



**HAL**  
open science

## Global monitoring of deep convection using passive microwave observations

Jean-François Rysman, Chantal Claud, Stavros Dafis

► **To cite this version:**

Jean-François Rysman, Chantal Claud, Stavros Dafis. Global monitoring of deep convection using passive microwave observations. *Atmospheric Research*, 2021, 247, pp.105244. 10.1016/j.atmosres.2020.105244 . hal-03109099

**HAL Id: hal-03109099**

**<https://hal.sorbonne-universite.fr/hal-03109099>**

Submitted on 13 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Global monitoring of deep convection using passive microwave observations

Jean-François Rysman\*<sup>1</sup>, Chantal Claud<sup>1</sup>, Stavros Dafis<sup>1,2</sup>

\*Corresponding Author : jfrysman@lmd.polytechnique.fr

1 - LMD/IPSL, École Polytechnique, Institut Polytechnique de Paris, ENS, PSL

Université, Sorbonne Université, CNRS, Palaiseau France

2 - National Observatory of Athens, Institute for Environmental Research and Sustainable Development, Vas. Pavlou & Metaxa, Athens, Greece

## Abstract

In this study, we present the DEEPSTORM (DEEP moiSt aTmospheric cOnvection from micRowave radioMeter) algorithm, able to retrieve ice water path (IWP) and to detect deep moist atmospheric convection (DC) from 80°S to 80°N using observations from four spaceborne passive microwave radiometers. DEEPSTORM is based on a machine learning approach and is fitted against observations from the CPR (Cloud Profiling Radar) spaceborne radar on-board CloudSat. IWP predictions show an average root mean square error of 0.27 kg/m<sup>2</sup> and a correlation index of 0.87. DC occurrence is detected with a probability of 59% and a false alarm rate of 24%. The prediction accuracy of IWP and DC is significantly better when the IWP exceeds 0.5 kg/m<sup>2</sup> showing that DEEPSTORM is well suited to detect and characterise the strongest DC events. Overall DC detection is more accurate in the tropics than in mid-latitudes while the IWP retrieval works better in the

mid-latitudes. Two examples illustrating the potential of DEEPSTORM are presented: the IWP is retrieved during Hurricane Matthew in 2016, and a climatology of DC occurrences between September 2016 and December 2016 is presented. This work will allow building a quasi-worldwide and 20-year long database of DC occurrence and intensity.

## **Keywords**

Deep moist atmospheric convection; passive microwave; atmospheric ice

## **Highlights**

- DEEPSTORM detects convection and retrieves ice with passive microwave observations
- DEEPSTORM detects well the strongest and more elevated convective clouds
- Frequencies suited for convection characterisation with radiometers are highlighted
- This work will lead to the development of a worldwide and 20-year long climatology

## **1. Introduction**

Deep moist atmospheric convection (DC) is a very frequent meteorological event that occurs near globally. DC plays a central role at meteorological scale by, for instance, homogenising, horizontally and vertically, the thermal energy received by the Earth's surface. DC also redistributes vertically the atmospheric gases and in particular the water vapour that is sometimes injected into the lower Stratosphere. At ground level, DC is an essential provider of rain and thus has a role for biodiversity and farming activity but also for the mitigation of atmospheric pollution events through air cleansing. DC is also

frequently associated with severe meteorological phenomena such as heavy precipitation, strong winds, hail, lightning and tornadoes (Brunner et al., 2007; Doswell, 2001). Often, those have an important societal impact with hundreds of casualties recorded every year in Europe and the USA, and probably many more worldwide. The economic cost of DC is also significant and is estimated between 1 and 8 billion of euros in Europe (Brooks & Dotzek, 2008; Hoeppe, 2016), and over 10 billion in the USA (Allen, 2018).

The role and impact of DC emphasise the need for comprehensive and long-term monitoring of this phenomenon. However, the global monitoring of DC is challenging because DC can occur in remote regions (e.g., ocean, isolated islands) where no ground observations are available and because it can sometimes be of relatively brief duration and small size (i.e.,  $\sim 30$  min and 20 km). Additionally, the last IPCC scientific report highlights the lack of global observations to study past and present trends in the context of climate change.

At present, spaceborne instruments appear to be the only available solution for a global monitoring of DC. In particular, active and passive microwave instruments are well adapted to this task thanks to the ability of microwave frequencies to probe within convective clouds and to detect atmospheric ice. Spaceborne radars can thus provide invaluable information on the vertical structure of microphysics in convective clouds. However, the spaceborne radars do not have large swaths ( $\sim 1.4$  km for CloudSat) and thus they sample only a limited fraction of severe convective clouds. Passive microwave radiometers (PMR) are also able to detect severe convection by taking advantage of the high frequency channels' ability to detect the scattering of Earth radiations by the ice

crystals (e.g., Cecil and Blankenship, 2012; Ferraro et al., 2015; Hong et al., 2005; Laviola and Levizzani, 2011). Unlike radars, PMR have a very large swath (up to 2000 km) and many have been launched for 30 years which ensures a long-term sampling of DC on a global scale. Hong et al., (2005) developed a diagnostic able to detect deep convection and convective overshooting clouds using high frequency channels which has been validated with airborne and spaceborne collocated observations (Hong et al., 2005; Rysman et al., 2016a; Rysman et al., 2016b). These diagnostics have been widely used in past years (Aumann and Ruzmaikin, 2013; Claud et al., 2012; Funatsu et al., 2007, 2008, 2009, 2012, 2018; Hong and Heygster, 2008; Sanò et al., 2015). However, the Hong et al., (2005) algorithm presents several limitations: first its efficiency decreases drastically over frozen soil and mountainous regions due to the surface contamination of the signal (Rysman et al., 2016b). Second, it has been designed specifically for the Advanced Microwave Sounding Unit (AMSU-B) radiometers and while it has been shown that it can be adapted to the MHS (Microwave Humidity Sounder) (which is a radiometer very similar to AMSU-B), it cannot be applied to newest generation instruments due to different channel frequency sets (e.g., SAPHIR (Soundeur Atmosphérique du Profil d’Humidité Intertropicale par Radiométrie), ATMS (Advanced Technology Microwave Sounder)) and/or different scanning strategies (GMI (Global precipitation measurement Microwave Imager)). Finally, the Hong et al., (2005) diagnostic does not inform on the intensity of the detected convection. As such it does not allow to distinguish powerful vertical convective towers and shallower cumulus clouds.

In this context, we developed an algorithm named DEEPSTORM (DEEP moiSt aTmospheric cONvection from micRowave radioMeter) able to overcome the

aforementioned difficulties. DEEPSTORM detects DC worldwide and retrieves the ice water path (IWP), which is used as a proxy of DC intensity, and is adapted for four different passive microwave radiometers. DEEPSTORM is trained, tuned and validated against coincident measurements of the CloudSat CPR radar. Datasets and methods are presented in Section 2. Evaluation of the algorithm is conducted in Section 3 and two applications of the algorithm are shown in Section 4. Discussion and conclusion are presented in the last two sections.

## **2. Data and methods**

### **Rationales**

This study aims at detecting deep moist atmospheric convection (DC) and retrieving ice water path (IWP) using observations from four different types of passive microwave radiometer : MHS (Microwave Humidity Sounder) on-board the National Oceanic and Atmospheric Administration - 18 (N18) platform (Bonsignori, 2007), SAPHIR (Sondeur Atmosphérique du Profil d'Humidité Intertropicale par Radiométrie) on-board the Megha-Tropiques (MT) satellite (Clain et al., 2015), ATMS (Advanced Technology Microwave Sounder) on-board the Suomi National Polar-orbiting Partnership (NPP) platform (Blackwell et al., 2011), GMI (GPM Microwave Imager) on-board the Global Precipitation Measurement – Core Observatory (GPM) satellite (Draper et al., 2015). MHS and ATMS operate on polar orbiting satellites, GMI has a drifting orbit ( $65^\circ$  inclination) and SAPHIR an equatorial orbit. All radiometers probe at several frequencies around the water vapour absorption line at 183.3 GHz and, except for SAPHIR, at about 89 GHz. ATMS and GMI also probe at lower frequencies, down to 10 GHz for GMI. In addition, several of GMI channels are both vertically and horizontally polarised (e.g., 166 GHz). All radiometers use

a cross-track scanning except GMI which uses a conical scanning at a constant viewing angle of  $53^\circ$ .

Passive microwave frequencies are sensitive to the presence of frozen hydrometeors in the troposphere and thus to the DC occurrence (e.g., Bennartz and Bauer, 2003; Deeter and Vivekanandan, 2005; Surussavadee and Staelin, 2006). In particular, a depression of brightness temperature (BT) in high frequency channels (e.g., Fig. 1. in Rysman et al., (2018)) can be a signature of frozen hydrometeors as those scatter Earth's radiation. However, several other factors such as the variation of humidity, atmospheric temperature and surface type also affect BT in a non-linear manner which makes it difficult, in some cases, to discriminate atmospheric convection from other features. Therefore the development of specific algorithms able to tackle these problems is of great importance. A further challenge is to characterise the intensity of atmospheric convection with passive microwave radiometers. A way to infer the convection intensity is to use atmospheric ice as a proxy. Indeed, a high magnitude of ice in a convective cell is a signature of strong and sustained vertical updrafts and thus of intense events. In addition, high contents of atmospheric ice are likely to yield lightning and heavy precipitation including hail but also strong surface wind gusts. For instance, Heymsfield et al., (2020) showed that 57% of all Earth's precipitation consists in precipitating ice that melts to produce rain. Therefore, retrieving IWP using BT can inform on the intensity of DC. However, the relationship between BT and IWP is non-linear and is affected by many other parameters such as humidity and temperature. Such high-dimensional problems, associated with a high volume of data, can be addressed using machine learning techniques. These techniques allow to characterise the high diversity of atmospheric conditions in which DC occurs and to

identify underlying patterns. It is for this purpose that we developed the DEEPSTORM (DEEP moiSt aTmospheric cOnvection from micRowave radioMeter) algorithm.

### **Coincidence databases**

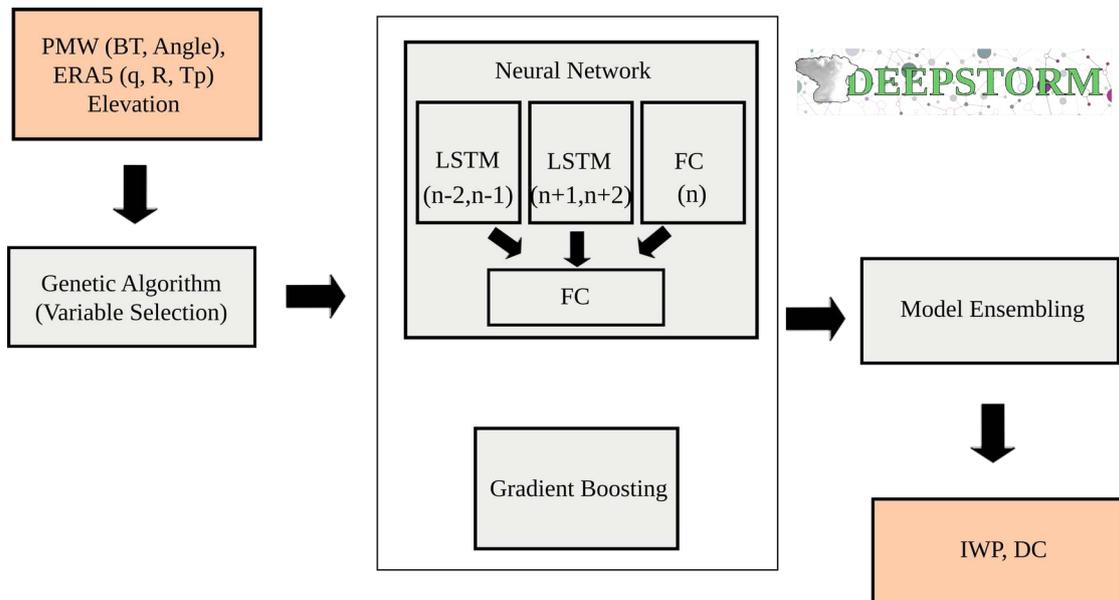
Building an algorithm able to detect DC and retrieve IWP requires first a reference dataset. In our study, we chose the CloudSat CPR (Cloud Profiling Radar) observations to train and evaluate the DEEPSTORM predictions. CPR is able to distinguish eight cloud types (2B-CLDCLASS product; Sassen and Wang (2008)) including Deep Convective Clouds used for DC detection, and to retrieve ice water content with the DARDAR-CLOUD product (Ceccaldi et al., 2013; Delanoë and Hogan, 2008, 2010) that we used to compute the IWP. Specifically, we built co-located databases between CPR observations and each radiometer observations, i.e., we identified every +/- 15 min CPR measurements embedded in each radiometer pixel for the period spanning between 2006 and 2017. This represents, for MHS on-board N18, more than 200 million of CPR pixels with 436 vertical levels each. Then, we averaged CPR measurements on the coincident radiometer pixel and we extracted several variables including the ice water path (IWP), the occurrence of DC, the altitude of the maximum ice content and the top and bottom altitude of the DC cloud. Note that, as soon as at least one CPR pixel detected DC, we flagged the whole associated radiometer pixel as DC. In addition, the co-located database includes all the BT and the viewing angles for each radiometer, the ground elevation, and the temperature ( $T_p$ ), the specific humidity ( $q$ ) and relative humidity ( $R$ ) profiles at 37 pressure levels (from 1 hPa to 1000 hPa) extracted from the ERA-5 reanalysis (Copernicus Climate Change Service, 2017). The nearest neighbour approach has been used to match the closest ERA-5 grid

point (available on a spatio-temporal grid of  $0.25^\circ \times 0.25^\circ$  and 6 h) to each passive microwave pixel.

Coincident measurements between CPR and the radiometers are unevenly spread over the globe with an increasing number with latitude due to their quasi polar orbits (except for SAPHIR). This could bias the DEEPSTORM algorithm. Therefore, we resampled the datasets by randomly taking a constant number of collocations for each degree of latitude for each radiometer. We finally obtained 4 databases with 300,000 (GMI), 400,000 (SAPHIR), 4 million (ATMS) and 13 million (MHS) coincidences.

### **DEEPSTORM algorithm**

When using machine learning algorithms, the larger the training dataset, the better the prediction accuracy. However, larger datasets increase greatly the computing time, thus in the variable selection and tuning procedures described thereafter, for which each model has been run hundreds of thousands of times, we limited each database to 300,000 randomly selected points. The full coincident databases are only used once for the final fitting of the models.



**Fig. 1. Scheme of the DEEPSTORM algorithm. First, input variables are provided to the genetic algorithm which selects the minimal set of variable required to detect DC and retrieve IWP. Then, these variables are provided to a neural network and a gradient boosting model. The neural network has 2 levels: the first level consists in 2 LSTM layers (32 neurons +16 neurons) and one FC layer (16 neurons) and the second level is made of one FC layer (32 neurons +8 neurons +1 neuron). The results of those two models are combined to obtain two output variables: the IWP and the DC occurrence.**

The first step was to select the appropriate input variables to predict DC and IWP (Fig. 1). As previously mentioned, more than 100 variables (BT, humidity ...) are available in the four coincident databases. Using a large number of input variables increases the probability to take into account all important source of information to reach the best

possible predictions, but it also makes the task more difficult for a machine learning algorithm to converge to an optimal solution. This is why, we reduced the number of input variables by keeping only variables related to DC or IWP using the gradient boosting variable importance information (Friedman et al., 2001). This reduces the number of input variables to about 30 (depending on the radiometer). However, several of the selected variables provide redundant information, e.g., the specific humidity variables at 1000 and 950 hPa levels are both related to the occurrence of DC but are strongly correlated and thus keeping only one can be enough. Thus, we used a genetic algorithm (Gen and Lin, 2007) that allowed to select, iteratively, the variable set that is able to predict DC and IWP with the highest accuracy (500 iterations have been performed). This leads to a minimal set of about 10 variables (depending on the type of radiometer, see results in Section 3 and supplementary material) that have then been used for predictions.

The second step was to identify the machine learning algorithm able to provide the most accurate prediction using those input variables (Fig. 1). We chose to use a neural network and a Gradient Boosting, two algorithms that have been recently acclaimed as among the most powerful algorithms for providing accurate predictions with very large datasets and are well adapted to meteorological data (e.g., Petković et al., 2019; Rysman et al., 2019; Sanò et al., 2018; Tang et al., 2018; Xiao et al., 2019). Using two (or more) algorithms rather than a single one to provide a prediction is called model ensembling. The advantage of this procedure is to reduce the bias of prediction outputs. Thus, in this study we averaged the outputs of the neural network and the Gradient Boosting models to provide the final DEEPSTORM output.

Neural networks consist of one or several layers chosen by the user depending of the characteristics of the input and output variables. A layer consists of two or more units (called neurons) connected to the units of the preceding and succeeding layers. The neural network modifies iteratively those units and the connections between those units to converge to a solution which is the closest possible to the expected output. In this study, we used two kinds of layer: the fully connected layer (FC) (also called dense layer) (LeCun et al., 2015) which is the more widespread kind of layer and the long short term memory (LSTM) layer (Gers et al., 1999). The main difference between those two layers is that the FC layer uses observations at a given pixel ( $n$ ) and predict output at a given pixel ( $n$ ) such as :

$$(BT1..j, q1..i, \dots)_n \rightarrow FC \rightarrow IWP_n \text{ or } DC_n$$

where  $j$  is the number of the channel,  $i$  is the number of vertical level, and  $n$  is a given passive microwave pixel. The LSTM is a layer in which the input signal is treated in sequence. For our case, it means that, in order to predict IWP or DC at a pixel  $n$ , the LSTM takes the input variables of the four neighbour pixels (e.g.,  $n-2, n-1, n+1, n+2$ ) in a sequence, i.e., the LSTM takes into account the fact that the pixel  $n-2$  is further away than pixel  $n-1$  from the pixel  $n$  to predict output in pixel  $n$  :

$$(BT1..j, q1..i, \dots)_{n-2, n-1, n+1, n+2} \rightarrow LSTM \rightarrow IWP_n \text{ or } DC_n$$

Using the neighbour pixels increases the overall prediction accuracy and reduces the possible noise since the data are autocorrelated, i.e., IWP at a pixel  $n$  is correlated to IWP in the surrounding pixels ( $n-2, n-1, n+1, n+2$ ). The Gradient Boosting (Chen and Guestrin, 2016) is an algorithm which, similarly to the random forest algorithm, combines the output of numerous decision trees (30 for the DEEPSTORM algorithm). It is based on the gradient boosting technique that consists of iteratively training models (i.e., several

decision trees) that address the weakness of the previous models. A central procedure in the Gradient Boosting algorithm is the tuning. Indeed, the Gradient Boosting algorithm behaviour is deeply affected by several parameters, named hyperparameters, that need to be tuned to reach an accurate prediction. In order to tune the Gradient Boosting and later to evaluate the predictions of the overall DEEPSTORM algorithm, we split the coincident database into two independent training and testing datasets of respectively 80% and 20% of the initial dataset. Note that the LSTM layers used in the neural network of DEEPSTORM require to keep neighbour pixels sorted sequentially. This is why the points of the training and testing datasets have been selected in order to preserve this sequencing. Using the training dataset, we tuned the hyperparameters of the gradient boosting algorithm using a Bayesian optimisation (Bischl et al., 2017).

Then, both algorithms (the neural network and the Gradient Boosting) have been fitted on the training dataset and we created a model ensembling, i.e., we averaged output from both algorithms (Fig. 1). Finally we evaluated the predictions accuracy using the following metrics for IWP : correlation index, bias, root mean square error (RMSE) and the fractional standard error percentage (FSE) defined as:

$$\text{FSE} = 100 \frac{\text{RMSE}}{\langle \text{IWP} \rangle}$$

where  $\langle \text{IWP} \rangle$  is the mean IWP.

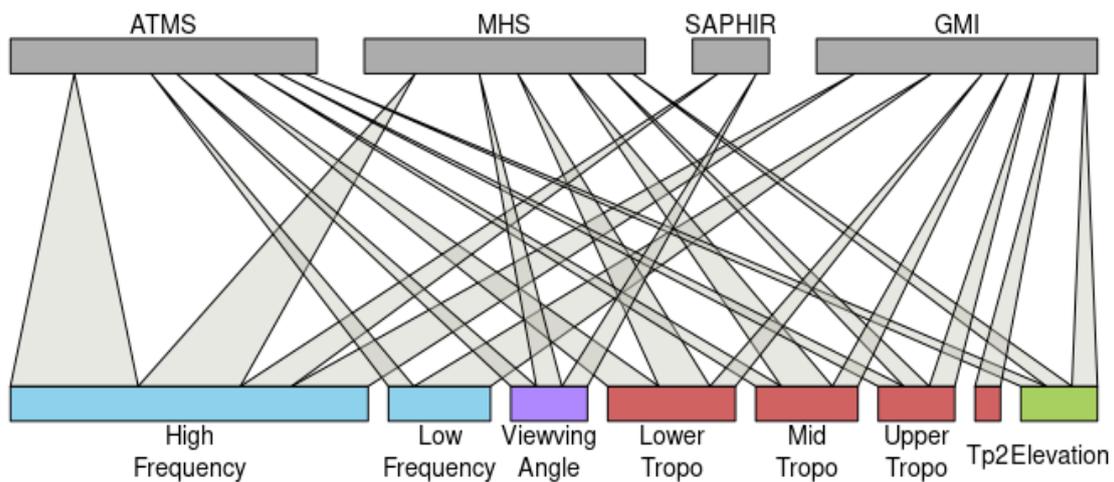
Probability of detection, false alarm rate and Heidke Skill Score (HSS) (or Cohen Kappa) were used to evaluate the DC detection algorithm :

$$\text{Heidke Skill Score (HSS)} = \frac{2(h.cn - fa.m)}{(h+m)(m+cn) + (h+fa)(fa+cn)}$$

where  $h$  is the hits (both prediction and reference detect an event),  $cn$  is the correct negatives (prediction and reference do not detect an event),  $m$  is the misses (prediction does not detect any event but the reference does) and  $fa$  is the false alarms (prediction detects an event but the reference does not). HSS ranges between -1 (worst prediction skill) and 1 (best prediction skill). In the results section, we separated the analyses between the Tropics ( $30^{\circ}\text{S}$ - $30^{\circ}\text{N}$ ) and Mid-latitudes ( $> 30^{\circ}\text{N}$  or  $< 30^{\circ}\text{S}$ ) since SAPHIR observations are only available in the Tropics.

### 3. Results

#### Ice water path retrieval

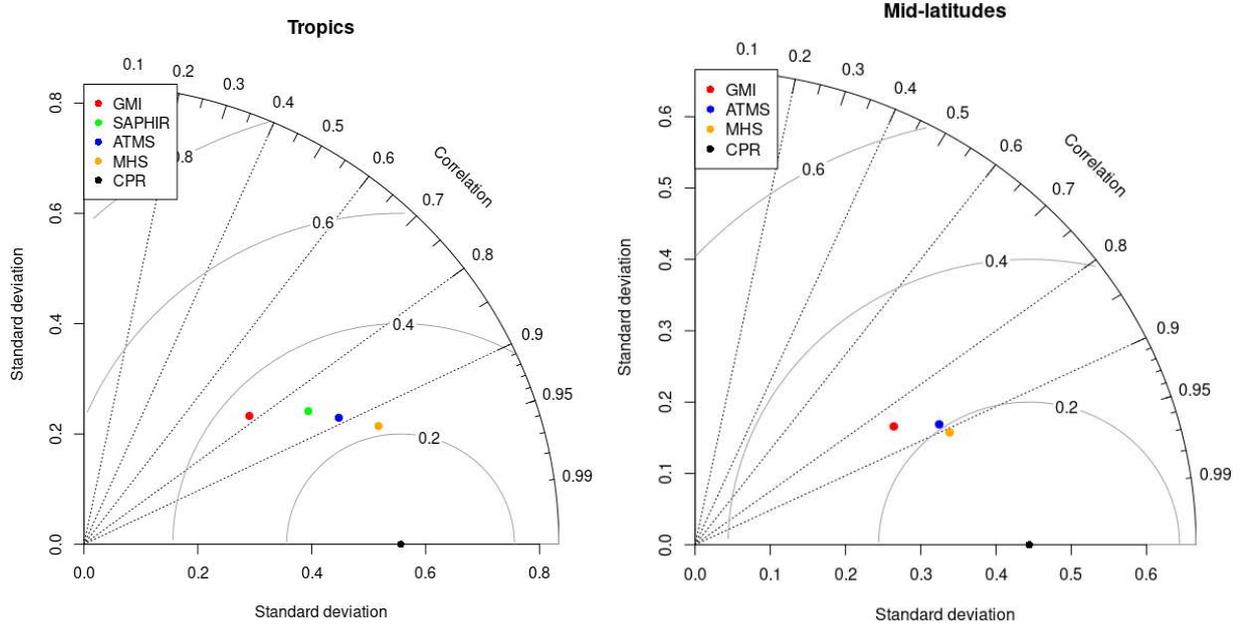


**Fig. 2. Variables used by each radiometer to retrieve IWP. Box and link sizes are proportional to the number of variables selected. BT variables are represented in blue, environmental variables ( $q$ ,  $R$ ,  $T_p$ ) are represented in red.  $T_p2$  stands for temperature at 2 hPa. High (Low) frequency BT correspond to channels with frequencies greater (lower) than 88 GHz. Lower**

**troposphere variables correspond to levels below 700 hPa, Mid troposphere variables corresponds to levels between 700 and 400 hPa and Upper troposphere variables corresponds to levels between 400 hPa and 100 hPa.**

Variable selection procedure identifies the minimal set of variables required to retrieve IWP, i.e., the set of variables that contains non-redundant information which is used by DEEPSTORM to provide the best prediction of the IWP. Eleven or twelve variables are used by DEEPSTORM for retrieving the IWP with MHS, ATMS and GMI while only three variables are used for SAPHIR (Table 1 – Supplementary Material). Fig. 2 shows that DEEPSTORM selects several high frequency channels for each radiometer (e.g., channels at 89.5, 165.5,  $183.31 \pm 7$ ,  $183.31 \pm 4.5$  and  $183.31 \pm 3$  GHz for ATMS). In addition some lower frequency channels are also selected by ATMS (1) and GMI (3) (MHS and SAPHIR do not have any low frequency channels). The viewing angle is also selected for each radiometer except for GMI for which the viewing angle is fixed. Ground elevation and environmental variables are selected for every radiometers except SAPHIR. Specifically, among all environmental variables selected, specific humidity is chosen 10 times over 12. SAPHIR requires much fewer variables to retrieve the IWP compared to the other radiometers (i.e., only the viewing angle and BT at  $183.31 \pm 11$  GHz and at  $183.31 \pm 4.2$  Ghz). Since SAPHIR probes only in the tropics this suggests a lesser variability in the tropical atmosphere. This hypothesis is emphasised by the fact that, when the variable selection procedure is conducted for the other radiometers with their observations restricted to the tropics, only 3 to 6 variables are used by DEEPSTORM to retrieve IWP (e.g., the viewing angle and the BT at 165.5 and  $183.31 \pm 7$  GHz are selected for ATMS). Finally, among the selected variables, we also identified the variable that provides the

most information for retrieving the IWP for each radiometer. Hence, the most informative variables are BT at  $183.31 \pm 11$  GHz, 165.5 GHz,  $183.31 \pm 7$  GHz and 157 GHz for SAPHIR, ATMS, GMI and MHS respectively.



**Fig. 3.** Taylor diagrams of GMI, SAPHIR, ATMS and MHS predictions of IWP with DEEPSTORM algorithm for tropics (a) and mid-latitudes (b) with respect to CPR observations. The correlation between prediction and CPR is indicated by the radial distance with respect to the y-axis and the RMSE is represented by the circle centred on reference (CPR) on the x-axis.

Then, we evaluated the skills of the algorithm for each sounder, separately for the tropics and the mid-latitudes, using Taylor Diagrams (Fig. 3). In Taylor Diagrams (Taylor, 2001), the correlation between prediction and reference values is indicated by the radial distance with respect to the y-axis and the RMSE is represented by the circle centred on the reference value on the x-axis. MHS reaches the highest accuracy in terms of correlation index (0.92 for the tropics and 0.91 for the mid-latitudes) and RMSE ( $0.23 \text{ kg/m}^2$  for the

tropics,  $0.18 \text{ kg/m}^2$  for the mid-latitudes). For the other instruments and in the tropics (mid-latitudes), we find a correlation index of 0.89 (0.89), 0.85 and 0.78 (0.85) and a RSME  $0.26$  ( $0.19$ ),  $0.35$ ,  $0.34$  ( $0.25$ )  $\text{kg/m}^2$  for respectively ATMS, SAPHIR and GMI.

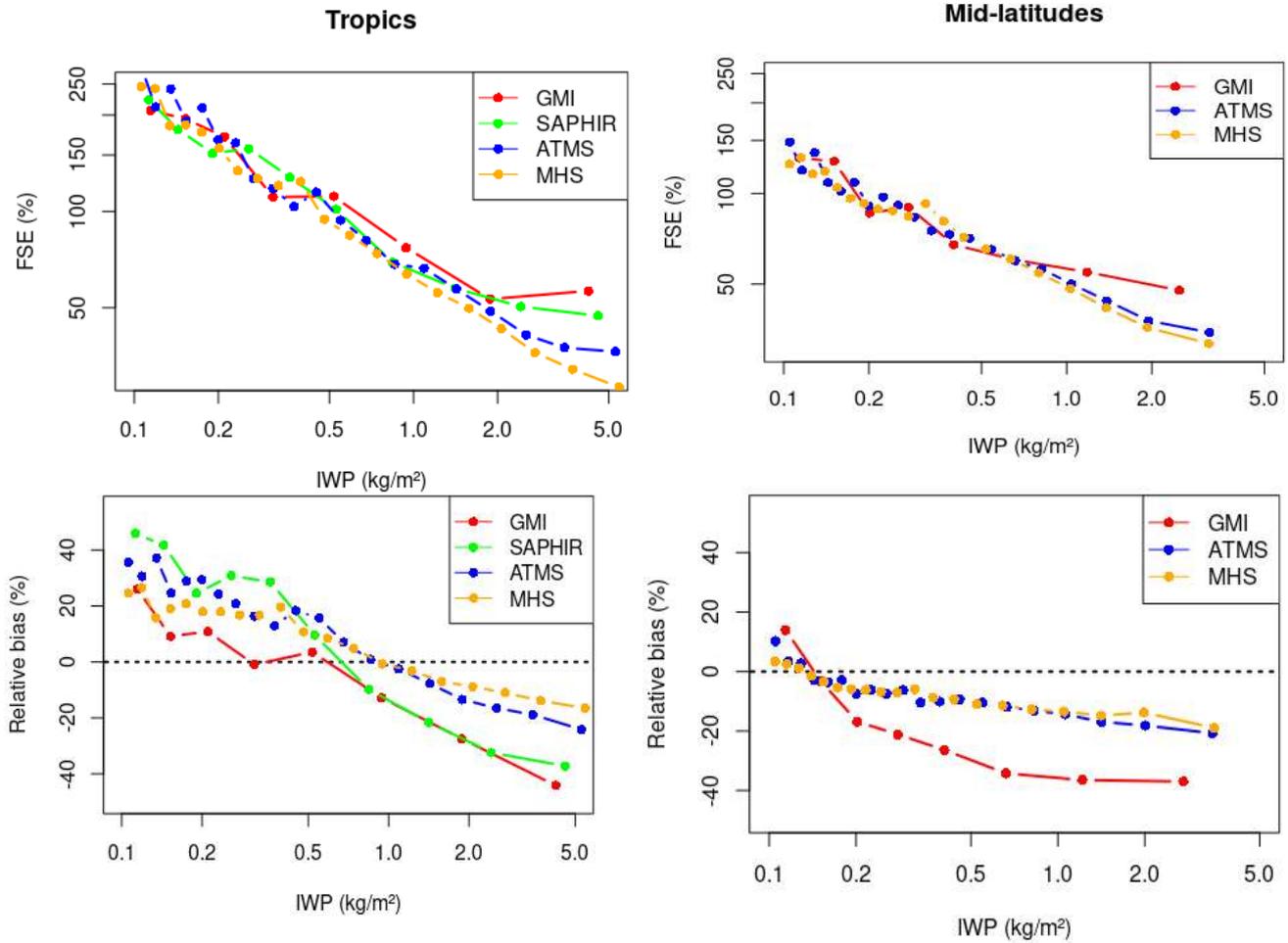
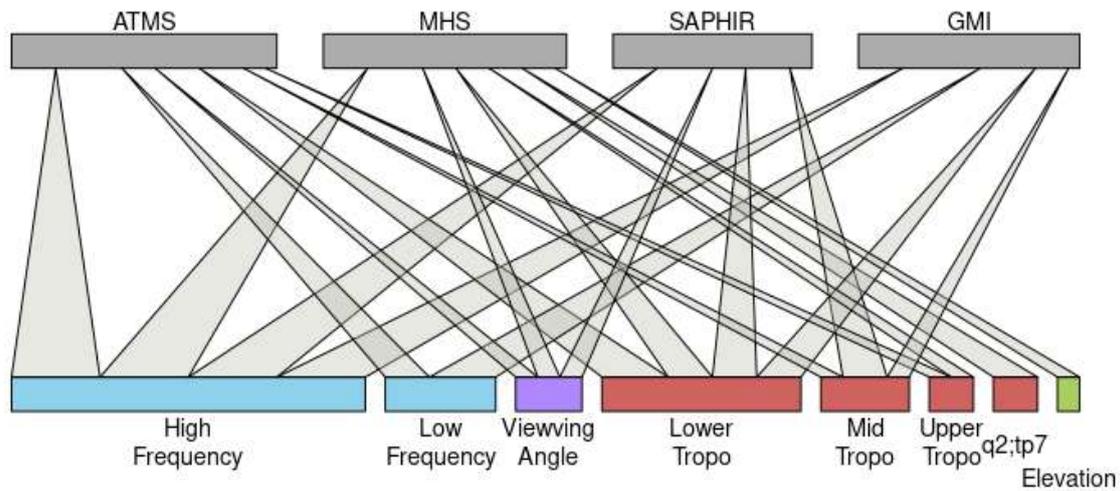


Fig. 4. (top left) FSE for tropics, (top right) FSE for mid-latitudes, (bottom left) relative bias for tropics and (bottom right) relative bias for mid-latitudes as a function of IWP (IWP and FSE are represented in log scale and relative bias is represented in linear scale). Observations have been split using IWP then the average FSE and relative bias have been computed for each bin.

In order to characterise the sensitivity of the predictions to the observed IWP, we computed the FSE and the relative bias as a function of the observed IWP (Fig. 4). Fig. 4 (top left and right) indicates that, for all sounders, the FSE decreases continuously as IWP increases. ATMS and MHS have a similar behaviour for IWP values lower than  $0.7 \text{ kg/m}^2$  with a FSE ranging from about 250% (150% in the mid-latitudes) to 70% (90% in the mid-latitudes), above MHS becomes slightly better than ATMS and reaches 28% for  $5.4 \text{ kg/m}^2$  in the Tropics. GMI's FSE is nearly as good as MHS and ATMS between 0.1 and  $0.7 \text{ kg/m}^2$  but when IWP is greater than  $0.7 \text{ kg/m}^2$ , FSE decreases much more slowly and reaches 50% (56% in the tropics) at  $2.7 \text{ kg/m}^2$  ( $4.2 \text{ kg/m}^2$  in the tropics). A similar behaviour is found for SAPHIR, with a FSE that ranges from 223% at  $0.1 \text{ kg/m}^2$  to 47% at  $4.6 \text{ kg/m}^2$ . Regarding relative bias, MHS and ATMS have a very low absolute bias in the mid-latitudes (lower than  $\pm 15\%$ ). Their relative bias is more substantial in the tropics (from 40% at  $0.1 \text{ kg/m}^2$  to about -20% at  $5.4 \text{ kg/m}^2$ ) and remains lower (in absolute value) for MHS regardless the IWP. SAPHIR predictions are slightly more biased for low IWP and, significantly more biased for higher values, as they underestimate true IWP up to -37% for  $4.6 \text{ kg/m}^2$ . GMI has the lowest absolute bias of all radiometers for low IWP in the tropics but it underestimates more strongly observed IWP in case of very high values. In the mid-latitudes its predictions are noticeably worse than MHS and ATMS reaching, for instance, -37% for  $2.7 \text{ kg/m}^2$ .

### **DC detection**



**Fig. 5. Same as Fig. 2 but for DC detection. q2 stands for the specific humidity at 2 hPa and tp7 stands for temperature at 7 hPa.**

The variables used by DEEPSTORM for the DC detection (Fig. 5 and Table 2 – Supplementary Material) are quite similar to those found in the IWP retrieving section (Fig. 2 and Table 1 – Supplementary Material), i.e., three to four high frequency channels are used for each radiometer, and two or three low frequency channels are used by ATMS and GMI. Specific and relative humidity variables in the low and mid troposphere are also used. Ground elevation is chosen by MHS and the viewing angle is used for every radiometer prediction except for GMI. It can be noted that more variables are used to detect DC with SAPHIR than for retrieving IWP (9 vs 3). In addition, the most informative variables for detecting DC are the same than for retrieving IWP (BT at 157 GHz, 165.5 GHz and  $183.31 \pm 11$  GHz for MHS, ATMS and SAPHIR) except for GMI for which it is the BT at 166.5 GHz (horizontal polarisation).

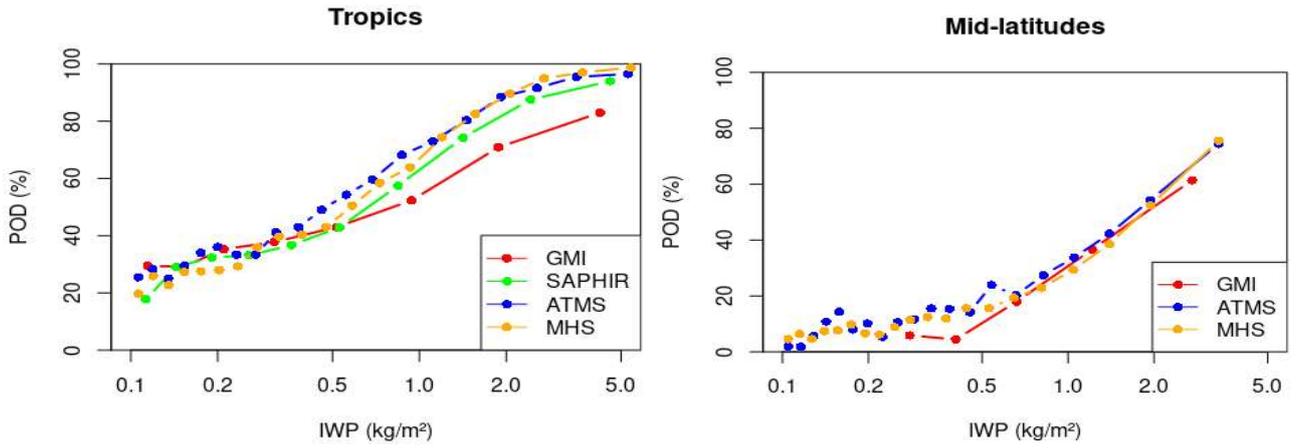
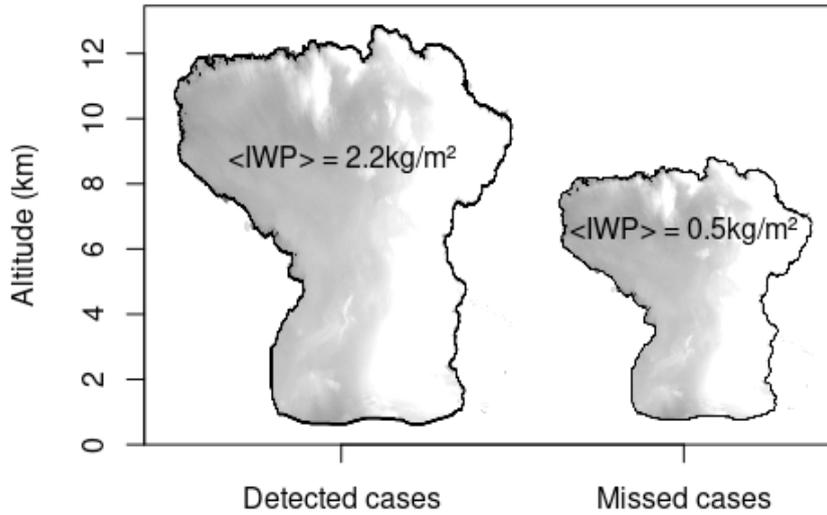


Fig. 6. (top) Probability of detection in the tropics (left) and mid-latitudes (right) as a function of ice water path (IWP is represented in log scale and POD is represented in linear scale). Observations have been split using IWP then the average POD has been computed for each bin.

We used the variables selected by the genetic algorithm to predict the occurrence of DC for each radiometer. In the tropics, MHS and ATMS have the highest HSS of all radiometers with 0.72, followed by SAPHIR and GMI with a HSS of 0.67. The probability of detection (POD) of DC occurrence is of 0.67, 0.68, 0.69 and 0.59 for SAPHIR, MHS, ATMS and GMI respectively, and the false alarm rate (FAR) is of 0.29, 0.20, 0.23, 0.2 for SAPHIR, MHS, ATMS and GMI respectively. It shows that, in the tropics, GMI detects less events than the other radiometers, and that SAPHIR tends to falsely detect non-DC cases more often. In mid-latitudes, the overall predictions are poorer: MHS (ATMS; GMI) has a HSS of 0.5 (0.51; 0.46), a POD of 0.38 (0.4; 0.34) and a FAR of 0.26 (0.29; 0.29). In Fig. 6, the sensitivity of the POD to the IWP is presented. All satellites have a similar behaviour in the tropics with a POD of about 20-30% for IWP at 0.1 kg/m<sup>2</sup>, i.e, for the shallowest atmospheric convection, and that increases with the IWP. For SAPHIR, ATMS and MHS, the POD exceeds 50% for IWP above 0.7 kg/m<sup>2</sup> and 90% for IWP above 4 kg/m<sup>2</sup>. For GMI, the POD is lower for all IWP with 50% at 1 kg/m<sup>2</sup> and about 78% at 4

kg/m<sup>2</sup>. In mid-latitudes, the detection rate is below 20% for IWP lower than 0.5 kg/m<sup>2</sup>, then it increases and overpasses 50% at about 1.8 kg/m<sup>2</sup>. It is slightly lower for GMI.



**Fig. 7. Median characteristics of detected and missed DC events by DEEPSTORM.**

In order to highlight the characteristics of the detected DC events, we extracted the median values of bottom and top altitudes of the clouds, and IWP of all detected and all missed DC (Fig. 7). Detected cases correspond to clouds with much higher vertical extent, i.e., 500 m to 13 km versus 700 m to 8.9 km for missed cases. In addition, detected cases have an median IWP about 4 times greater than the one of missed cases.

#### **4. Applications of DEEPSTORM**

In spite of the aforementioned limitations for weak cases, DEEPSTORM is able to detect DC and retrieve IWP at a quasi global scale. In this section, we present two applications of the DEEPSTORM algorithm. First, we analyse the IWP distribution and evolution within the Hurricane Matthew, then we show a quasi global climatology of DC occurrence.

## Hurricane Matthew

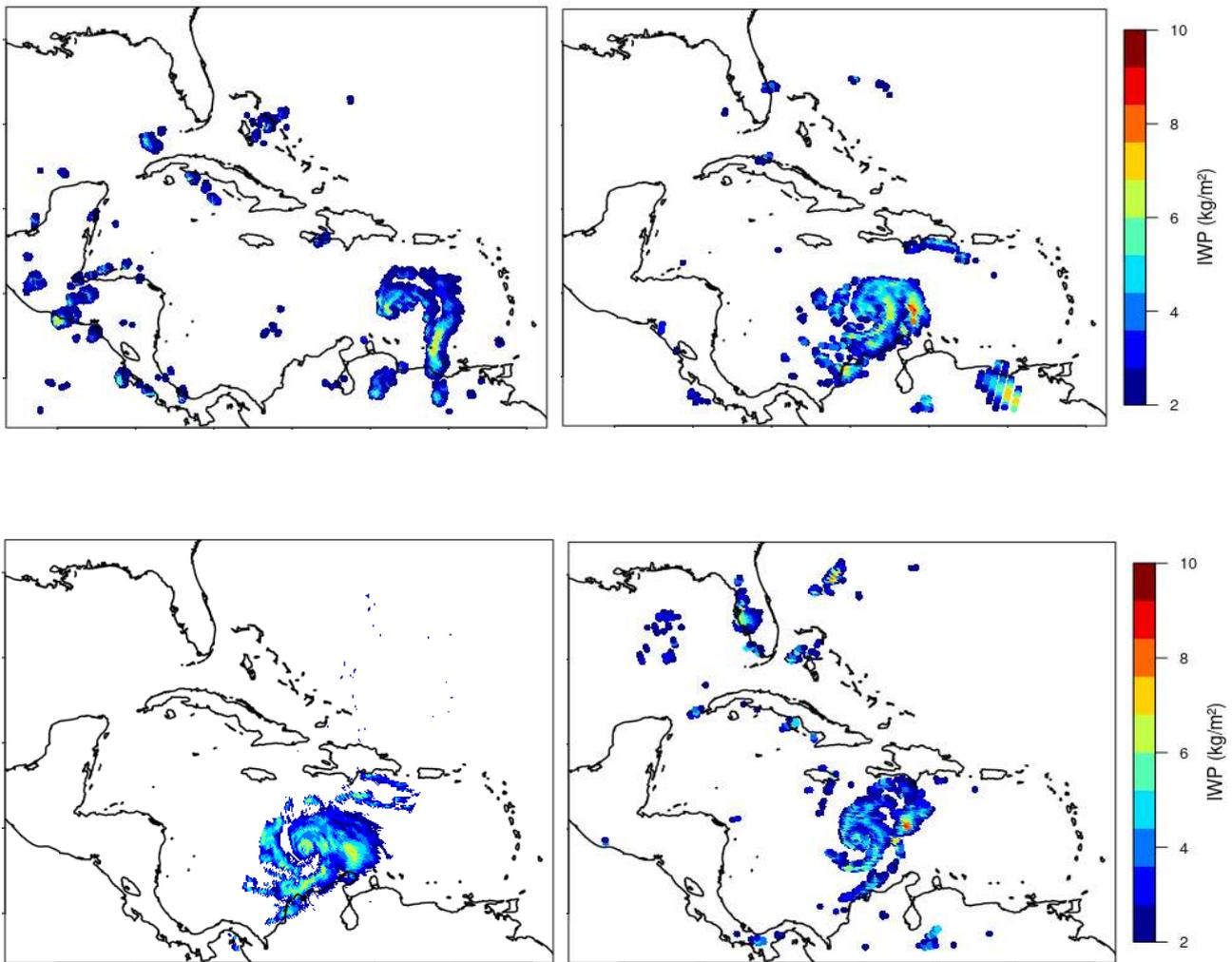
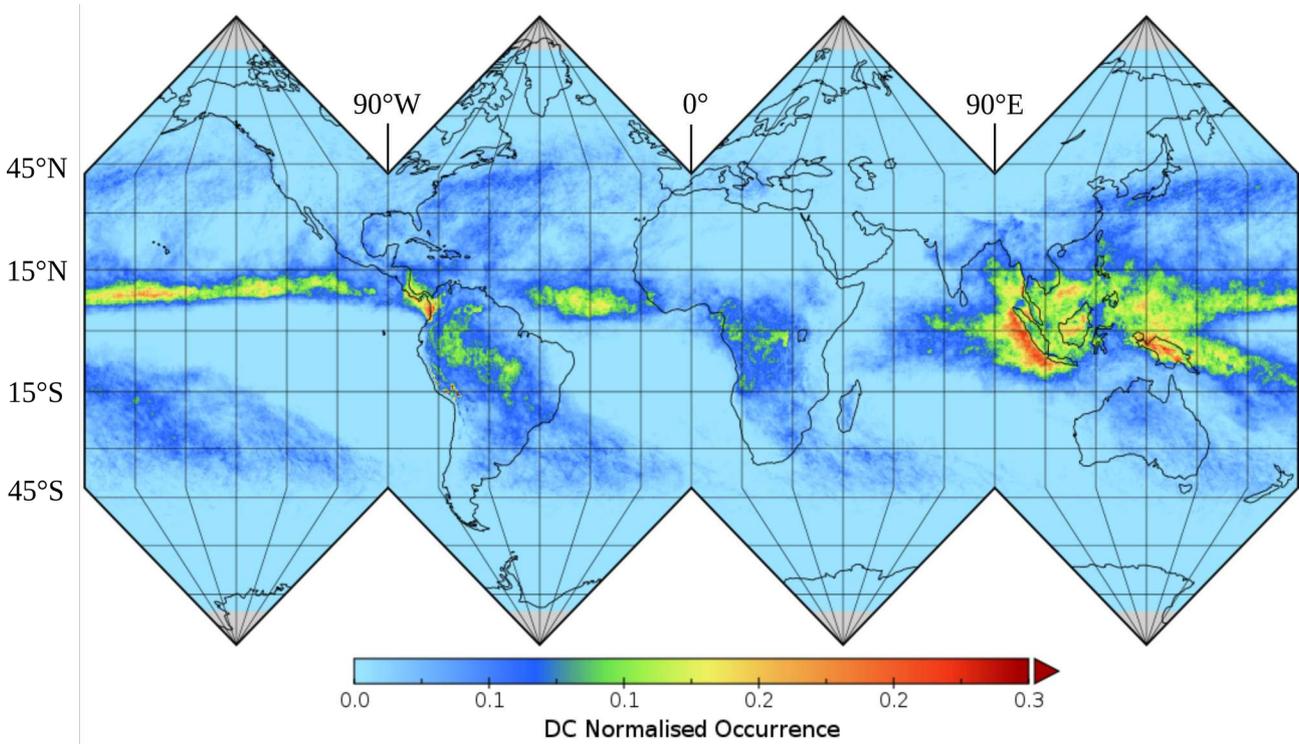


Fig. 8. IWP retrieval for the Hurricane Matthew on 2016-09-30 04:12Z with SAPHIR (top left), 2016-10-02 06:33Z with ATMS (top right), 2016-10-02 09:40Z with GMI (bottom left), 2016-10-02 23:13Z with MHS (bottom right).

Hurricane Matthew was a major hurricane of the 2016 Atlantic hurricane season that led to more than 500 casualties (Stewart, 2017). It emerged on the eastern side of Lesser Antilles on September 28, 2016 then moved following a constant latitude (around 15° N) to the Caribbean Sea and from the October 2, 2016 it began to move northwardly, travelling

between Cuba and Haiti on the 4 and 5 October 2016 and reached Florida on the October 6, 2016 in the afternoon (Miller et al., 2018; Stewart, 2017; Guimond et al., 2020). This hurricane was observed by at least 6 passive microwave radiometers (MHS onboard Metop A/B and N18, ATMS onboard NPP, GMI onboard GPM-core and SAPHIR onboard MT) with more than 100 overpasses between the September 28 and the October 6, 2016. In Fig. 8, we selected four timesteps (one for each type of radiometer) for which IWP has been retrieved using DEEPSTORM. First, we selected an observation on the September 30, 2016 from SAPHIR (Fig. 8 (a)) just after the hurricane had entered in the Caribbean Sea. At this time, Matthew presented a comma shape with two IWP maxima of about  $7 \text{ kg/m}^2$  located in the northwestern part and in the southeastern part of the hurricane. The three other panels of Fig. 8 show the evolution of Matthew during October 2, 2016: in the morning ATMS observed the cyclone that was wrapped up around an “eye” centred around  $14^\circ\text{N}$ ,  $74^\circ\text{W}$ . The active part of the hurricane was found to the southern and eastern region of the hurricane. Interestingly, the strongest IWP values occurred in the outer active region of the hurricane with values reaching  $9 \text{ kg/m}^2$ . This is also true for GMI and MHS observations. Three hours later (bottom left panel of Fig. 8) Matthew’s “eye” was clearly formed as highlighted by the  $10 \times 10 \text{ km}$  ice-free region. The overall structure did not change much compared to the previous time step but the IWP maximum was lower, reaching  $7 \text{ kg/m}^2$ . On the final time step (bottom right panel) Matthew moved northwestwardly with the active part about to reach the Dominican Republic and Haiti. Two elongated and narrow bands can be seen in the southern part of the hurricane and the most active part was found in the eastern part with a IWP up to  $9 \text{ kg/m}^2$ . This case study highlights that DEEPSTORM can be used for the tracking of a hurricane during its whole life time without night interruption, that can happen with visible frequency measurements.

## DC occurrence climatology



**Fig. 9.** DC normalised occurrence from September 2016 to December 2016 retrieved using observations from MHS on-board N18, SAPHIR on-board MT, ATMS on-board NPP, GMI on-board GPM-Core.

In Fig. 9 the DC normalised occurrence, i.e., the number of DC occurrence over the total number of observations, in a  $0.2^\circ \times 0.2^\circ$  regular grid, from September to December 2016 (i.e., summer of the southern hemisphere), between  $80^\circ\text{S}$  and  $80^\circ\text{N}$  for MHS on-board N18, SAPHIR on-board MT, ATMS on-board NPP, GMI on-board GPM, is shown. A value of 0.3 means that DC has been detected 30% of time during this period. The maximas of DC normalised occurrence are found between  $15^\circ\text{S}$  and  $15^\circ\text{N}$ . Specifically, the areas with the

highest rates are found in the maritime continent with maxima over 0.3 on the western coast of Sumatra and Java islands and in Papua New Guinea mainland, in the eastern coast of Colombia and in a thin band along  $8^{\circ}\text{N}$  in the western Pacific. High values around 0.15 are also found in the Central Atlantic, and over South America and Central Africa. In addition, DC are detected 10% of time over the Ocean between  $45^{\circ}\text{S}$  and  $15^{\circ}\text{S}$  and between  $15^{\circ}\text{N}$  and  $45^{\circ}\text{N}$ . Finally, DC are found less than 1% of time over  $60^{\circ}\text{N}$  and never over  $60^{\circ}\text{S}$ . Several studies (Aumann and Ruzmaikin, 2013; Hong and Heygster, 2008; Hong et al., 2008; Liu et al., 2007, 2008; Liu and Liu, 2016; Wu et al., 2020) investigated deep convection and convective clouds distribution in tropical regions (and globally for Liu and Liu, 2016) using space observations (precipitation radar and/or microwave and infrared radiometers). The DC distribution, highlighted in Fig. 9, matches qualitatively very well with those presented in these studies. Similarly to our results, the Maritime Continent, Central and South America and Africa are identified, in all studies, as the regions with the highest rate of convective activity. It can also be noted that more DC occurrences are detected in eastern Pacific and central Atlantic in Fig. 9 than in other studies. This could be related to the different time frame (September to December 2016) and DC definition used in this work.

## 5. Discussion

Results highlight the capability of DEEPSTORM to detect DC and retrieve IWP. In this section, we discuss further those results in particular to understand why some variables have been selected by DEEPSTORM, and to investigate the possible origin of the discrepancies in prediction accuracy between radiometers.

## Variable selection

DEEPSTORM unsurprisingly selects BT from high-frequency channels which are sensitive to frozen hydrometeor occurrence in the troposphere. In addition, the top informative variables for each radiometer are always BT with frequencies between 157 GHz and  $183.31 \pm 7$  GHz. On the other hand, DEEPSTORM never selects the channels closest to the water vapour absorption line (e.g.,  $183.31 \pm 1.0$  GHz for ATMS). This means that these channels probe too high in the atmosphere to provide useful information about convection in the troposphere or that the information contained in these channels is already provided by other channels. Channels with frequencies around 89 GHz are nearly always selected but the polarisation of channel is never used in DEEPSTORM (e.g., 166 GHz V and H for GMI). This is surprising since, as polarisation is a signature of deep convection (Gong and Wu, 2017), we could have expected that the algorithm would use it. This means that information contained in one channel (166 GHz V or H) associated with some other channels is enough to detect DC and that the difference of polarisation between the two polarised channels does not provide additional information for discriminating between DC and not DC cases. It is also worth highlighting that some lower frequency channels are selected (when available). For instance, 50.3 GHz and 23.8 GHz channels have been selected for detection DC with ATMS. These channels are use by DEEPSTORM probably because they indicate, over ocean, the occurrence of rainfall associated with DC through the warm emission of liquid precipitation process (e.g., Kummerow et al., 2001).

Regarding environmental variables, humidity appears as the most essential variable for both IWP retrieval and DC detection. It is selected at 3 ranges of altitude : close to the surface, in mid-troposphere and in upper-troposphere. This is not a surprise as a moist

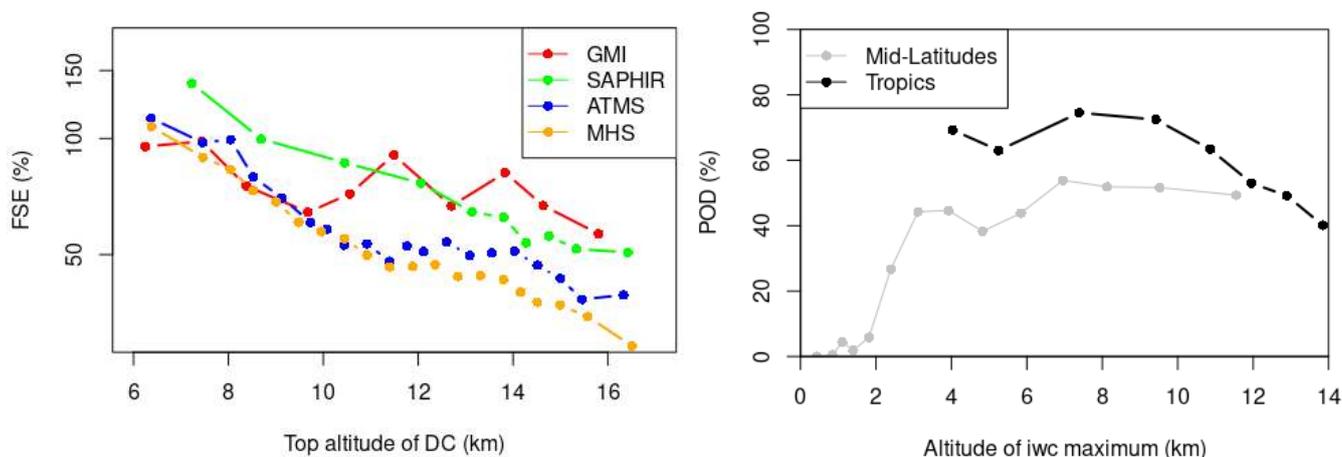
atmosphere is favourable to DC and large IWP. In addition, we highlighted that HF channels are used for the predictions and those channels are highly sensitive to the presence of humidity. Thus, having information about tropospheric humidity helps DEEPSTORM to use more insightfully the signal measured by those channels. Notably, no environmental variables are selected for retrieving IWP with SAPHIR radiometer. In addition, no environmental variable is neither selected by the other radiometers (except for MHS which only selects specific humidity at 950 hPa) when coincidences databases are limited to the tropics (not shown in the results section). This means that retrieving IWP in the tropics only requires BT temperatures measurements. This is probably related to the fact that atmospheric humidity is always very high in the Tropics and thus the HF channels are not subject to surface contamination as they can be when humidity is low in the mid-latitudes.

Overall, this variable selection procedure allowed to identify key variables to retrieve IWP and to detect DC for each radiometer. Such a data-driven selection of variables could, for instance, be used to help choosing the appropriate channel frequencies for future spatial missions devoted to convection.

### **Radiometers performance**

In Section 3, we also evaluated the ability of each radiometer to predict IWP and DC occurrence. MHS shows the highest RMSE and the lowest bias for predicting IWP and has the best HSS for DC detection. These good scores could be due to the fact that MHS is the radiometer with the lowest noise equivalent delta temperature ( $NE\Delta T$ ) for high frequency channels (0.41 K). The  $NE\Delta T$  of high-frequency channels of ATMS, GMI and SAPHIR

are 0.74, 1.31 and 1.32 K respectively. Given the importance of the high-frequency channels for retrieving IWP and detecting DC, the low  $NE\Delta T$  can explain why DEEPSTORM performs better for MHS. In addition, the coincident database of MHS (~13 million of observations) is much larger than the ones of ATMS (~4 million), SAPHIR (~400,000) and GMI (~300,000). This improves the prediction of DEEPSTORM as providing more observations allows the algorithm to adapt to a larger gamut of situations. On the other hand, GMI predictions are on average poorer than the other instruments. This is surprising as one could have expected that the instrument with the highest spatial resolution (5 km for the high frequency channels vs 15 km for MHS) would perform better in retrieving IWP and detecting DC. The explanation could lie in a parallax issue as GMI scans with a constant viewing angle at  $53^\circ$  while CPR scans at nadir.



**Fig. 10. (left) FSE as a function of top altitude of DC for the 4 radiometers and (right) Probability of detection as a function of altitude of ice water content maximum.**

This hypothesis is supported in Fig. 10 (left) which presents the FSE as a function of the maximal altitude reached by DC clouds. We see that when the DC clouds are more elevated the FSE is smaller for MHS, ATMS and SAPHIR but that, for GMI, this relationship is only true until 10 km. Above the FSE remains constant around 70%. This

likely means that when DC clouds reaches high altitudes, more ice is located in higher altitudes and the parallax problem between CPR and GMI becomes acute. This is particularly problematic in the tropics where deep convective clouds can be very elevated. On the other hand, GMI is as good as MHS and NPP for altitudes below 10 km meaning that it is more appropriate for the shallower mid-latitudes DC. This is also why GMI show a better correlation index and a better RMSE in the mid-latitudes than in the tropics (Fig. 4).

In addition, results in Fig. 6 highlight the better DC detection skills in the tropics than in mid-latitudes. This is partially due to the lowest detection rate when convective clouds are shallow. This is illustrated in Fig. 10 (right) where it appears that the probability of detection of DC is considerably reduced when the altitude of maximum of ice is below 3.5 km ( $\text{POD} < 40\%$ ). This figure also shows why it is easier to detect DC in the tropics since no convective clouds with a maximum ice altitude is found below 4 km while those shallow clouds represent 68% of DC cases in the mid-latitudes. We can also notice in Fig. 10 (right) that when the altitude of ice water content maximum is very elevated ( $> 10$  km), convective clouds are more difficult to detect. This means that overshooting clouds could be particularly challenging to detect with passive microwave radiometers.

Finally, it is also interesting to understand why SAPHIR has the lowest accuracy for IWP retrieval. It is certainly related to the fact that SAPHIR only has channels with frequencies close to the water vapour absorption line at 183.3 GHz which probe only in the upper layers of the cloud and miss the possible strong IWC in the lower layers. This is confirmed

by the fact that the variable selection procedure retains, for the other radiometers, lower frequency variables (e.g., 89 GHz) to retrieve IWP.

## 6. Conclusion

In this study, the DEEPSTORM (DEEP moiSt aTmospheric cOnvection from micRowave radioMeter) algorithm, able to detect deep convection (DC) and retrieve ice water path (IWP) using passive radiometer measurements, is presented. In order to train DEEPSTORM, we used collocated measurements between four passive microwave radiometers (ATMS, MHS, SAPHIR and GMI) and the spaceborne radar CPR (Cloud Profiling Radar) on-board CloudSat. Specifically, DEEPSTORM uses brightness temperatures and viewing angle from the passive microwave radiometers, ground elevation, and vertical profiles of specific humidity, relative humidity and temperature from ERA-5 reanalysis. The variables necessary to retrieve IWP and to predict DC occurrence are identified, then a neural network and a gradient boosting algorithm are fitted, tuned and combined to provide the predictions based on 80% of the collocation datasets. Finally, we used the remaining 20% of the data to evaluate the output of DEEPSTORM. Regarding IWP retrievals, DEEPSTORM predictions show a RMSE ranging between 0.21 kg/m<sup>2</sup> (for MHS) and 0.34 kg/m<sup>2</sup> (for SAPHIR) and that the overall accuracy increases strongly with the IWP for all radiometers. In the tropics, DC predictions reach a HSS of 0.72 for MHS and ATMS and 0.67 for GMI and SAPHIR. In the mid-latitudes, predictions are poorer with an average HSS of 0.49. Similarly to IWP retrieval, the POD (FAR) increases (decreases) with IWP. We highlighted that DEEPSTORM detects mainly the most vertically extended convective clouds and those containing the heaviest ice content and

that DC detection (IWP retrieval) is more (less) accurate in the tropics than in mid-latitudes. Eventually, we illustrated DEEPSTORM potential with two applications : by retrieving IWP of Hurricane Matthew using observations of SAPHIR, ATMS, MHS and GMI, then by constructing a global climatology of DC occurrence between September and December 2016. These results illustrate the potential of DEEPSTORM for better characterising the DC distribution and intensity. In addition, as DEEPSTORM can be used for MHS, ATMS, GMI and SAPHIR radiometers, it will be possible to characterise DC occurrence and intensity at a quasi global scale on a 20-year period as the first radiometer of this type have been placed in orbit in 1999 and since then more than 10 instruments have been launched.

DEEPSTORM has several limitations in addition to those presented in the results (e.g., for weak DC events) that are discussed in the following. First, DEEPSTORM fully relies on two products (DARDAR for IWP and 2B-CLDCLASS for DC detection) based on CloudSat CPR radar observations. Yet, those products have several sources of bias which in turn can bias DEEPSTORM algorithm. First, CPR does not probe close to the surface due to ground clutter contamination. Therefore measurements from the lowest layers of the convective clouds can be missed, which could artificially reduce IWP when clouds are close to the Earth's surface. The IWC retrieved by DARDAR is also subjected to uncertainties as CPR measured reflectivity using a single frequency and thus the inversion problem is not fully constrained. This is why DARDAR requires the use of an empirical relationship using the radar reflectivity factor and temperature implying uncertainties in IWC estimates (Delanoë and Hogan, 2008). In addition to limitations associated with CloudSat observations, some uncertainties arise from the building of the coincidence

dataset. Indeed, the aggregation of CPR pixels on the radiometer pixels lead to cases where a radiometer pixel may contain both CPR pixels with and without DC. In these case, we arbitrarily decided that the whole radiometer pixel was associated with DC which could be a factor of uncertainty. The same issue appears with IWP computation : in some cases, ice water content from CPR was highly heterogeneous vertically and horizontally within the corresponding radiometer pixel. This can also affect the retrieval procedure. Finally, the machine learning procedure itself could be a source of uncertainties. Machine learning models, from their design, have some limitations that tend to bias the results. For instance some of them are prone to overfit, some have difficulties to predict extreme values, some do not handle well non-linear relationship between input and output data. Thus no perfect model exists and our choice to select gradient boosting and neural networks model implies some bias in the final results. In order to mitigate this problem, we evaluated several machine learning models and input variable combinations before obtaining the final algorithm described in this article.

Several avenues of improvement will be considered in the future developments of DEEPSTORM. First, it would be interesting to add new observations to extend the training dataset. For instance, it could be interesting to use the DPR (Dual Polarization Radar) on-board GPM-core since it provides observations since 2014 on a 100 km-wide swath with two different frequencies (13.6 and 35.55 GHz). Including this dataset could increase drastically the number of coincidence with radiometer observations. Yet, as we showed with GMI, parallax can be a significant issue when collocating instruments with different viewing angles. In this article, we also highlighted the difficulties to retrieve IWP in the strongest cases. In particular, a negative bias has been identified for all radiometers.

It could therefore be interesting to develop a specific model for the strongest events. We also showed that lower frequency channels can be useful to characterize convection. Therefore, it might be worthwhile to include in DEEPSTORM the measurements of AMSU-A radiometers that are installed on-board the same satellites that carry AMSU-B and MHS. Finally, this work will lead to the building of a long term database of DC occurrence and associated IWP. We will use observations from all radiometers for which DEEPSTORM has been developed (MHS/AMSU-B, SAPHIR, ATMS and GMI). The first instrument has been launched in 1999 and since 2002 at least 3 radiometers fly conjointly which will ensure a good temporal sampling together with a good spatial resolution of the DC on a two decades period. Such a long term database will be of great interest for climatic studies.

### **Acknowledgments and Funding**

This work is supported by the Direction Général de l'Armement (PRECIPCLOUD-SAT project) and by the Centre national d'études spatiales. The authors want to express their gratitude to Julie Carreau for her valuable suggestions and the two reviewers of this article for their constructive and insightful comments. The datasets used in this study are available through the French Mixed Service Unit project ICARE/Climserv/Aeris (ATMS, MHS, MT, DARDAR, ERA5), the NASA PPS website <ftp://arthurhou.pps.eosdis.nasa.gov> (GPM products) and the CloudSat website <ftp.cloudsat.cira.colostate.edu> (CPR products).

### **Tools**

This project has been conducted using the R software (version 3.5 (RCoreTeam, 2013)) with the following packages: rhdf5 (Fischer, 2015), ncdf4 (Pierce, 2012), rpart (Therneau

et al., 2015), data.table (Dowle et al., 2014), hydroGOF (Zambrano-Bigiarini, 2014), nabor (Elseberg et al., 2012), snow (Tierney et al., 2008), snowfall (Knaus, 2010), mlr (Bischl et al., 2016), mlrMBO (Bischl et al., 2017), xgboost (T. Chen et al., 2018), keras (Allaire and Chollet, 2019) and tensorflow (Allaire and Tang, 2019).

## References

- Allaire, J. J., Chollet, F. (2019). *keras: R Interface to “Keras.”* Retrieved from <https://CRAN.R-project.org/package=keras>
- Allaire, J. J., Tang, Y. (2019). *tensorflow: R Interface to “TensorFlow.”* Retrieved from <https://CRAN.R-project.org/package=tensorflow>
- Allen, J. T. (2018). Climate Change and Severe Thunderstorms. In J. T. Allen, *Oxford Research Encyclopedia of Climate Science*. Oxford University Press.  
<https://doi.org/10.1093/acrefore/9780190228620.013.62>
- Aumann, H. H., Ruzmaikin, A. (2013). Frequency of deep convective clouds in the tropical zone from 10 years of AIRS data. *Atmospheric Chemistry and Physics*, 13(21), 10795–10806. <https://doi.org/10.5194/acp-13-10795-2013>
- Bennartz, R., Bauer, P. (2003). Sensitivity of microwave radiances at 85–183 GHz to precipitating ice particles. *Radio Science*, 38(4).  
<https://doi.org/10.1029/2002RS002626>
- Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., et al. (2016). mlr: Machine Learning in R. *Journal of Machine Learning Research*, 17(170), 1–5.
- Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J., Lang, M. (2017). *mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions*.
- Blackwell, W. J., Chidester, L., Kim, E. J., Leslie, R. V., Lyu, C.-H., Mo, T. (2011). NPP ATMS prelaunch performance assessment and sensor data record validation. In *2011 IEEE International Geoscience and Remote Sensing Symposium* (pp. 32–34). IEEE.

- Bonsignori, R. (2007). The Microwave Humidity Sounder (MHS): in-orbit performance assessment. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (Vol. 6744, p. 0). <https://doi.org/10.1117/12.737986>
- Brooks, H. E., Dotzek, N. (2008). The spatial distribution of severe convective storms and an analysis of their secular changes. *Climate Extremes and Society*, *35*, 53.
- Brunner, J. C., Ackerman, S. A., Bachmeier, A. S., Rabin, R. M. (2007). A Quantitative Analysis of the Enhanced-V Feature in Relation to Severe Weather. *Weather and Forecasting*, *22*(4), 853–872. <https://doi.org/10.1175/WAF1022.1>
- Ceccaldi, M., Delanoë, J., Hogan, R. J., Pounder, N. L., Protat, A., Pelon, J. (2013). From CloudSat-CALIPSO to EarthCare: Evolution of the DARDAR cloud classification and its comparison to airborne radar-lidar observations. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7962–7981.
- Cecil, D. J., Blankenship, C. B. (2012). Toward a Global Climatology of Severe Hailstorms as Estimated by Satellite Passive Microwave Imagers. *Journal of Climate*, *25*(2), 687–703. <https://doi.org/10.1175/JCLI-D-11-00130.1>
- Chen, T., Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16* (pp. 785–794). San Francisco, California, USA: ACM Press. <https://doi.org/10.1145/2939672.2939785>
- Clain, G., Brogniez, H., Payne, V. H., John, V. O., Luo, M. (2015). An Assessment of SAPHIR Calibration Using Quality Tropical Soundings. *Journal of Atmospheric and Oceanic Technology*, *32*(1), 61–78. <https://doi.org/10.1175/JTECH-D-14-00054.1>
- Claud, C., Alhammoud, B., Funatsu, B. M., Lebeau-pin Brossier, C., Chaboureau, J.-P., Béranger, K., Drobinski, P. (2012). A high resolution climatology of precipitation and deep convection over the Mediterranean region from operational satellite microwave data: development and application to the evaluation of model uncertainties. *Natural Hazards and Earth System Sciences*, *12*(3), 785–798. <https://doi.org/10.5194/nhess-12-785-2012>

- Copernicus Climate Change Service (C3S) (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate.
- Deeter, M., Vivekanandan, J. (2005). AMSU-B observations of mixed-phase clouds over land. *Journal of Applied Meteorology*, 44(1), 72–85.
- Delanoë, J., Hogan, R. J. (2008). A variational scheme for retrieving ice cloud properties from combined radar, lidar and infrared radiometer. *J. Geophys. Res.*, 113, 1682–1698. <https://doi.org/10.1029/2007JD009000>
- Delanoë, J., Hogan, R. J. (2010). Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research*, 115. <https://doi.org/10.1029/2009JD012346>
- Doswell, C. A. (2001). Severe Convective Storms—An Overview. In C. A. Doswell (Ed.), *Severe Convective Storms* (pp. 1–26). Boston, MA: American Meteorological Society. [https://doi.org/10.1007/978-1-935704-06-5\\_1](https://doi.org/10.1007/978-1-935704-06-5_1)
- Dowle, M., Short, T., Lianoglou, S., Saporta, R., Srinivasan, A., Antonyan, E. (2014). data. table: Extension of data. frame.
- Draper, D. W., Newell, D. A., Wentz, F. J., Krimchansky, S., Skofronick-Jackson, G. M. (2015). The Global Precipitation Measurement (GPM) Microwave Imager (GMI): Instrument Overview and Early On-Orbit Performance. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7), 3452–3462. <https://doi.org/10.1109/JSTARS.2015.2403303>
- Elseberg, J., Magnenat, S., Siegart, R., Nüchter, A. (2012). Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *Journal of Software Engineering for Robotics*, 3(1), 2–12.
- Ferraro, R., Beauchamp, J., Cecil, D., Heymsfield, G. (2015). A prototype hail detection algorithm and hail climatology developed with the advanced microwave sounding unit (AMSU). *Atmospheric Research*, 163, 24–35. <https://doi.org/10.1016/j.atmosres.2014.08.010>
- Fischer, B. (2015). rhdf5-HDF5 interface for R. *R# Package Version*, 2(0).

- Friedman, J., Hastie, T., Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1). Springer series in statistics New York.
- Funatsu, B. M., Claud, C., Chaboureau, J.-P. (2007). Potential of Advanced Microwave Sounding Unit to identify precipitating systems and associated upper-level features in the Mediterranean region: Case studies. *Journal of Geophysical Research*, *112*(D17). <https://doi.org/10.1029/2006JD008297>
- Funatsu, B. M., Claud, C., Chaboureau, J.-P. (2008). A 6-year AMSU-based climatology of upper-level troughs and associated precipitation distribution in the Mediterranean region. *Journal of Geophysical Research*, *113*(D15). <https://doi.org/10.1029/2008JD009918>
- Funatsu, B. M., Claud, C., Chaboureau, J.-P. (2009). Comparison between the Large-Scale Environments of Moderate and Intense Precipitating Systems in the Mediterranean Region. *Monthly Weather Review*, *137*(11), 3933–3959. <https://doi.org/10.1175/2009MWR2922.1>
- Funatsu, B. M., Dubreuil, V., Claud, C., Arvor, D., Gan, M. A. (2012). Convective activity in Mato Grosso state (Brazil) from microwave satellite observations: Comparisons between AMSU and TRMM data sets: CONVECTION IN MATO GROSSO FROM MICROWAVE. *Journal of Geophysical Research: Atmospheres*, *117*(D16), <https://doi.org/10.1029/2011JD017259>
- Funatsu, B. M., Rysman, J.-F., Claud, C., Chaboureau, J.-P. (2018). Deep convective clouds distribution over the Mediterranean region from AMSU-B/MHS observations. *Atmospheric Research*, *207*, 122–135.
- Gen, M., Lin, L. (2007). Genetic algorithms. *Wiley Encyclopedia of Computer Science and Engineering*, 1–15.
- Gers, F. A., Schmidhuber, J., Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.
- Gong, J., Wu, D. L. (2017). Microphysical properties of frozen particles inferred from Global Precipitation Measurement (GPM) Microwave Imager (GMI) polarimetric

measurements. *Atmospheric Chemistry and Physics*, 17(4), 2741–2757.

<https://doi.org/10.5194/acp-17-2741-2017>

- Guimond, S. R., Reasor, P. D., Heymsfield, G. M., McLinden, M. M. (2020). The dynamics of vortex rossby waves and secondary eyewall development in hurricane Matthew (2016): New insights from radar measurements. *Journal of the Atmospheric Sciences*. <https://doi.org/10.1175/JAS-D-19-0284.1>
- Heymsfield, A. J., Schmitt, C., Chen, C.-C.-J., Bansemmer, A., Gettelman, A., Field, P. R., Liu, C. (2020). Contributions of the Liquid and Ice Phases to Global Surface Precipitation: Observations and Global Climate Modeling. *Journal of the Atmospheric Sciences*, 77(8), 2629–2648. <https://doi.org/10.1175/JAS-D-19-0352.1>
- Hoeppe, P. (2016). Trends in weather related disasters – Consequences for insurers and society. *Weather and Climate Extremes*, 11, 70–79. <https://doi.org/10.1016/j.wace.2015.10.002>
- Hong, G., Heygster, G., Miao, J., Kunzi, K. (2005). Detection of tropical deep convective clouds from AMSU-B water vapor channels measurements. *Journal of Geophysical Research*, 110(D5). <https://doi.org/10.1029/2004JD004949>
- Hong, G., Heygster, G., Notholt, J., Buehler, S. A. (2008). Interannual to Diurnal Variations in Tropical and Subtropical Deep Convective Clouds and Convective Overshooting from Seven Years of AMSU-B Measurements. *Journal of Climate*, 21(17), 4168–4189. <https://doi.org/10.1175/2008JCLI1911.1>
- Hong, G., Heygster, G. (2008). Intense Tropical Thunderstorms Detected by the Special Sensor Microwave Imager/Sounder. *IEEE Transactions on Geoscience and Remote Sensing*, 46(4), 996–1005. <https://doi.org/10.1109/TGRS.2008.915750>
- Knaus, J. (2010). snowfall: Easier cluster computing (based on snow). *R Package Version*, 1.
- Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., et al. (2001). The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors. *Journal of Applied Meteorology*,

40(11), 1801–1820. [https://doi.org/10.1175/1520-0450\(2001\)040<1801:TEOTGP>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2)

Laviola, S., Levizzani, V. (2011). The 183-WSL fast rain rate retrieval algorithm: Part I: Retrieval design. *Atmospheric Research*, 99, 443–461.

<https://doi.org/10.1016/j.atmosres.2010.11.013>

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

<https://doi.org/10.1038/nature14539>

Liu, C., Zipser, E. J., Nesbitt, S. W. (2007). Global Distribution of Tropical Deep Convection: Different Perspectives from TRMM Infrared and Radar Data. *Journal of Climate*, 20(3), 489–503. <https://doi.org/10.1175/JCLI4023.1>

Liu, C., Zipser, E. J., Cecil, D. J., Nesbitt, S. W., Sherwood, S. (2008). A Cloud and Precipitation Feature Database from Nine Years of TRMM Observations. *Journal of Applied Meteorology and Climatology*, 47(10), 2712–2728.

<https://doi.org/10.1175/2008JAMC1890.1>

Liu, N., Liu, C. (2016). Global distribution of deep convection reaching tropopause in 1 year GPM observations: TROPOPAUSE-REACHING CONVECTION IN GPM OBSERVATION. *Journal of Geophysical Research: Atmospheres*, 121(8), 3824–3842. <https://doi.org/10.1002/2015JD024430>

Miller, S. D., Straka III, W. C., Yue, J., Seaman, C. J., Xu, S., Elvidge, C. D., et al. (2018). The dark side of hurricane matthew: Unique perspectives from the VIIRS day/night band. *Bulletin of the American Meteorological Society*, 99(12), 2561–2574.

Petković, V., Orescanin, M., Kirstetter, P., Kummerow, C., Ferraro, R. (2019). Enhancing PMW Satellite Precipitation Estimation: Detecting Convective Class. *Journal of Atmospheric and Oceanic Technology*, 36(12), 2349–2363.

<https://doi.org/10.1175/JTECH-D-19-0008.1>

Pierce, D. (2012). ncd4: Interface to Unidata netCDF (version 4 or earlier) format data files. *R Package*, URL [Http://CRAN.R-Project.Org/Package=Ncdf4](http://CRAN.R-Project.Org/Package=Ncdf4).

RCoreTeam. (2013). R: A language and environment for statistical computing.

- Rysman, JF, Panegrossi, G., Sanò, P., Marra, A. C., Dietrich, S., Milani, L., et al. (2019). Retrieving Surface Snowfall With the GPM Microwave Imager: A New Module for the SLALOM Algorithm. *Geophysical Research Letters*, *46*(22), 13593–13601. <https://doi.org/10.1029/2019GL084576>
- Rysman, JF, Claud, C., Chaboureau, J.-P., Delanoë, J., Funatsu, B. M. (2016a). Severe convection in the Mediterranean from microwave observations and a convection-permitting model: Severe Convection in the Mediterranean. *Quarterly Journal of the Royal Meteorological Society*, *142*, 43–55. <https://doi.org/10.1002/qj.2611>
- Rysman, JF, Claud, C., Delanoë, J. (2016b). Monitoring deep convection and convective overshooting from 60 S to 60 N using MHS: a Cloudsat/CALIPSO-based assessment. *IEEE Geoscience and Remote Sensing Letters*, *14*(2), 159–163.
- Rysman, JF, Berthou, S., Claud, C., Drobinski, P., Chaboureau, J.-P., Delanoë, J. (2018). Potential of microwave observations for the evaluation of rainfall and convection in a regional climate model in the frame of HyMeX and MED-CORDEX. *Climate Dynamics*, *51*(3), 837–855. <https://doi.org/10.1007/s00382-016-3203-7>
- Sanò, P., Panegrossi, G., Casella, D., Di Paola, F., Milani, L., Mugnai, A., et al. (2015). The Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for AMSU/MHS observations: description and application to European case studies. *Atmospheric Measurement Techniques*, *8*(2), 837–857. <https://doi.org/10.5194/amt-8-837-2015>
- Sanò, P., Panegrossi, G., Casella, D., Marra, A., D’Adderio, L., Rysman, J., Dietrich, S. (2018). The Passive Microwave Neural Network Precipitation Retrieval (PNPR) Algorithm for the CONICAL Scanning Global Microwave Imager (GMI) Radiometer. *Remote Sensing*, *10*(7), 1122. <https://doi.org/10.3390/rs10071122>
- Sassen, K., Wang, Z. (2008). Classifying clouds around the globe with the CloudSat radar: 1-year of results. *Geophysical Research Letters*, *35*(4)
- Stewart, S. (2017). National Hurricane Center Tropical Cyclone Report Hurricane Matthew (AL142016) 28 September–9 October 2016). *National Hurricane Center*.

*Available Online at [Http://Www. Nhc. Noaa. Gov/Data/Tcr/AL142016\\_Matthew. Pdf](http://www.nhc.noaa.gov/Data/Tcr/AL142016_Matthew.Pdf).*

- Surussavadee, C., Staelin, D. H. (2006). Comparison of AMSU millimeter-wave satellite observations, MM5/TBSCAT predicted radiances, and electromagnetic models for hydrometeors. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), 2667–2678.
- Tang, G., Long, D., Behrangi, A., Wang, C., Hong, Y. (2018). Exploring Deep Neural Networks to Retrieve Rain and Snow in High Latitudes Using Multisensor and Reanalysis Data. *Water Resources Research*, 54(10), 8253–8278.  
<https://doi.org/10.1029/2018WR023830>
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192.  
<https://doi.org/10.1029/2000JD900719>
- Therneau, T., Atkinson, B., Ripley, B. (2015). *rpart: Recursive Partitioning and Regression Trees. R package version 4.1–10*.
- Tierney, L., Rossini, A. J., Li, N., Sevcikova, H. (2008). snow: Simple network of workstations. *R Package Version 0.3-3, URL [Http://CRAN. R-Project. Org/Package= Snow](http://CRAN.R-Project.Org/Package=snow)*.
- Wu, X., Yuan, T., Qie, K., Luo, J. (2020). Geographical distribution of extreme deep and intense convective storms on Earth. *Atmospheric Research*, 235, 104789.  
<https://doi.org/10.1016/j.atmosres.2019.104789>
- Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., Chen, Z. (2019). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sensing of Environment*, 233, 111358.  
<https://doi.org/10.1016/j.rse.2019.111358>
- Zambrano-Bigiarini, M. (2014). hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. *R Package Version 0.3-8*.

Supplementary material

| Radiometer | Variables selected   |
|------------|--|
| MHS        | viewing angle, elevation, R125, q650, q700, q750, q1000, <b>BT157</b> , BT183.31 $\pm$ 3.0, BT190.311, BT89                                    |
| ATMS       | viewing angle, elevation, q300, q600, q775, q1000, BT89.5, <b>BT165.5</b> , BT183.31 $\pm$ 7.0, BT183.31 $\pm$ 4.5, BT183.31 $\pm$ 3.0, BT50.3 |
| SAPHIR     | viewing angle, <b>BT183.31 <math>\pm</math> 11</b> , BT183.31 $\pm$ 4.2  |
| GMI        | Elevation, tp2, R400, q700, q975, BT10.65V, BT166.5H, <b>BT183.31 <math>\pm</math> 7</b> , BT18.7V, BT36.5H, BT89V                             |

Table 1. Variables needed to retrieve IWP for each radiometer. R stands for relative humidity, q stands for specific humidity, tp stands for temperature.

All these variables are followed by a number that indicates the corresponding pressure level (hPa). BT stands for brightness temperature and is followed by the corresponding frequency (Ghz) and the polarisation (for GMI). The most informative variables are indicated in red.

| Radiometer | Variables selected  |
|------------|---|
| MHS        | viewing angle, elevation, tp7, tp300, R1000, q2, q1000, <b>BT157</b> , BT183.31 $\pm$ 3.0, BT190.311, BT89                                |
| ATMS       | viewing angle, tp975, q100, q700, q950, q975, <b>BT165.5</b> , BT183.31 $\pm$ 4.5, BT183.31 $\pm$ 3.0, BT183.31 $\pm$ 1.8, BT23.8, BT50.3 |
| SAPHIR     | viewing angle, R500, R600, R750, q925, <b>BT183.31 <math>\pm</math> 11</b> , BT183.31 $\pm$ 6.6, BT183.31 $\pm$ 4.2, BT183.31 $\pm$ 2.7   |

|     |  |
|-----|--|
| GMI | R450, q925, q1000, BT10.65V, <b>BT166.5H</b> ,<br>BT183.31 ± 3, BT183.31 ± 7, BT36.5H,<br>BT36.5V, BT89V |
|-----|--|

Table 2 Variables needed to detect DC for each radiometer. R stands for relative humidity, q stands for specific humidity, tp stands for temperature.

All these variables are followed by a number that indicates the corresponding pressure level (hPa). BT stands for brightness temperature and is followed by the corresponding frequency (Ghz) and the polarisation (for GMI). The most informative variables are indicated in red.