The 183-WSL fast rain rate retrieval algorithm. Part II: Validation using ground radar measurements

Sante Laviola[a], Vincenzo Levizzani[a], Elsa Cattani[a], Chris Kidd[b]

[a]*CNR-ISAC, Bologna, Italy*
[b]*NASA-GSFC and Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA*

**Corresponding author:* CNR-ISAC, via Gobetti 101, I-40129 Bologna, Italy. Tel: +39-051-6398019. Fax: +39-051-6399658. *E-mail address*: s.laviola@isac.cnr.it.

Manuscript submitted to
Atmospheric Research

ARTICLE

ABSTRACT

The Water vapour Strong Lines at 183 GHz (183-WSL) algorithm is a method for the retrieval of rain rates and precipitation type classification (convective/stratiform), that makes use of the water vapor absorption lines centered at 183.31 GHz of the Advanced Microwave Sounding Unit module B (AMSU-B) and of the Microwave Humidity Sounder (MHS) flying on NOAA-15-18 and NOAA-19/Metop-A satellite series, respectively. The characteristics of this algorithm were described in Part I of this paper together with comparisons against analogous precipitation products. The focus of Part II is the analysis of the performance of the 183-WSL technique based on surface radar measurements. The "ground truth" dataset consists of 2.5 years of rainfall intensity fields from the NIMROD European radar network which covers North-Western Europe. The investigation of the 183-WSL retrieval performance is based on a twofold approach: 1) the dichotomous statistic is used to evaluate the capabilities of the method to identify rain and no-rain clouds; 2) the accuracy statistic is applied to quantify the errors in the estimation of rain rates.

The results reveal that the 183-WSL technique shows good skills in the detection of rain/no-rain areas and in the quantification of rain rate intensities. The categorical analysis shows annual values of the POD, FAR and HK indices varying in the range 0.80-0.82, 0.33–0.36 and 0.39-0.46, respectively. The RMSE value is 2.8 mm h^{-1} for the whole period despite an overestimation in the retrieved rain rates. Of note is the distribution of the 183-WSL monthly mean rain rate with respect to radar: the seasonal fluctuations of the average rainfalls measured by radar are reproduced by the 183-WSL. However, the retrieval method appears to suffer for the winter seasonal conditions especially when the soil is partially frozen and the surface emissivity drastically changes. This fact is verified observing the discrepancy distribution diagrams where

the 183-WSL performs better during the warm months, while during the winter time the discrepancies with radar measurements tends to maximum values. A stable behavior of the 183-WSL algorithm is demonstrated over the whole study period with an overall overestimation for rain rates intensities lower than 1 mm h^{-1}. This threshold is crucial especially in wintertime where the low precipitation regime is difficult to be classified.

Keywords: Microwave, precipitation, satellites, validationradar, winter.

1. Introduction

A number of studies have tackled the problem of quantifying the accuracy of satellite-based rainfall estimation methods using ground measurements (gauges to radars) or the spaceborne Precipitation Radar (PR) onboard the Tropical Rainfall Measuring Mission (TRMM). As pointed out by Michaelides et al. (2009) the verification task must account for whether general rainfall statistics or accurate location and timing is assessed. This is necessary for devising applications of satellite-based precipitation estimations for a wide variety of expanding applications (e.g., Kucera et al., 2012). Several methods were proposed for the verification of these retrievals and an extended description can be found in Ebert (2007, 2011). For the purpose of this study it is worth noting a few findings of recent validation studies. Lin and Hou (2008) showed that over land the AMSU-B rain retrievals are comparable in quality to those from conically scanning radiometers for instantaneous rain rates between 1.0 and 10.0 mm h^{-1}. Lin and Hou (2012) noted that for passive microwave sensors with pixel sizes ranging from 14 to 16 km the minimum detectable rain rates are around 1 mm h^{-1}. Zhou et al. (2008) found that satellite products over China are comparable to rain gauge data in revealing the spatial patterns of JJA precipitation amount, frequency, and intensity, with pattern correlation coefficients ranging from 0.66 to 0.94. Tian et al. (2009) reported, among other things, that the satellite rainfall estimation methods in general miss a significant amount of light precipitation up to 40%. Shen et al. (2010) found that the performance of the satellite products varies for different regions and different precipitation regimes, with better comparison statistics observed over wet regions and for warm seasons. Finally, Kidd et al. (2012) conducted an extensive validation of high resolution satellite precipitation products over northwest Europe finding that they exhibit a seasonal cycle in

correlation, bias ratio, probability of detection, and false alarm ratio, with poorer statistics during the winter.

In Part 1 of this paper (Laviola and Levizzani, 2011) the retrieval design of the Water vapour Strong Lines at 183 GHz (183-WSL) method is described. The method is conceived for the Advanced Microwave Sounding Unit-B/Microwave Humidity Sounder (AMSU-B/MHS) sensors currently flying on the National Ocean and Atmospheric Administration (NOAA) and MetOp satellites. The detection of the rain systems and the retrieved rainfall rates were shown to be in qualitative agreement with the TRMM 2A12 (Kummerow at al., 1998) and the Goddard Profiling Algorithm (GPROF) (Kummerow at al., 2001) products. Laviola (2011) and Laviola et al. (2012) also presented a first validation of the 183-WSL estimation capabilities at midlatitudes and over the Baltic Regions, where the experimental module for the snow cover detection was also tested.

The validation results are detailed over a 2.5-year period (hereafter called “three years”). The analysis is based on the surface rainfall fields measured by the NIMROD radar network over North-Western Europe during 2006, 2007 and 7 months of 2008. The spatial domain covers various precipitation regimes, including light stratiform precipitation, snowy precipitation during the winter storms, frontal rain bands, and orographic convection. The relatively long temporal extent of the analysis is thus exploited to investigate the algorithm sensitivity to the different types of precipitation and to explore the seasonal and intra-seasonal variations. The 183-WSL performances in terms of detection skill and rain rate quantification is assessed following a standard approach (Wilks, 1995): 1) categorical statistics used to evaluate the algorithm capabilities to discern rain from no-rain areas; 2) accuracy statistics to quantify the degree of reliability of the 183-WSL estimates by computing the numerical errors associated to the retrieved rain rates. The analysis is completed with the sensitivity evaluation to light rain rate

regimes considering the threshold at 1 mm h^{-1}. Although the 1 mm $h^{-}1$ threshold is often used to mark the accuracy threshold limit of the satellite-based rain retrieval methods, in this study it is adopted to separate light from moderate rain rates (e.g., Lin and Hou, 2012) in order to assess the algorithm skills in terms of total retrieved rainfall and rain type as classified by the 183-WSL modules for the detection of convective and stratiform rainfall.

2. Validation of the 183-WSL performances

In this section the validation methodologies and corresponding results of the comparison between the 183-WSL rainfall fields and the NIMROD surface measurements used as ground "truth" (Sauvageot, 1994) are described. NIMROD is an European C-band radar network for weather analysis and nowcasting, which covers the United Kingdom, France, Germany, the Netherlands, and Belgium (Golding, 1998, 2000; Harrison et al., 1998). All measurements are calibrated and corrected for unrealistic rain values and spurious effects such as ground clutter, and are available for each radar location or as image composite every 5 or 15 minutes and at a nominal 1 or 5 km spatial resolution, respectively. The archive is managed by the British Atmospheric Data Center (BADC) and all measurements are archived in a binary format (UK Met Office, BADC: http://badc.nerc.ac.uk/data/nimrod). In this study the 15-minute composite dataset at 5 km grid is used over the domain reported in Fig. 1; during the processing steps all data were remapped over a 16 km × 16 km grid to be compatible with the spatial resolution of the 183-WSL rain product.

2.1. Categorical statistics

The categorical statistics are applied to verify the ability of the 183-WSL algorithm to reproduce the precipitation spatial domain observed by radar measurements. Table 1 reports the number of rain/no-rain pixel pairs used in the analysis. Conditional indices representing the algorithm performance were calculated. The categorical verification data are presented in a 2 × 2 contingency matrix (Table 2) where all possible combinations of the estimated-observed pairs define the basis for the statistical skill score calculation:

- False Alarm Ratio: $FAR = \frac{b}{(a+b)}$
- Probability Of Detection: $POD = \frac{a}{(a+c)}$
- Frequency Bias Score [0 to∞, 1 = unbiased]: $FBS = \frac{(a+b)}{(a+c)}$
- Critical Success Index: $CSI = \frac{a}{(a+b+c)}$
- Percentage of Corrects: $PC = \frac{(a+d)}{n}$
- Probability Of Rejection: $POR = \frac{d}{(b+d)}$
- False Rejection Rate: $FRR = \frac{c}{(c+d)}$
- Hanssen-Kuiper discriminant [-1;1] (1 = perfect score): $HK = \frac{a \cdot b - b \cdot d}{(a+c) \cdot (b+d)}$

An account of the hits, false alarms, missies and correct negatives referring to the entire statistics is shown in Table 3.

The statistical scores in Table 4 show that the algorithm results are substantially coherent among themselves revealing a systematic good performance during the whole study period. The FAR and POD, that measure the events misclassified as true and those successfully estimated as true, are a measure of the overall performances. For the 183-WSL algorithm low FAR values (0.34, 0.33, and 0.36) and high POD values (0.80, 0.82, and 0.80) were found for 2006, 2007 and 2008, respectively. The values of other categorical indices are as follows for the three years in sequence: 0.72, 0.71 and 0.69 for PC; 0.65, 0.61 and 0.59 for POR; 0.20, 0.22 and 0.24 for FRR. These latter scores demonstrate that the method effectively distinguishes rain from no rain episodes and support the FAR and POD indications. In fact, using the FRR index as an overall indicator, the percentage of false rejected data comes out very low when compared with the magnitude of the PC index. Furthermore, the CSI, which unlike FAR and POD takes into account both false alarms and missed events, reinforces the above findings showing that the values of CSI are always greater than 50%. The categorical analysis can be summarized by the HK index, which evaluates the capability of the method to distinguish “events” from “no-events”.. Over the three years the index ranges from 0.46 for 2006 to 0.39 for the seven months of 2008. The dichotomous statistics also reveals that the 183-WSL algorithm tends to consistently overestimate the retrieved precipitating areas, as demonstrated by the FBS index, which evaluates the ratio between estimated and observed “events” and regularly exhibits values greater than 1.

The dichotomous analysis highlights the properties of the algorithm correlated to a relatively large dataset where the possible seasonal fluctuations are drastically smoothed. To investigate the algorithm sensitivity to the seasonal conditions the three year dataset is organized into four

seasons. Fig. 2 shows the histograms of FAR, POD, HK and FBS for winter and summer, being representative of very diverse weather conditions during the year . In general, the 183-WSL algorithm shows the same behavior during the two seasons throughout the study period. The FAR index is slightly better during the winter time, while the FBS is higher during the warm season. However, the FBS is slightly higher than 1, which represents the unbiased condition, thus revealing a very low bias of the algorithm throughout the whole validation period. In the first case, this is due to the screening procedure of the 183-WSL, which screens out some of the false rainy signals during the cold months mostly caused by frozen soils or sea ice. In summer time the observed overestimation can be justified with the presence of a systematic error, which increases with precipitation intensity. In other words, the less stringent threshold on snow/ice surfaces applied in summer with respect to winter allows for non-precipitating clouds to pass through the screening procedure of the algorithm and be misclassified as precipitation, hence the overestimation. However, note that the FAR and FBS values in summer are not significantly higher than those in winter, thus indicating an overall stable behavior of the algorithm throughout the year. This latter fact is confirmed by POD and HK that appear to be stable with values greater than 0.78 and 0.45, respectively.

2.2. Accuracy statistics

The quantification of the algorithm performance in estimating the rainfall rates is done by the accuracy statistics, which quantifies the errors associated with the algorithm numerical output. By considering N as the total number of data (see Table 1), RR^{sat} and RR^{NRD} as the rain rates for satellite and radar pixels, respectively, the following scalars were calculated:

- Mean value (M): $M = \frac{1}{N}\sum_{i=1}^{N} RR_i^{sat}$
- Mean Error (ME): $ME = \frac{1}{N}\sum_{i=1}^{N}\left(RR_i^{sat} - RR_i^{NRD}\right)$
- Mean Absolute Error (MAE): $MAE = \frac{1}{N}\sum_{i=1}^{N}\left|RR_i^{sat} - RR_i^{NRD}\right|$
- Root Mean Square Error (RMSE): $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(RR_i^{sat} - RR_i^{NRD}\right)^2}$

The statistical results summarized in Table 5 show that all the accuracy error figures are reasonably low with an increasing magnitude from ME to RMSE. However, the value of ME, which basically measures the bias value of the estimated rainfall, as compared with the mean value (M) of the radar rain rates is very small during the study period. The same holds for MAE that is the average over the verification sample of the absolute values of the differences between the rain retrievals and the corresponding observed average radar rain values (M= 1.46 , 1.45 , and 1.37 mm h^{-1}).

Even though the ME gives a rough measure of the mean error distribution with respect to other indices, it describes the overall overestimation of the 183-WSL method during the validation period. Depending on the geographical location and seasonal conditions an average overestimation in the range 0.5 → 1.4 mm h^{-1} is found, which is coherent with the ME values. A more complete investigation of the error distribution in terms of differences between satellite and radar rain rates is shown in Fig. 3. The radar-satellite difference trend confirms the discussed systematic overestimation and that the rain rate quantitative retrievals are closer to the radar values during the warm months. These results are justified by considering that the winter threshold test reduces the possible false alarms in the cold season acting directly on the brightness temperatures before the actual estimation of rain rate. As shown in Fig. 2, however,

the winter FAR values are around 0.3, which suggests that a residual number of pixels escape the thresholding. These latter pixels, when processed by the rainfall estimation module, give raise to a very high precipitation intensity that is by no means real and due to the scattering from the snow covered surface. This generates the observed overestimation of the retrieved rain rates. On the contrary, during the summer time when the screening module is relaxed, with respect to the winter period, the algorithm processes a larger number of pixels (higher FAR), but these pixels that survive the thresholding are less influenced by the surface than in the winter. This induces a better quantitative rain rate estimation by the 183-WSL as seen in Fig. 3.

The RMSE is very sensitive to the large errors and increases from zero to large positive values as a function of increasing discrepancies between estimates and observations. The RMSE values can be assumed as the maximum that the 183-WSL algorithm can theoretically achieve. The monthly distribution of M, ME, and RMSE are presented in Fig. 4. The diagrams in the left column compare the mean rain rate values of the 183-WSL with those of NIMROD showing that the retrieval method follows the radar rainfall measurements accurately reproducing the seasonal trend. The surface rainfall measured by the radar increase during the summertime from the minimum values observed for January to the maximum intensities measured in June-July. The ME diagrams in the middle column of Fig. 4 determine the previously discussed overestimation. The RMSE follows the same trend the same tendency of M monthly values while exhibiting an estimation error, which increases with increasing precipitation.

2.3. Sensitivity analysis with a 1 mm h^{-1} threshold test

Recent studies by Laviola et al. (2012) supported by previous findings (Ebert et al., 2007; Dinku et al., 2010) hint towards the adoption of the rain intensity value at 1 mm h^{-1} as a good indicator to evaluate the algorithm sensitivity to low rain intensities. This threshold represents an objective

limit for the satellite rain retrieval algorithms, but it represents a good compromise for the discrimination between light and moderate precipitation at mid-latitudes. The sensitivity analysis compares the 183-WSL estimates with the NIMROD measurements by considering two intensity classes calculated on the basis of radar-derived rain rates: class-1 ($0 < RR < 1$ mm h^{-1}); class-2 ($RR \geq 1$ mm h^{-1}). Table 6 shows the percentage of rain-bearing pixels per each of the two intensity classes. Note that the maximum values correspond to the estimated rain rates greater than 1 mm h^{-1} thus confirming the overestimation spotted by the accuracy statistics. In particular, to only about 20% (20.47% for 2006, 21% for 2007, and 19.09 for 2008) of the pixels are correctly classified by the 183-WSL when radar observes rainfalls in class-1, i.e. for rain rates lower than 1 mm h^{-1}. The rest of rainfall is misclassified class-2, i.e. among rain rates higher than 1 mm h^{-1} (56.97% for 2006, 58.45 for 2007, and 57.78 for 2008). Further analysis on the algorithm (not reported here) have demonstrated that its sensitivity significantly decreases for rain rates lower than 0.5 mm h^{-1}, when the retrieval reliability becomes very low, thus suggesting to set this value as the effective limit of the method.

When considering class-2 the algorithm performance notably increase showing a low percentage of misclassification (13.79% for 2006, 13.49% for 2007, and 12.44% for 2008) while the majority of pixels are correctly classified as rain rates > 1 mm h^{-1} (71.08% for 2006, 72.05% for 2007, and 70.10% for 2008).

The sensitivity threshold of 1 mm h^{-1} is depicted on the discrepancy diagrams in Fig. 3 (solid line). The discrepancy between radar and satellite of the monthly mean rain rates is seasonally dependent: the maximum discrepancy is found during the winter months, while the displacement values reach the minimum during the summer months. It is conceivable that in wintertime the atmospheric profile associated with the possibility of frozen soil severely affects the rain

estimation in the low intensity domain. It can thus be concluded that the threshold at 1 mm h^{-1} has two major implications: 1) it marks the sensitivity limit between light and very light/absence of precipitation, and 2) it suggests that the seasonal conditions affect the performance of the method and in turn this reveals the impact on the classification of rain type.

3. Precipitation type classification

The 183-WSL method differentiates the total rain rates in stratiform (183-WSLS module) and convective (183-WSLC module) types. An analysis of the two contributions for each season is useful to delineate the algorithm strengths and weakness by considering the natural interannual variations. In other words, the link between the precipitation type categories and the estimated rain rates need to be explored by taking into account that stratiform and convective precipitation are often associated to light/moderate and moderate/intense rain rates, respectively. Therefore, keeping in mind that the physical structure of the 183-WSL method is based on the emitted radiation by rain drops, the more "direct" rain retrieval inferred during the stratiform events can be assessed with respect to convective storms, where the scattered radiation by frozen hydrometeors extinguishes the upwelling radiation emitted by the rain volume.

The analysis is based on twenty intensity classes each 1 mm h^{-1} wide and is performed by counting the number of rainy pixels belonging to each class. The resulting histograms give an indication on the dynamics of the method in the estimation of rain rates providing at same time an evaluation of the impact of seasonal effects. Fig. 5 shows the mean values of the rain rate distribution during the four seasons for each of the three years: December-January-February (DJF), March-April-May (MAM), June-July-August (JJA), and September-October-November (SON). Note that only the cold months (DJF and partially MAM) present a separation between

the distribution lines for the intensity classes higher than 8 mm h^{-1}. It needs to be considered that at the latitudes of the NIMROD network rain rates higher than 10 mm h^{-1} are rarely measured during the winter/springtime so that the high intensity classes are often populated by a few elements which renders them almost insignificant from the statistical point of view. During JJA all lines overlap describing an uniformity in the retrieval and a good dynamic range in the estimated rain rates (all classes are densely populated). This homogeneity means that the summer conditions with more or less intense convection, which favor more intense rains, are well described in such ideal working conditions for the algorithm and the retrieval results are thus consistent for all three years.

Finally, for SON, when only the 2006 and 2007 data were available the distribution of rain shows a similar behavior and the small differences can be neglected.

The examination of the diagrams in Fig. 5 highlights the seasonal dependence of the 183-WSL method and suggests that the method is strongly responsive to the increasing precipitation rates. Therefore, it is interesting to analyze the capabilities of the algorithm in discriminating the different types of rain both in the winter and the summer time. Once more, two seasons are selected where the maximum differences in the precipitation regime are likely to occur.

The analysis is summarized in Fig. 6. Considering the physical basis of the method, the variation of the temperature and humidity profiles can impact on the 183-WSL retrievals especially at high latitudes where the surface may be frozen for a long time. The diagrams display the differences between the two distributions: while for the 183-WSLS the two lines almost entirely overlap, for the 183-WSLC they are clearly separated. The 183-WSL classifies the low rain rate pixels as stratiform both in winter and summer (see right column of Fig. 6) by populating the classes up to 5-6 mm h^{-1}, considering that the majority of precipitation observed

by radar is lower than 5 mm h^{-1} at these latitudes. On the contrary, the convective precipitation being generally more intense than stratiform and often characterized by the presence of vertical development and ice hydrometeors at cloud top, the more intense precipitation classes are more populated than in the stratiform case. At the same time, the algorithm shows a seasonal dependence that reflects in the separation of the two curves on the left of Fig. 6. This is due to the scattering of ice at cloud top that is more relevant in summer when intense convective events are more frequent. Also, the freezing level in summer is located ≅ 4 km while in winter it is very variable, but likely to be between 0 and 2 km, thus affecting the depth of the ice within the troposphere.

In summary, the 183-WSL rain product combines the strengths of the modules 183-WSLC and 183-WSLS and detects the full range of precipitation intensities from a few mm h^{-1} generally associated to winter rain rates to the heavy rains during the summer severe storms.

4. Summary and conclusions

Three years of rainfall intensities from the NIMROD European radar network were used to validate the 183-WSL satellite rain retrieval method and evaluate its sensitivity to different precipitation types. The validation is done at the instantaneous scale by using pixel-by-pixel rainfall rates with a spatial resolution of 0.15° latitude/longitude over the area between 40º and 65º N and 12° W and 20º E.

The statistics were computed to assess the algorithm performance, both to identify the precipitating areas and to quantify the estimation errors in the rain rate retrievals. The results demonstrate the high sensitivity of the 183-WSL in detecting the occurrence of rainfall showing high skill scores throughout the study period. The FAR and POD indices exhibit values around

0.30 and 0.80 respectively, while the HK is over 0.40. Nevertheless, the value of the FBS reveals that the method tends to overestimate the number of pixels with precipitation. The overestimation observed in the categorical analysis is confirmed by the accuracy statistics where the error is associated with the algorithm products. The scalars used to quantify the displacement from the radar measurements reach values in the range 0.9-1.1 mm h^{-1} for ME to 2.9-3.0 mm h^{-1} for the RMSE, which imply an average discrepancy of about 1 mm h^{-1} in terms of mean rain rates.

Furthermore, as a consequence of the accuracy statistics and bearing in mind the physical characteristics of the retrieval method, its capabilities to retrieve very low rain intensities were explored by using the sensitivity threshold at 1 mm h^{-1}, which can be assumed as a valid detection threshold for satellite retrieval algorithms. Results demonstrate that there is a overestimation over 50% for rainfall rates lower than 1 mm h^{-1} and good performances are found for precipitation intensities greater than 1 mm h^{-1}.

A further investigation in terms of discrepancy values clarifies that the weakness of the algorithm increase during the winter season when the atmospheric conditions impact more on the water vapor distribution thus drastically changing the weighting functions of channels used in the algorithm. In the summer time the retrieval skills significantly increase showing a discrepancy values lower than 0.5 mm h^{-1}. However, even when the algorithm exhibits its worse performances the displacement with the ground truth is lower than 1.5 mm h^{-1}.

Finally, since the 183-WSL algorithm classifies the precipitation type, the impact of the seasonal conditions on the classification modules was also investigated. The stratiform precipitation shows a similar distribution both in the winter and in the summer while the convective rain classification reveals a better dynamic range distinguishing the rain rates in the

two seasons. This is related to the different genesis of these type of rain. Stratiform rain being generally associated to low rain intensities forces the algorithm to work more in the emission than the scattering mode so that the retrieved rain rates are more directly linked to the extinction process of radiation by the rain volume and are less contrasted with the background. Therefore, this low perturbation of the nominal value of the radiation induces the retrieval of very low rain rates. On the contrary, during convective weather the rainfall is largely formed by melting of ice hydrometeors and the high precipitation intensity is correlated to the scattering of ice, which strongly contrasts with the background radiation.

An important topic of future studies will be the recalibration of the 183-WSL equations on the basis of the present results with the aim to reduce the errors of the current version. A possible development avenue is indicated by the study of Xie and Xiong (2011) who have developed a conceptual model to correct and remove the bias error from the satellite-based rainfall estimations by merging in a gauge-based analysis. Furthermore, a series of experiments during the wintertime at mid and high latitudes is planned in order to better evaluate the impact of the frozen soil and solid precipitation on the final estimation, which is now currently not well defined.

Acknowledgments

The work was supported by EUMETSAT's "Satellite Application Facility on support to Hydrology and Operational Water Management", and by the European Commission 7th Framework Programme FP7-2010-1.1-04 GMES project Global Water Scarcity Information Service (GLOWASIS) GA No. 262255.

5. References

Dinku, T., Ruiz, F., Connor, S. J., Ceccato, P., 2010. Validation and intercomparison of satellite rainfall estimates over Columbia. J. Appl. Meteor. Climatol. 49, 1004-1014.

Ebert, E. E., 2007. Methods for verifying satellite precipitation estimates. In: Measuring precipitation from space. Levizzani, V., Bauer, P., Turk, F. J. (Eds), Advances in Global Change Research, 28, Springer, p. 345-356.

Ebert, E. E., 2011. Validation / intercomparison of daily satellite precipitation estimates -- An IPWG project. [available at http://cawcr.gov.au/projects/SatRainVal/validation-intercomparison.html]

Ebert, E. E., Janowiak, J. E., Kidd, C., 2007. Comparison of near real time precipitation estimates from satellite observations and numerical models. Bull. Amer. Meteor. Soc. 88, 47-64.

Golding, B. W., 1998. Nimrod: A system for generating automated very short range forecasts. Meterorol. Appl. 5, 1-16.

Golding, B. W., 2000. Quantitative precipitation forecasting in the UK. J. Hydrol. 239, 286-305.

Harrison, D. L., Driscoll, S. J., Kitchen, M., 1998. Improving precipitation estimates from weather radar using quality control and correction techniques. Observation Based Products, MetOffice, Technical Report No.13, p. 10. [available at http://badc.nerc.ac.uk/cgi-bin/data_browser/data_browser/badc/ukmo-nimrod/doc/nimrod_radar_processing.pdf]

Kidd, C., Bauer, P., Turk, F. J., Huffman, G. J., Joyce, R., Hsu, K.-L., Braithwaite, D., 2012. Intercomparison of high-resolution precipitation products over Northwest Europe. J. Hydrometeor. 13, 67-83.

Kucera, P. A., Ebert, E. E., Turk, F. J., Levizzani, V., Kirschbaum, D., Tapiador, F. J., Loew, A., Borsche, M., 2012. Precipitation from space: Advancing Earth system science. Bull. Amer. Meteor. Soc., doi:10.1175/BAMS-D-11-00171.1.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., Simpson, J., 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. J. Atmos. Oceanic Technol. 15, 809-817.

Kummerow, C. D., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., Ferraro, R., Petty, G., Shin, D.-B., Wilheit, T. T., 2001. The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. J. Appl. Meteor. 40, 1801-1820.

Laviola, S., 2011. Validation and evaluation of different AVHRR and AMSU/MHS based precipitation retrieval algorithms. Nowcasting-SAF Visiting Scientist Activity Final Report,

EUMETSAT-SMHI, 35 pp. [available at https://www.nwcsaf.org/HD/files/vsadoc/VSA_technical_report_20110906.pdf]

Laviola, S., Levizzani, V., 2011. The 183-WSL fast rain rate retrieval algorithm. Part I: Retrieval design. Atmos. Res. 99, 443-461.

Laviola, S., Levizzani, V., Cattani, E. Kidd, C., 2012. First validation of retrieved rain rates and snow cover mask of the 183-WSL retrieval method. 12th Specialist Meeting of Microwave Radiometry and Remote Sensing of the Environment, IEEE, E-ISBN doi:10.1109/MicroRad.2012.6185242.

Lin, X., Hou, A. Y., 2008. Evaluation of coincident passive microwave rainfall estimates using TRMM PR and ground measurements as references. J. Appl. Meteor. Climatol. 47, 3170-3187.

Lin, X., Hou, A. Y., 2012. Estimation of rain intensity spectra over the Continental United States using ground radar–gauge measurements. J. Climate 25, 1901-1915.

Michaelides, S., Levizzani, V., Anagnostou, E. N., Bauer, P., Kasparis, T., Lane, J. E., 2009. Precipitation: measurement, remote sensing, climatology and modeling. Atmos. Res. 94, 512-533.

Sauvageot, H., 1994: The probability density function of rain rate and the estimation of rainfall by area integrals. J. Appl. Meteor. 33, 1255–1262.

Shen, Y., Xiong, A., Wang, Y., Xie, P., 2010. Performance of high-resolution satellite precipitation products over China. J. Geophys. Res. 115, D02114, doi:02110.01029/02009JD012097.

Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., Hsu, K.-L., Turk, F. J., Garcia, M., Zeng, J., 2009. Component analysis of errors in satellite-based precipitation estimates. J. Geophys. Res. 114, D24101, doi:24110.21029/22009JD011949.

Wilks, D. S., 1995. Statistical methods in the atmospheric sciences. An introduction. Academic Press, San Diego, 467 pp.

Xie, P. P, Xiong, A. Y., 2011. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. J. Geophys. Res. 116, D21106, doi: 10.1029/ 2011JD016118.

Zhou, T., Yu, R., Chen, H., Dai, A., Pan, Y., 2008. Summer precipitation frequency, intensity, and diurnal cycle over China: A comparison of satellite data with rain gauge observations. J. Climate 21, 3997-4010.

Table captions

Table 1. Number of satellite-derived rain/no-rain pairs for each year used in the statistical analysis.

Table 2. Contingency table with *a*=hits, *b*=false alarms, *c*=misses and *d*=correct negatives. Partial sums are the *marginal total* of two data sets and *n* is the sample size.

Table 3. Synopsis of hits, false alarms, misses and correct negatives of the whole validation dataset.

Table 4. Categorical statistics comparing the performances of pixel-by-pixel rain data.

Table 5 Pixel-by-pixel accuracy statistics by comparison of the rain data pairs from Table 1. Top to bottom the mean error (ME), the mean absolute error (MAE), the root mean square error (RMSE) and the mean are listed in mm h^{-1}. The last row indicates the mean values of 183-WSL/radar data for the whole study period.

Table 6. Sensitivity analysis with the 1 mm h^{-1} threshold test.

Figure captions

Fig.1. NIMROD weather radar network spatial coverage.

Fig 2. Mean values of the most significant categorical indices for winter (black) and summer (dotted) for the 3-year dataset. The FAR index is slightly smaller during the winter time where the screening threshold tests of the 183-WSL mitigate the false alarms due to the frozen soil. On the other hand, the Bias score increases in the warm season. The POD and HK indices are more or less stable in both seasons.

Fig 3. Monthly mean of the radar-satellite rain rate discrepancies for 2006 (top), 2007 (middle), and 2008 (bottom). The 183-WSL performs better during the summer season when the average value of discrepancies is of -0.69 mm h^{-1} with respect to the winter when the same values becomes -1.35 mm h^{-1}. The solid line marking the 1 mm h^{-1} threshold refers to the algorithm performance limit as discussed in section 2.

Fig 4. Monthly distribution of M, ME, and RMSE for 2006 (top), 2007 (middle), and 2008 (bottom).

Fig 5. Intensity classes for the 183-WSL mean rain rates during the three years: DJF (top-left), MAM (top-right), JJA (bottom-left), and SON (bottom-right). The distribution of the retrieved rain rates for the DJF months is divided in two intensity regions: the first where a similar trend for the three years is found up to 7-8 mm h^{-1} and the second where different rain distribution for high intensity rain is found. The differences in this second class are lower in MAM. During the warm months, when precipitation is often associated to convective characteristics, the retrieved rain rates display no evident difference for all intensity classes.

Fig 6. Mean rain rate intensity classes classified by the convective (183-WSLC, left) and stratiform (183-WSLS, right) retrieval modules during the winter and summer seasons 2006 (top), 2007 (middle), and 2008 (bottom). Note the low number of pixels belonging to high intensity classes.

Tables

Table 1. Number of satellite-derived rain/no-rain pairs for each year used in the statistical analysis.

Year	No-Rain	Rain	Tot. data
2006	468681	386840	855521
2007	426859	411224	838083
2008 (7 months)	265127	291441	556668

Table 2. Contingency table with *a*=hits, *b*=false alarms, *c*=misses and *d*=correct negatives. Partial sums are the *marginal total* of two data sets and *n* is the sample size.

		OBSERVED		
		YES	NO	
ESTIMATE	YES	a	b	a+b
	NO	c	d	c+d
		a+c	b+d	n=a+b+c+d

Table 3. Synopsis of hits, false alarms, misses and correct negatives of the whole validation dataset.

Year	Hits	False alarms	Misses	Correct negatives
2006	311199	163379	75641	305302
2007	337107	166156	74117	260703
2008 (7 months)	212399	119910	52728	171631

Table 4. Categorical statistics comparing the performances of pixel-by-pixel rain data.

	2006	**2007**	**2008**
FAR	0.34	0.33	0.36
FBS	1.23	1.22	1.25
POD	0.80	0.82	0.80
CSI	0.57	0.58	0.55
PC	0.72	0.71	0.69
POR	0.65	0.61	0.59
FRR	0.20	0.22	0.24
HK	0.46	0.43	0.39

Table 5. Pixel-by-pixel accuracy statistics by comparison of the rain data pairs from Table 1. Top to bottom the mean error (ME), the mean absolute error (MAE), the root mean square error (RMSE) and the mean are listed in mm h^{-1}. The last row indicates the mean values of 183-WSL/radar data for the whole study period.

	2006	**2007**	**2008**
ME	0.9	0.8	1.1
MAE	1.8	1.8	1.9
RMSE	2.9	2.8	3.0
M	2.41 / 1.46	2.39 / 1.45	2.50 / 1.37

Table 6. Sensitivity analysis with the 1 mm h^{-1} threshold test.

Observed (2006)

Estimated	0-1	> 1
0-1	20.47	13.79
> 1	56.97	71.08

Observed (2007)

Estimated	0-1	> 1
0-1	21.00	13.49
> 1	58.45	72.05

Observed (2008)

Estimated	0-1	> 1
0-1	19.09	12.44
> 1	57.78	70.10

Figures

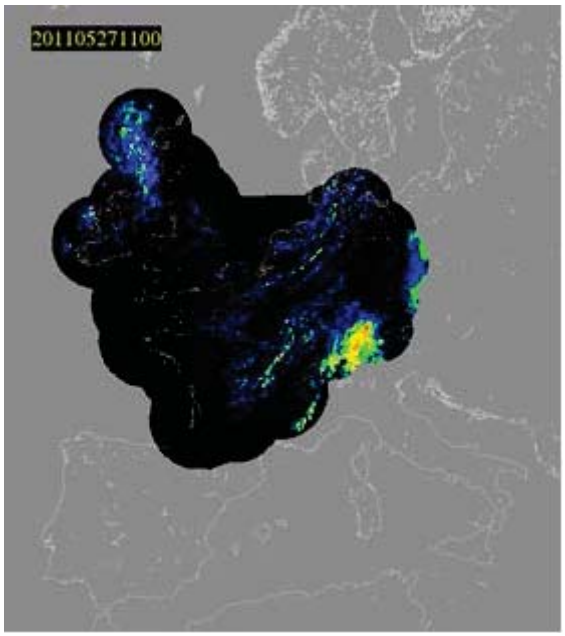

Fig.1. NIMROD weather radar network spatial coverage.

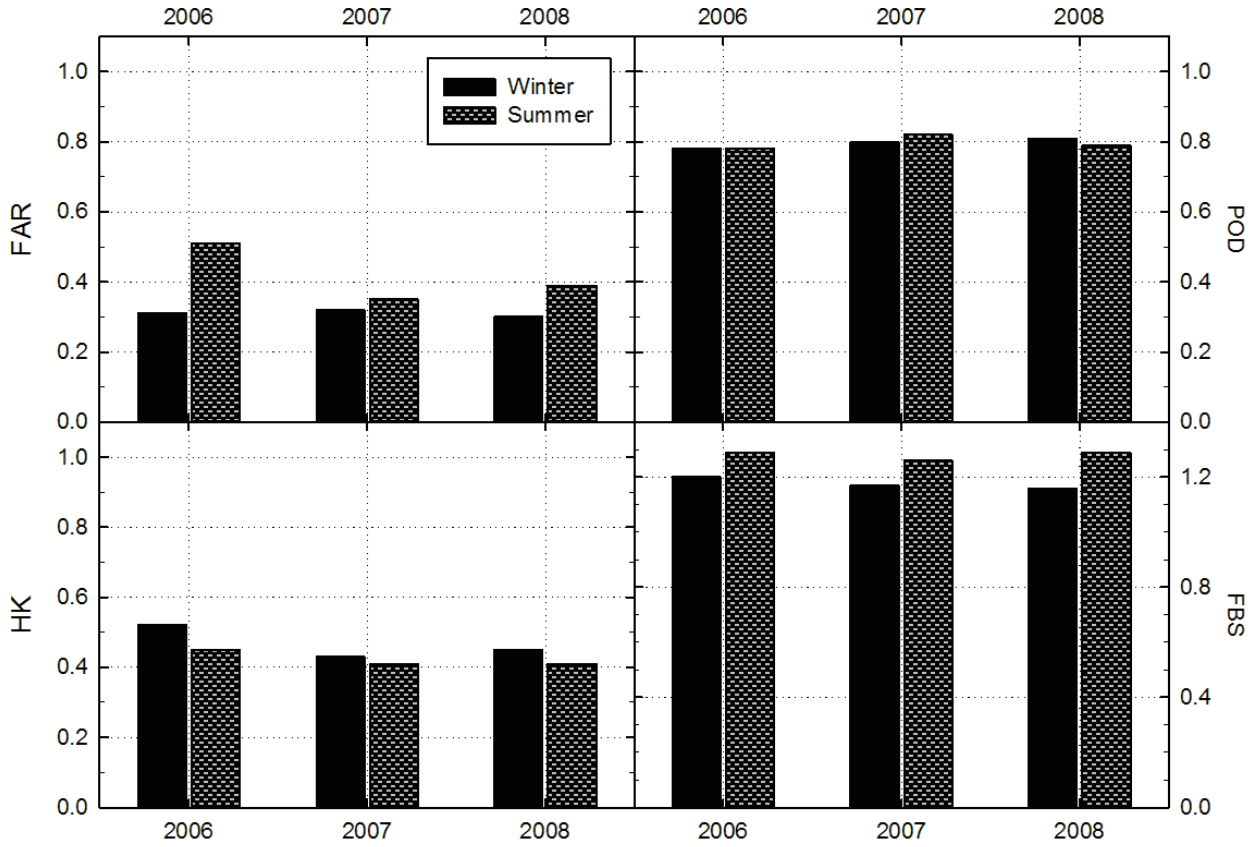

Fig 2. Mean values of the most significant categorical indices for winter (black) and summer (dotted) for the 3-year dataset. The FAR index is slightly smaller during the winter time where the screening threshold tests of the 183-WSL mitigate the false alarms due to the frozen soil. On the other hand, the Bias score increases in the warm season. The POD and HK indices are more or less stable in both seasons.

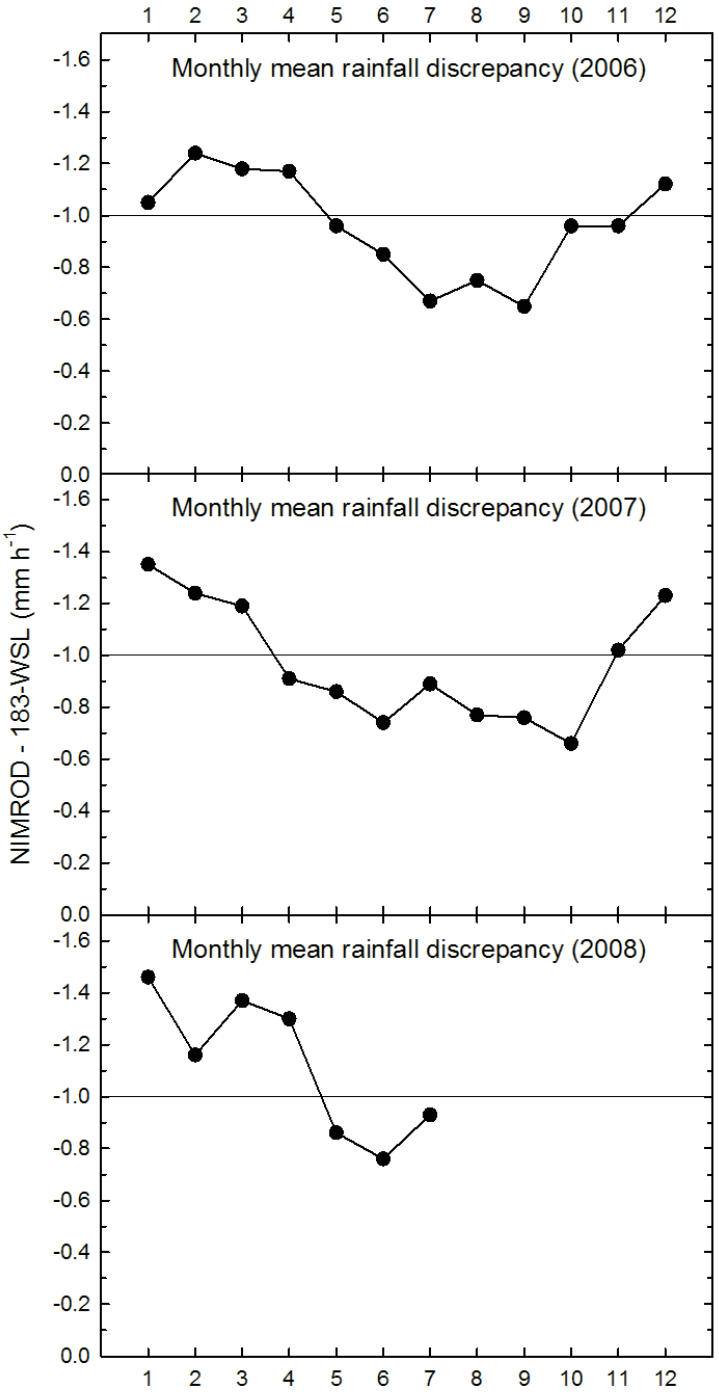

Fig 3. Monthly mean of the radar-satellite rain rate discrepancies for 2006 (top), 2007 (middle), and 2008 (bottom). The 183-WSL performs better during the summer season when the average value of discrepancies is of -0.69 mm h^{-1} with respect to the winter when the same values becomes -1.35 mm h^{-1}. The solid line marking the 1 mm h^{-1} threshold refers to the algorithm performance limit as discussed in section 2.

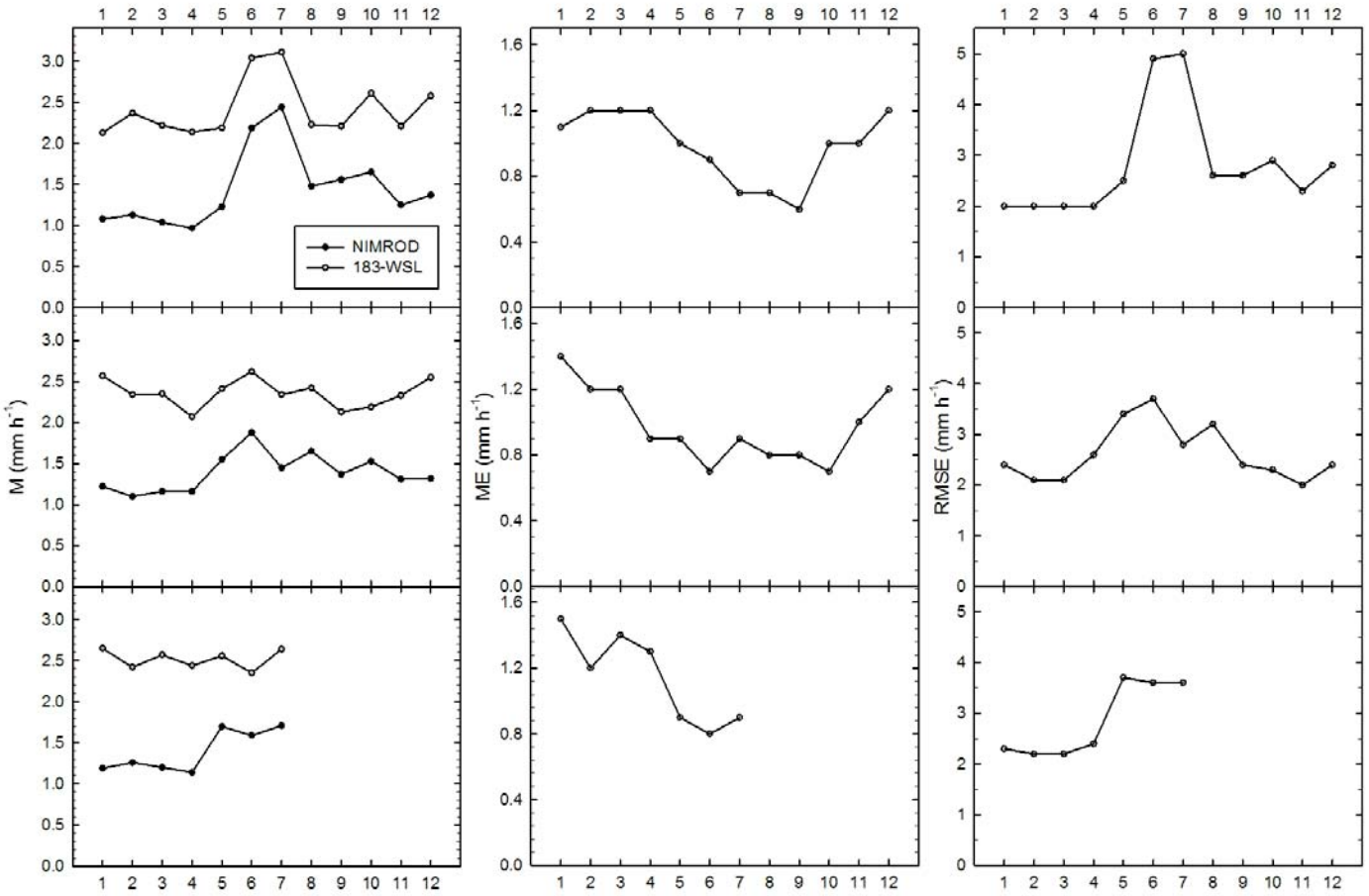

Fig 4. Monthly distribution of M, ME, and RMSE for 2006 (top), 2007 (middle), and 2008 (bottom).

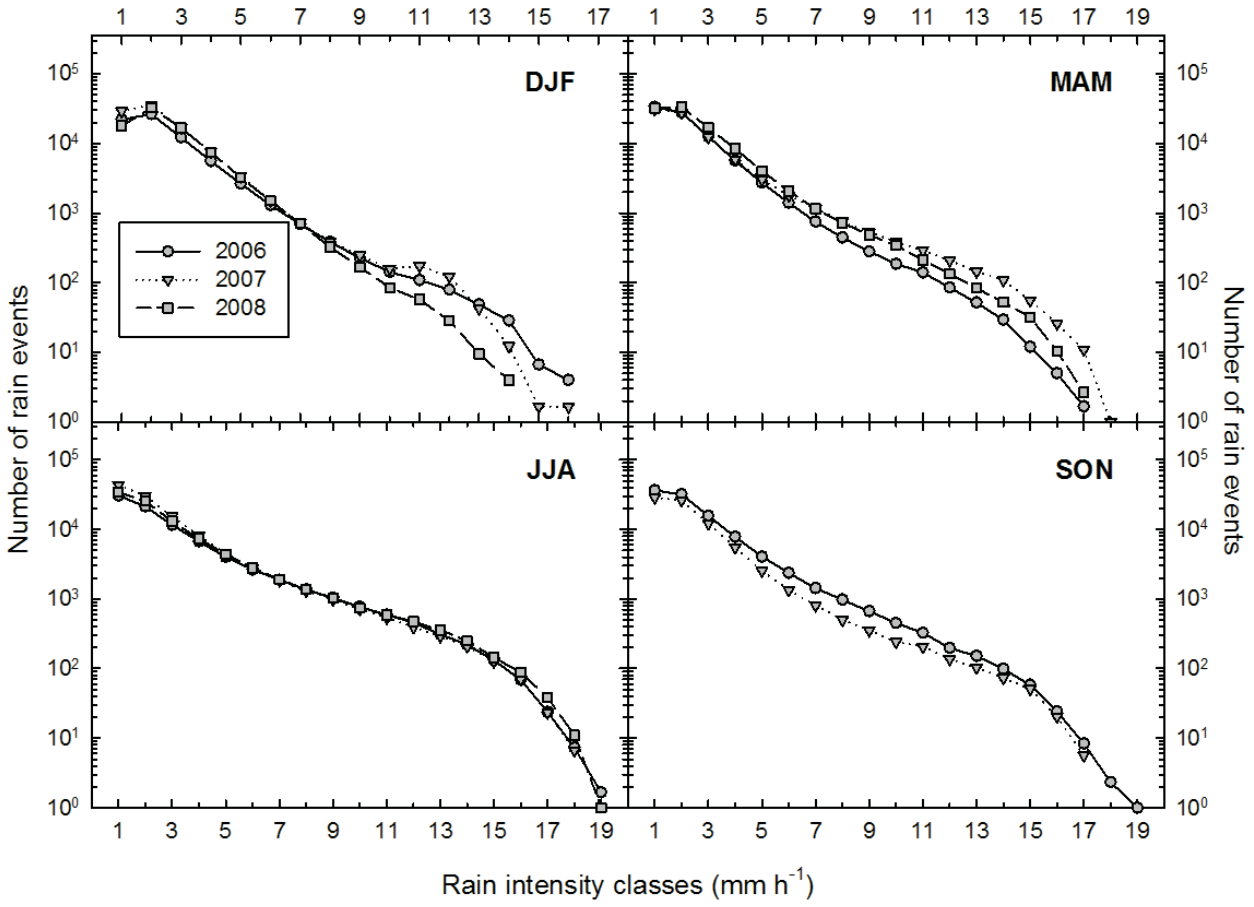

Fig 5. Intensity classes for the 183-WSL mean rain rates during the three years: DJF (top-left), MAM (top-right), JJA (bottom-left), and SON (bottom-right). The distribution of the retrieved rain rates for the DJF months is divided in two intensity regions: the first where a similar trend for the three years is found up to 7-8 mm h^{-1} and the second where different rain distribution for high intensity rain is found. The differences in this second class are lower in MAM. During the warm months, when precipitation is often associated to convective characteristics, the retrieved rain rates display no evident difference for all intensity classes.

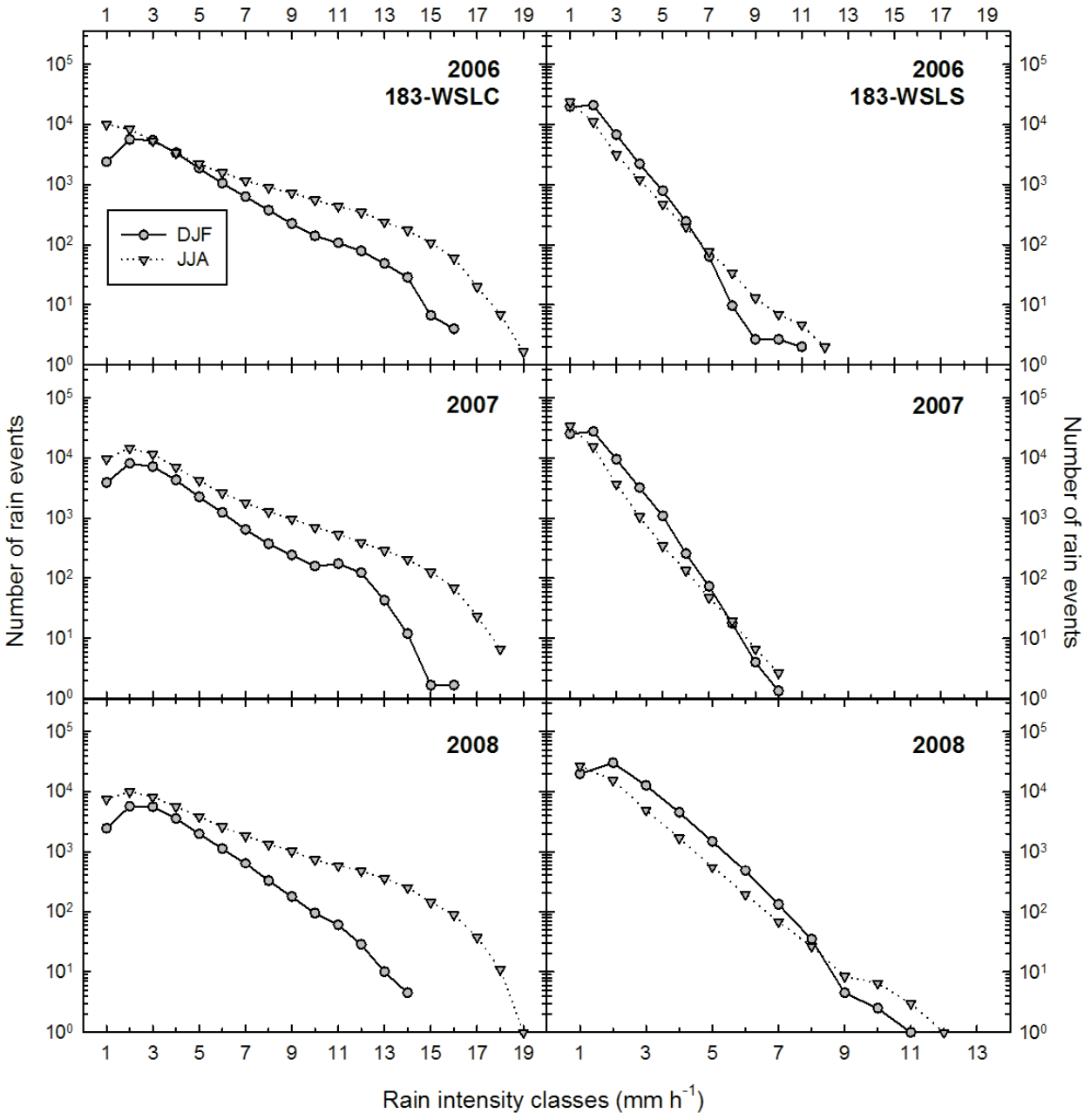

Fig 6. Mean rain rate intensity classes classified by the convective (183-WSLC, left) and stratiform (183-WSLS, right) retrieval modules during the winter and summer seasons 2006 (top), 2007 (middle), and 2008 (bottom). Note the low number of pixels belonging to high intensity classes.