656609

[47] J.-P. Chaboureau, N. Söhne, J.-P. Pinty, I. Meirold-Mautner, E. Defer, C. Prigent, J. R. Pardo, M. Mech, and S. Crewell, "A Midlatitude Precipitating Cloud Database Validated with Satellite Observations," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 5, pp. 1337–1353, May 2008, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology. [Online]. Available: https://journals.ametsoc.org/view/journals/apme/47/5/2007jamc1731.1.xml

[48] T. Dauhut, J.-P. Chaboureau, J. Escobar, and P. Mascart, "Large-eddy simulations of Hector the convector making the stratosphere wetter," *Atmospheric Science Letters*, vol. 16, no. 2, pp. 135–140, 2015, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/asl2.534. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/asl2.534

[49] R. Saunders, J. Hocking, E. Turner, P. Rayer, D. Rundle, P. Brunel, J. Vidot, P. Roquet, M. Matricardi, A. Geer, N. Bormann, and C. Lupu, "An update on the RTTOV fast radiative transfer model (currently at version 12)," *Geoscientific Model Development*, vol. 11, no. 7, pp. 2717–2737, Jul. 2018, publisher: Copernicus GmbH. [Online]. Available: https://gmd.copernicus.org/articles/11/2717/2018/

[50] A. J. Geer, P. Bauer, K. Lonitz, V. Barlakas, P. Eriksson, J. Mendrok, A. Doherty, J. Hocking, and P. Chambon, "Bulk hydrometeor optical properties for microwave and sub-millimetre radiative transfer in RTTOV-SCATT v13.0," *Geoscientific Model Development*, vol. 14, no. 12, pp. 7497–7526, Dec. 2021, publisher: Copernicus GmbH. [Online]. Available: https://gmd.copernicus.org/articles/14/7497/2021/

[51] Z. Feng, X. Dong, B. Xi, C. Schumacher, P. Minnis, and M. Khaiyer, "Top-of-atmosphere radiation budget of convective core/stratiform rain and anvil clouds from deep convective systems," *Journal of Geophysical Research: Atmospheres*, vol. 116, no. D23, 2011, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011JD016451. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016451

[52] P. V. Hobbs, T. J. Matejka, P. H. Herzegh, J. D. Locatelli, and R. A. Houze, Jr., "The Mesoscale and Microscale Structure and Organization of Clouds and Precipitation in Midlatitude Cyclones. I: A Case Study of a Cold Front." *Journal of the Atmospheric Sciences*, vol. 37, pp. 568–596, Mar. 1980, aDS Bibcode: 1980JAtS...37..568H. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1980JAtS...37..568H

[53] R. A. Houze Jr., "Observed structure of mesoscale convective systems and implications for large-scale heating," *Quarterly Journal of the Royal Meteorological Society*, vol. 115, no. 487, pp. 425–461, 1989, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49711548702. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711548702

[54] ——, *Cloud Dynamics*, Mar. 1994. [Online]. Available: https://shop.elsevier.com/books/cloud-dynamics/houze-jr/978-0-12-356881-6

[55] R. A. Houze, "Stratiform Precipitation in Regions of Convection: A Meteorological Paradox?" *Bulletin of the American Meteorological Society*, vol. 78, no. 10, pp. 2179–2196, Oct. 1997, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/78/10/1520-0477_1997_078_2179_spiroc_2_0_co_2.xml

[56] R. A. Houze Jr., "Mesoscale convective systems," *Reviews of Geophysics*, vol. 42, no. 4, 2004, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004RG000150. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1029/2004RG000150

[57] R. Houze, "100 Years of Research on Mesoscale Convective Systems," *Meteorological Monographs*, vol. 59, Jun. 2018.

[58] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[59] J. H. Friedman, "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, publisher: Institute of Mathematical Statistics. [Online]. Available: https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-5/Greedy-function-approximation-A-gradient-boosting-machine/10.1214/aos/1013203451.full

[60] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY: Springer, 2009. [Online]. Available: http://link.springer.com/10.1007/978-0-387-84858-7

[61] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 785–794. [Online]. Available: https://dl.acm.org/doi/10.1145/2939672.2939785

[62] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: a highly efficient gradient boosting decision tree," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., Dec. 2017, pp. 3149–3157.

[63] S. Liu, M. Li, Z. Zhang, B. Xiao, and X. Cao, "Multimodal Ground-Based Cloud Classification Using Joint Fusion Convolutional Neural Network," *Remote Sensing*, vol. 10, no. 6, p. 822, Jun. 2018, number: 6 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2072-4292/10/6/822

[64] S. Liu, M. Li, Z. Zhang, B. Xiao, and T. S. Durrani, "Multi-Evidence and Multi-Modal Fusion Network for Ground-Based Cloud Recognition," *Remote Sensing*, vol. 12, no. 3, p. 464, Jan. 2020, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2072-4292/12/3/464

[65] J. Zhang, P. Liu, F. Zhang, H. Iwabuchi, A. Moura, and V. Albuquerque, "Ensemble Meteorological Cloud Classification Meets Internet of Dependable and Controllable Things," *IEEE Internet of Things Journal*, vol. PP, pp. 1–1, Dec. 2020.

Supplement material

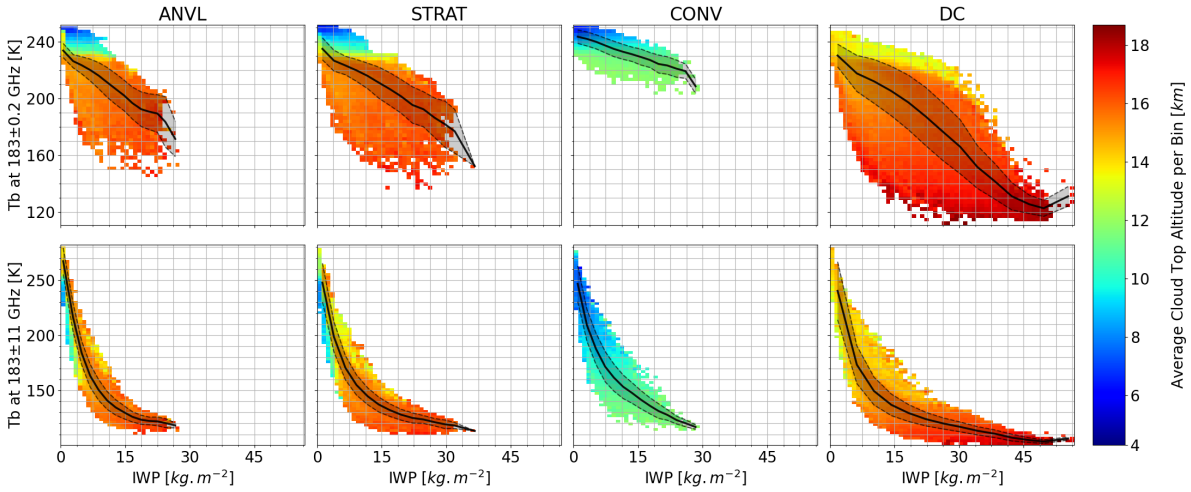

Figure 11: Scatter plots of average Tb at the center and the wings of the 183 GHz line, depending of the mean IWP. Colors represent the average cloud top altitude.