H-SAF VISITING SCIENTIST FINAL REPORT

"VALIDATION OF H-SAF MICROWAVE AND INFRARED PRECIPITATION PRODUCTS WITH TRMM PRODUCTS"

by

M. Montopoli[1,2]

[1]*H-SAF visiting scientist at DPC Italy*
[2] *Sapienza Univ. of Rome*

Task 2
(T0+6)
version 1.0

ACKNOWLEDGEMENTS

This work has received various contributions.

A special thank goes to Dr. Puca and Dr. Rinollo and Dr. Vulpiani for their brilliant technical advices and passion on the subject of this work as well as to the Department of Civil Protection (DPC) hosting institution and especially to Dr. Dr. S. Meroi and Dr. P. Pagliara for they courtesy and interest that they have been showing in supporting HSAF activities within the Office II - Hydrological and anthropic risks.

I am also grateful to CNR-ISAC staff for their assistance and in particular Dr S. Dietrich , Dr E. Cattani, Dr. D. Casella, Dr. P. Sano` and Dr. G. Panegrossi, for the fruitful scientific discussions we had during the VS activity.

A special thank goes for the "Centro Nazionale di Meteorologia e Climatologia Aeronautica (CNMCA), Ufficio Generale Spazio Aereo e Meteorologia (USAM), Aeronautica Militare Italiana (ITAF)" and in particular to Cap. D. Melfi, Maj. D. Biron and Lt.Col. F. Zauli for their technical advices for HSAF data interpretation and Dr. P. Rosci, Dr. F. Gattari, and Col. L. De Leonibus for their assistance to speed-up the bureaucratic procedures.

I'm grateful to Prof. Marzano and DIET department, Sapienza university of Rome, for the support given for the participation to the 2013 HSAF meeting in Budapest and Prof. M. Pierdicca for his interesting suggestion on triple colocation technique.

I wish to thank EUMETSAT, which is supporting HSAF project and the related validation activities.

1

2 DOCUMENT CHANGE RECORD

NAME SURNAME	DATE	COMMENTS	VERSION NUMBER
Mario Montopoli	June 4th 2014	Firs draft	0
Mario Montopoli	June 20 2014	Text modifications after critical review hold on June 19th 2014 (skype session) on technical aspects and final conclusions. Contributions from CNMCA / CNR-ISAC / DPC. Added figures 2 and 3.	1

3 ACRONIMS AND ABBREVIATIONS

AMSU	*Advanced Microwave sounding Unit*
ANN	*Artificial Neural Network*
BT	*Brightness temperature [K]*
CDRD	*Cloud Dynamics and Radiation Database*
DMSP	*Defense Military Satellite Program*
GPROF	*Goddard Profiling Algorithm*
H-SAF:	*Satellite Application Facility on Support to Operational Hydrology and Water Management*
HPs	*Hypothesis*
IFOV	*Instantaneous Field Of View*
LST	*Local Solar Time*
MHS	*Microwave Humidity Sounder*
MSG	*Meteosat Second Generation*
MTG	*Meteosat third Generation*
NRLT	*Naval Research Laboratory Technique*
PDF	*Probability density function*
PMW	*Passive MicroWave*
PNPR	*PMW-Neural Network Precipitation Retrieval*
R_{H03}	*Rain rate [mm/h] from H03 product*
RMSE	*Root Mean Square Error*
R_{TMI}	*Rain rate [mm/h] from TRMM Microwave imager 2A12 v7 product*
R_{TPR}	*Rain rate [mm/h] from TRMM Precipitation radar 2A25 v7 product*
SEVIRI	*Spinning Enhanced Visible and Infrared Imager*
SSMIS	*Special Sensor Microwave Imager-Sounder*
TMI	*TRMM Microwave Imager*
TPR	*TRMM Precipitation Radar*
TRMM	*Tropical Rainfall Measuring Mission*

4 INTRODUCTION

This report summarizes the work carried out within the HSAF visiting scientist activity (ref. H-SAF_VSA_13_04) and provide a whole picture of the performance of HSAF rain product during the 12 months period between June 2012 and May 2013, over the North African and Mediterranean territory. Among the existing HSAF rain products, a prototype of the infrared based PR-OBS-03 also known as "H03" is tested. H03 is a blended product that combines infrared observations and microwave rain rate estimations, thus taking advantage of the higher temporal sampling of geostationary infrared sensors and the penetration capabilities of microwaves.

The score evaluation of H03 has a more practical impact for hydrological purpose where higher temporal sampling is required for temporal tracking of rain accumulations. For the reason mentioned above, among the various HSAS rain products, we focused our attention on H03.

Given the scarcely dense network of ground stations in Africa, the difficult to get access to field campaign data that occurred in past years in Africa and the lack of detailed data quality documentation in that area, (see *Montopoli VS report for task n.1*), we decided to deal with the validation using TRMM data and in particular the rain products from TRMM Microwave Imager (TMI, 2A12 v7) and the TRMM-Precipitation radar (TPR, 2A25 v7). The good level of interdependencies among H03, TMI and TPR, due to the different nature of the sensors used and their different rain rate retrieval algorithms, allowed performing validation both in dual and triple colocation mode. While the former is the standard way to compare couple of dataset choosing one of the two as truth, the latter way is a tool that does not require choosing a "true" source for the comparison. However, even though the triple colocation technique can be considered as a new approach in the HSAF program, it requires several hypotheses to work properly which validity is not always easy to check. A detailed appendix and formulation on the triple colocation basics is given through the main text.

Overall results indicate a systematic underestimation of the prototype H03 rain rate even though this has less impact on stratiform then convective rain regimes. The strength of the H03 underestimation error varies too as a function of the dual or triple colocation approach as well as by spatial scale of analysis. Recently introduced quality index of H03 product is also taken into account in the validation analysis. Results showed that is has a little impact on the error score considered while it does not bring benefit on the quality of the skill scores (e.g. probability of detection and false alarms).

5 OBJECTIVES

Among the operational goals of H-SAF there is the need to assign a measure of accuracy to the developed products. This is a crucial step to help the final users to fruitfully interpret the advantages and drawbacks of the H-SAF operational products.

The measure product accuracy is something related with error structure analysis. In our context, the error is defined as a measure of the discrepancy (i.e. a difference or a ratio depending if the error is thought to be multiplicative or additive) between the H-SAF operational products of rain precipitation (i.e. the satellite rain retrievals considering their various versions and degree of complexity) and independent reference records of rain precipitation not directly involved in the H-SAF product build-up (usually from punctual rain gauges or ground weather radar or other satellite active/passive observations).

Within the aforementioned context this reports has two main objectives as follows:

Objective 1: Evaluate the error structure to be associated to a prototype of H03 products.

Objective 2: Provide the algorithm developers with a benchmark for future rain product's developments and refinements.

Objective 3: Investigate the complexity, reliability and practical implementation of the triple colocation methodology for error score evaluation.

6 HSAF AND TRMM RAIN PRODUCTS

This section gives a synthetic description of the HSAF and TRMM rain products involved in this analysis. Each subsection refers to a specific rain product with a clear indication of input and output quantities, spatial and temporal sampling properties and algorithm principles.

6.1 H-SAF PR-OBS1 PRODUCT (H01)

6.1.1 Input observations and output quantities

Inputs: Currently PR-OBS1 makes use of PMW observations from ***SSMIS.*** SSMIS is a conically scanning microwave radiometer flying on DMSP satellites F16/F17/F18 series at 833 km altitude with available channels [GHz/polarization]: 9.35/H–V; 22.235/V; 37/H–V; 50.3/H; 52.8/H; 53.596/H; 54.4/H; 55.5/H; 57.29/RC; 59.4/RC; 60.792668 ± 0.357892/RC; 60.792668 ± 0.357892 ± 0.002/RC; 60.792668 ± 0.357892 ± 0.0055/RC; 60.792668 ± 0.357892 ± 0.016/RC; 60.792668 ± 0.357892 ± 0.05/RC; 63.283248 ± 0.285271/RC; 91.655 ± 0.9/H-V; 150 ± 1.25/H; 183.31 ± 1/H; 183.31 ± 3/H; 183.31 ± 6.6/H

PR-OBS1 makes use of a predefined database of possible retrievals (i.e. profiles and surface rain intensities). In order to reduce ambiguity in the PR-OBS1 retrievals, other geophysical inputs guide the algorithm towards selecting database members that are most representative of an observed scene. They are, in the general scheme, seven: (1 and 2) vertical velocities at 700 and 500hPa; (3) equivalent potential temperature at surface; (4) convective available potential energy; (5) moisture flux 50 hPa above surface; (6) freezing level height; and (7) surface height.

Outputs**:** instantaneous surface precipitation rate [mm/h] and its phase flag [liquid, frozen, mixed or unknown]; vertical profiles of microphysical species 2-water/4-ice consisting of mixing ratio (q) profiles of cloud droplets (qc), rain drops (qr), pristine crystals (qi), snow pellets/flakes (qs), ice aggregates (qa), and graupel/hail particles (qg);

6.1.2 Spatial, temporal sampling and coverage

Spatial horizontal resolution:	13.2 × 15.5 km^2 deconvolved, (30 × 30 km^2 original).
Temporal sampling (mid latitudes)*:	minimum 2 times a day (with 1 sat.); maximum: 6 times a day (with 3 sat.) Descending overpasses at approximately 05:30, 05:50 and 08:10 (LST). Ascending overpasses at approximately 17:30, 17:50 and 20:10 (LST).
Coverage:	H-SAF area [25 N–75N] latitude, [25W–45E] longitude
Future Coverage:	Full disk [65 N–65N] latitude, [65W–65E] longitude

(*) Note that since the radiometer scan swaths extend over more and more time zones, as latitude increases in either hemisphere, the actual local times at which a given location on the earth's surface is observed by a given radiometer may change considerably (even hours).

6.1.3 Algorithm synthetic description

The PR-OBS-1 product is based on a new methodology, which is founded on the ***Cloud Dynamics and Radiation Database*** methodology (***CDRD***) and the ***Bayesian solution*** method for the retrieval of precipitation. CDRD represents a modified and improved methodology applied to the Cloud Radiation Database (CRD) originally developed by Smith et al. 1992.

The key issue concerning CRD-type algorithms is that it is likely to produce ambiguous solutions. In the CRD process of obtaining a solution, an individual observed radiometer BT vector must be compared to the entire set of simulated BTs in the knowledge CRD database. This is done in order to select candidate microphysical profile solutions (which may include all profiles in the database) based on the proximity of the modeled BT vector quantities to the corresponding measured BT vector quantity. The ambiguity stems from the fact that multiple solutions are possible because different vertical profile structures of microphysical hydrometeors can lead to exactly or nearly exactly the same BT vector.

The novelty in the CDRD methodology is the use of additional parameters that better isolate candidate profile solution subsets used by the CDRD algorithm's Bayesian solver to acquire unique solutions. The additional parameters used in CDRD are optimal ***metrological and geophysical variables*** (also refereed as tags). The specific meteorological-geophysical tags have been chosen from those produced by the underlying cloud resolving model, but with a restriction that they must have direct counterparts in the observational world such that the same procedure used on the BT vectors involving proximity testing can be employed with the meteorological-geophysical tags when seeking to constrain the algorithm solution subsets

The metrological and geophysical variables tags are used to refine the selection of the candidate microphysical profiles (i.e. the possible solutions) used for the Bayesian retrieval. Then, the essence of CDRD is that it uses the optimal meteorological-geophysical parameter tags to assist in the process of guiding the algorithm in finding

microphysical profile precipitation solutions that are congruent with the environments encompassing the satellite BT measurements used in the retrievals.

In the CDRD framework, the measured optimal meteorological constraint parameters are obtained from NOAA's National Centers for Environmental Prediction (NCEP) Global Forecast System (NCEP-GFS) or from ECMWF's operational forecast model, whereas the modeled set of optional tags are obtained from the model simulations used to populate the CDRD.

While the use of an only CRD framework (i.e. BT guidance only) leads to pervasive non-uniqueness problems in finding rainfall solutions, in the CDRD framework (TB + tag guidance) mitigates against non-uniqueness problems through improved environmental constraints. Thus, the main objective in migrating from the CRD to CDRD design has been to reduce non-uniqueness effects that have affected the CRD schemes. Basically, in the CDRD scheme the candidate profiles, which are found to be not consistent with the ambient environmental conditions for given observational situations, are discarded (Casella et al. 2013). Constrain the possible solutions to a restricted subset is introduced to mitigate the ambiguities that would otherwise arise in the precipitation retrieval.

An integral part of the CDRD algorithm is a precipitation-screening scheme used to determine if precipitation exists in a given IFOV before the Bayesian solver is invoked. This is done to avoid an outcome of the solver producing nonzero positive values of surface precipitation rate even when precipitation does not actually exist in the IFOV.

6.2 H-SAF PR-OBS2 PRODUCT (H02)

6.2.1 Input observations and output quantities

Inputs**.** Currently PR-OBS2 makes use of PMW observations from ***AMSU-A*** and ***MHS.*** The channel used from AMSU-A are [GHz±Δ*f*]: 50.3 ± 50, 52.8 ± 105, 53.596 ± 115 where ±Δ*f* represents double symmetric sideband frequency positions whereas the MHS channels used are [GHz±Δ*f*]: 89 ± 0.9, 150 ± 0.9, 183.31 ± 1, 183.31 ± 3, 183.31 ± 7.

In order to reduce ambiguity, other geophysical inputs (i.e., latitude, terrain height, surface type, season) guide the algorithm towards selecting training members that are most representative of an observed scene. Additionally, Pixel/scan geometry factors are taken as input for limb smearing compensation due to the variation of the cross track view angles of AMSU and MHS.

Outputs**.** Surface precipitation rate [mm/h] and its phase flag [liquid, frozen, mixed or unknown]; the quality index for each pixel [assigned as poor, fair, good or unknown].

6.2.2 Spatial, temporal sampling and coverage

- Spatial horizontal resolution:	26×52 km^2 (at scan edge), 16×16 km^2 (at nadir).
- Temporal sampling:	minimum of 8 distinct times (i.e., twice before 01:30 LST, twice before 09:30 LST, twice after 13:30 LST and twice after 21:30 LST).
- Current Coverage:	H-SAF area [25 N–75N] latitude, [25W–45E] longitude
- Future Coverage:	Full disk [65 N–65N] latitude, [65W–65E] longitude

6.2.3 Algorithm synthetic description

PR-OBS2 is based on the ***PMW-Neural Network Precipitation Retrieval (PNPR)*** algorithm. The PNPR algorithm is based on a new optimal three-layers Artificial Neural Network (ANN), which is trained using the database from the same numerical model simulations and the same radiative transfer system that are used for populating the CDRD (Casella et al. 2013).

It is worth noting that while the CDRD algorithm for the PR-OBS1 product uses its database repeatedly to obtain a priori potential solution profiles for its Bayesian solver, the PNPR algorithm, used for PR-OBS2, uses the database only once during the training process. Another aspect to note is the consistency between PNPR and CDRD approach since they make use of the same *a priori* information (i.e. the same database of microphysical profiles) to develop the functional relationships needed between the inputs (i.e., TBs, geographical/seasonal factors, pixel view angle) and the output retrievals (i.e., surface precipitation rate, phase flag).

The use of an ANN solver relies on its convenient application when dealing with cross-track scanning radiometers. This is because in the cross-track scanning strategies the view angle is continuously changing introducing view angle-dependent scenes. This would be an additional degree of freedom when using the CDRD

and Bayesian approach that leads to systematic errors. When view-angle dependent errors enter the Bayesian retrievals, they complicate how systematic error should be expressed and impose a reduced confidence in formulating Bayesian probabilities. It has been found (Casella et al. 2013) that an ANN is able to overcome some of the view angle-dependent uncertainties at the expense of less physics based retrieval. However, to reduce its uncertainty the ANN is trained to a variety of cross-track scanning angles implying that a given cloud structure must be associated to several BT vectors (one or each of the scan radiometer view angles). This introduce an additional complexity in the PNPR training database since it has to be larger than that used in the CDRD case. Thus, the correspondence between BT vectors, along with their associated hydrometeor structures and surface precipitation rates, is complicated by the dependence of spatial resolution along a radiometer scan due to the varying view angle.

6.3 H-SAF PR-OBS3 PRODUCT (H03)

6.3.1 Input observations and output quantities

Input. Infrared BTs at 10.8 μm from GEO MSG ***SEVIRI*** satellite and rain intensity estimations from ***PMW*** satellite sensors is used as input.

Output. Map of rain rate [mm/h].

6.3.2 Spatial, temporal sampling and coverage

- Spatial horizontal resolution: variable from approx. 8 × 8 and 3.5 × 3.5 km^2
- Temporal sampling: Every 15 min
- Current Coverage: H-SAF area [25 N–75N] latitude, [25W–45E] longitude
- Future Coverage: Full disk [65 N–65N] latitude, [65W–65E] longitude

6.3.3 Algorithm synthetic description

The standard PR-OBS3 algorithm combines the temporally-rich information from the SEVIRI infrared (IR) geostationary observations together with the more quantitative, but less frequent, rainfall information from the passive microwave polar orbiting satellites (i.e. H-SAF from the products PR-OBS1 and PR-OBS2).

The infrared and PMW observations are blended using the probability matching technique. The probability matching approach used for PR-OBS3 was originally developed at the US Naval Research Laboratory (NRL) and therefore it is referred to as ***NRL Technique*** (***NRLT***).

The probability matching aims to find a relationship $y=f(x)$ to convert (or calibrate) a variable "x" into another one "y" imposing that the histograms of x and y are equal. In our case $x=BT$ [K] at IR and $y=R$ rain rate in [mm/h] estimated from PMW instruments. Thus, it is essential to have enough samples of x and y to built their histograms in a robust way. For this reason the choice of the spatial and temporal domain for the IR-PMW combined approach is a critical factor. For example, instantaneous calibrations based upon nearly coincident IR-PMW reflect the change in calibration over shorts periods. However, the use of instantaneous calibration would result in a relatively few data pairs and consequently a larger spatial domain is required to ensure an adequate sampled-size.

The NRLT processing is triggered as soon as a new slot of SEVIRI data at 10.8 μm is available. As a second step, the identification of the PMW measurements coincident in time and space with the TBs at 10.8 μm of the currently processed SEVIRI image is performed. The coincident data are subsequently located in a geographical latitude/longitude grid, and for each grid box the histogram of the IR BTs and that of the corresponding PMW rain rates are built and then combined by means of a probabilistic histogram matching technique (Calheiros and Zawadzki, 1987) to produce geo-located IR-BT vs. PMW rain-rate relationships. These relationships are then used to assign a rainfall intensity value to each SEVIRI pixel. As soon as a grid box is refreshed with new data, the corresponding relationship is renewed using updated IR-TB and PMW rain-rate histograms. Relationships older than 24 h with respect to the acquisition time of the IR TB are considered unreliable and consequently no rainfall intensity values are assigned until a refresh of the relationship is done.

The key point of this technique is thus to provide instantaneous rainfall estimations at the GEO spatial and temporal scales, which are consistent with the nature and development of the precipitating cloud systems, by overcoming the scarcity of PMW overpasses with the more frequent GEO slots and the weak connection between the rain intensity and IR BTs with the calibration of the IR BTs by the PMW rain rates. Note that in order to apply the IR-PMW blending, PR-OBS1 or PR-OBS2 or both can be used to feed the PR-OBS3 algorithm.

6.4 TRMM –TMI 2A12 RAIN PRODUCT (TMI)

6.4.1 Input observations and output quantities

Input: Microwave BTs at 10.7, 19.4, 21.3, 37.0 85.0 GHz and polarization H and V from TRMM satellite.
Output: Map of rain rate [mm/h] on the TRMM-TMI swath, land coverage map, rain quality map.

6.4.2 Spatial, temporal sampling and coverage

- Spatial resolution: variable with channels from ~6×8 km^2 at 85 GHz to ~72×43 km^2 at 10.7 GHz. On average it is approximately: 12 × 12 km^2.
- Temporal sampling: 16 orbit per day, 92.5 minutes per orbit with a swath of 878 km (post burst).
- Current Coverage: [38 S–38N] latitude, [180W–180E] longitude

6.4.3 Algorithm synthetic description

The TRMM Microwave Imager (TMI) passively collects rain information using nine channels at five microwave frequencies: 10.7, 19.4, 21.3, 37.0 and 85.5 GHz. The 21.3 GHz is the only channel that is not dually polarized (only the vertical channel is available at 21.3 GHz). Over land, the TMI rain algorithms only rise rain information from the two 85.5 GHz "scattering" channels. The lower frequency channels more directly probe the precipitable water in the lower regions of the cloud, but over land these channels become contaminated by variations in the microwave emissions from the earth's surface (i.e., non-homogeneous background). Spencer (1989), Conner and Petty (1998) and others have shown that the high frequency scattering channels are correlated with surface rain rates and so can be used as an estimator of rain rate, but since the rain information is communicated by ice-scattering processes occurring above the freezing layer, the relationship between brightness temperature and rain rate is more uncertain and is sensitive to the specific characteristics of the observed rain system. Instantaneous TMI rain rates are generated using the Goddard Profiling Algorithm (GPROF) (Kummerow et al 2001). The algorithm has continued to evolve and significant improvements to the algorithm are described in Ryu, et, a1., (2010) and Olson et al. (2006). The horizontal effective field of views of the TMI sensor varies as a function of channel's frequencies GHz/(km^2) = 10.7/ (72×43), 19.4/(30×18), 21.3/(23×18), 37.0/(16×19) and 85.5/(8× 6) (values refer to post-boost conditions in 2001). The TMI footprint in the Level II rain products is about 150 km^2 (Wolff and Fisher 2008); however, TMI rain rates are determined from passive microwave radiances collected at five different frequencies, which span a broad range of geo- physical scales and collectively probe the brightness temperature structure of the atmosphere at different depths. The Level II TMI footprint cannot therefore be thought of as representing a fundamental physical scale, but rather results from an empirical optimization of the rain information covering several different geophysical scales (Olson et al. 2006).

GPROF uses a Bayesian approach to match the observed brightness temperatures from a pixel with those from a database of simulated hydrometeor profiles from cloud-resolving-model simulations, coupled with a microwave radiative transfer model. Over ocean, the usefulness of both emission and scattering channels allows the efficient use of a Bayesian method to select hydrometeor profiles in the database from the full suite of brightness temperatures, where the physics of the algorithm relies on both emission-based and scattering-based retrievals. Over land, only scattering frequencies are useful for rain estimation, which yields too little information to use a Bayesian profile selection technique in GPROF because of the high and variable microwave emissivity of land surfaces. Ice-scattering techniques have a major disadvantage in that they are inherently empirical because of the unknown phase, density, size distribution, shape, and orientation of the ice particles in the sample volume, as well as the empirical relationship used to relate the ice scattering optical depth aloft and the rain rate underlying the ice layer.

The 2A-12 Version 7 product, "TMI Profiling", generates surface rainfall and vertical hydrometeor profiles on a pixel by pixel basis from the TRMM Microwave Imager (TMI) brightness temperature data using the Goddard Profiling algorithm GPROF2008. Because the vertical information comes from a radiometer, it is not written out in independent vertical layers like the TRMM Precipitation Radar. Instead, the output is referenced to one of 100 typical structures for each hydrometeor or heating profiles. These vertical structures are referenced as clusters in the output structure. Vertical hydrometeor profiles can be reconstructed to 28 layers by knowing the cluster number (i.e. shape) of the profiles and a scale factor that is written for each pixel.

6.5 TRMM –TPR 2A25 RAIN PRODUCT (TPR)

6.5.1 Input observations and output quantities

Input: TRMM-PR reflectivity [dBZ] 3D volumes at 13.8 GHz, freezing height from 2A23.
Output: Map of rain rate [mm/h] on the TRMM-PR swath, vertical profiles of reflectivity, rain regimes classification.

6.5.2 Spatial, temporal sampling and coverage

- Spatial resolution: 5 x 5 km^2.
- Temporal sampling: 16 orbit per day, 92.5 minutes per orbit with a swath of 878 km (post burst).
- Current Coverage: [37 S–37N] latitude, [180W–180E] longitude

6.5.3 Algorithm synthetic description

The objective of the 2A25 product is to correct for the rain attenuation in measured radar reflectivity and to estimate the instantaneous three-dimensional distribution of rain from the TRMM PR (TPR) data. The estimates of attenuation-corrected radar reflectivity factor and rainfall rate are given at each resolution cell of the PR. The estimated near-surface rainfall rate and average rainfall rate between the two predefined altitudes (2 and 4 km) are also calculated for each beam position.

Estimates of rain rate from PR require a means by which the radar signal attenuation can be corrected. One of the methods available is the surface reference technique in which the radar surface return in rain-free areas is used as a reference against which the path-integrated attenuation is obtained. Despite the simplicity of the basic concept, an assessment of the reliability of the technique is difficult because the statistical properties of the surface return depend not only on surface type (land/ocean) and incidence angle, but also on the detailed nature of the surface scattering. 2A25 basically uses the surface reference technique (*Menegini et al., 2000*) to estimate the total path loss of radar signal's power due to rain attenuation and a hybrid of the Hitschfeld-Bordan method (*Iguchi and Meneghini, 1994 and Menegini et al., 2004, 2013*) to correct for the vertical true radar reflectivity (*Z*) profile. The vertical rain profile (*R*) is then calculated from the estimated true *Z* profile by using an appropriate Z-R relationship $R=aZ^b$. One major difference from the method described in Meneghini 1994 is that in order to deal with the uncertainties in measurements of the scattering cross section of surface as well as the rain echoes, a probabilistic method is used.

7 STATE OF THE ART IN THE H-SAF RAIN PRODUCT VALIDATION

In this section a review of recent literature and their achievements that involved the validation of HSAF precipitation products is discussed. Note that each author give different definition of error score (for example average root mean square error (RMS) is not the same as the average RMS), then the comparison among score error numbers from different works should be carefully done. Lastly, note that, so far all HSAF validation efforts have been done in a dual colocation mode which means that one retrieval is compared against the reference one which is assumed to be error free.

7.1.1 Mugnai et al, 2013

Mugnai et al., 2013b showed comparisons of retrievals from the CDRD algorithm (used in PR-OBS1) and those from TRMM TMI facility algorithm 2a12-v7 (also called GPROF) with the reference retrieval from the TRMM PR facility algorithm 2a25-v7. Retrievals from both algorithms GPROF and CDRD have been obtained throughout an annual cycle (2010) for the North Africa [25N–36N, 25W–45E].

Results indicate that in a root mean square error (rmse) monthly framework,

- Over ocean, CDRD (average rmse=1.24 mm/h) and GPROF (average rmse=1.27 mm/h) algorithms are nearly equivalent.
- Over land, rainfall (nonsnowing surface conditions), the CDRD algorithm produces significant improvements (average rmse=2.94 mm/h) with respect GPROF (average rmse=3.42 mm/h) between 10 to 55 % depending on mean monthly rainfall accumulations, with an overall annual improvement of 16 %.
- Over land, summer season show a slightly increase of the "rmse" with respect winter seasons.

- On an annual basis, the CDRD algorithm relative to the GPROF algorithm exhibits moderate improvements in correlation coefficient with respect to TRMM-PR retrievals for both land and oceanic rainfall (0.65 vs. 0.60 and 0.59 vs. 0.54, respectively).
- In summary it is apparent that the CDRD algorithm is competitive with GPROF for ocean applications, and provides a significant improvement for continental applications. More exhaustive tests are underway to corroborate this result.

7.1.2 Sano et al. 2013

Sano' et al., 2013b Compared the CDRD (used in PR-OBS1) and its previous version (CRD) with rain radar retrievals. They used the Polar 55C Doppler C-band radar system at CNR/ISAC-Rome, converting the radar reflectivity into rain rate and using polarimetric variables to guarantee, with a sufficient accuracy, the calibration of the radar reflectivity. Two meteorologically distinct precipitation case studies in the Lazio region in the vicinity of Rome (land areas only) have been conducted to test the viability of the CDRD approach. The first case study (July 2nd 2009) involves convectively intense thunderstorms producing heavy rainfall. The second one (November 4, 2008) is a mix of convective and stratiform regimes.

Results indicate that

- For case study 1 (convective) RMSE and CorrCoeff are respectively 1.70 mm/h and 0.84
- For case study 2 (convective/stratiform): RMSE and CorrCoeff are respectively 1.93 mm/h and 0.88

7.1.3 Rinollo et al., 2013

Rinollo et al., 2013 introduced the concept of quality index when dealing with radar measurements for validating satellite retrievals. In that work, the precipitation field derived by radar data was compared with the PR-OBS-3 precipitation product, with varying thresholds of quality index. The impact of the introduction of the quality index was then evaluated. The dataset considered for the validation study was consisting of 12 case studies, each of 1 day, corresponding to relevant meteorological events of different seasons and precipitation regimes between summer 2009 and spring 2011.

Results showed that

- PR-RMSE (a relative RMSE here introduced) is reduced from values between 2.5 and 3 to values around 1 when the quality threshold is increased from 0 (no threshold) to 0.8. Fractional standard error also decreases, from values around 2 to values around 1.5 in the same span of the quality threshold. This can be interpreted as the signature of a strong impact of the radar data quality on the validation of satellite-based rainfall estimation using radar data as reference.

7.1.4 Puca et al 2014

Puca et al., 2014, analysed 1 year of data (July 2011 to June 2012) within the H-SAF European area using radar and gauge observations to make comparison with PR-OBS1, PR-OBS2 and PR-OBS3 products. Ground and airborne data have been spatially and temporally aliened using the temporal nearest neighbour approach and the spatial GRISO interpolator. GRISO was applied on gauge data to map them into a 5 x 5 km^2 Cartesian regular grid.

Continuous (RMSE, MAE, BIAS, CorrCoeff) and multy-category statistical scores (POD, FAR, CSI) were considered as validation indicators. In order to make validation comparison more robust, statistical scores were calculated using discretized rain precipitations into four or six rain classes depending if rain records are treated as instantaneous rate [mm/h] or accumulated rain [mm].

Results showed a

- A better performance in the winter seasons with respect to the summer of PR-OBS-1-2-3 when compared to either radar or gauge observations. Rrmse, using radar in winter/summer for PR-OBS for i-th HSAF product in [mm/h] (RMSE_W/S_*i*) is RMSE_W/S_1= 1.05/2.77; RMSE_W/S_2= 0.90/1.53; RMSE_W/S_3= 1.74/5.48;
- A good capability of PR-OBS-1-2-3 to detect low values of rain rates as opposed to a bad performance to catch high precipitation levels. .

Note: The worse performance during the summer might be also due to the convective regime of the precipitation during that period. Convective cells might extend horizontally less that typical MW-IVOF causing Non Uniform Beam Filling effects. This is unlikely to happen during the winter where precipitation is more stratiform filling

the typical MW-IVOF in a way that is more homogeneous. Thus, would be interesting to investigate the role of NUBF introducing a MW-NUBF quality index in the validation analysis.

8 METODOLOGY OF VALIDATION

This section describes the methodology of validation adopted in this work. They are essentially: i) the dual colocation and ii) the triple colocation technique. Both methodologies will be applied to the same dataset of rain precipitation product form H03, TMI and TPR, previously collocated over the target area as explained in section 8.3.

8.1 DUAL COLOCATION

The ***dual colocation*** (or dual comparison) is the usual way to make validations among different measurements of the same true (unknown) state. In this context, one measurement system is usually taken as reference and regarded as truth for the validation of the other measurement systems. This methodology has the limitation due to the fact that the reference system considered as truth has an unavoidable error that is not taken into account in the dual validation comparison. In fact, unfortunately, all observation systems contain errors. This means that we cannot assume that one measurement represents the true state and calibrate the other against it.

For example, assume we have a distribution of "true" states indicated by the variable θ, with expected variance $\sigma_\theta^2=<\theta^2>$ and two independent measurement systems x_1 and x_2, with error variances of $\sigma^2_{\delta i}=<(\theta-x_i)^2>$ where i=1 or 2 and the symbol "<•>" is the average operator. If the distribution of true values and the error distributions are normal, one can show that, for fixed x_1, (which is, for example, the measurement assumed as reference for validation), the average of x_2, which is the average of the measurements under validation, is not equal to the reference value x_1, i.e.: $<x_2>\neq x_1$, but it holds: $<x_2>= x_1 \bullet \sigma_\theta^2(\sigma_\theta^2+ \sigma^2_{\delta 1})^{-1}$ (see *Stoffelen, 1998*). Thus, only when the reference system (i.e. system 1 in our example) is error free (i.e. in the ideal case where $\sigma^2_{\delta 1}=0$), we satisfy the condition $<x_2>= x_1$. In all the other (actual) cases, we have a pseudo-bias term represented by the quantity $\sigma_\theta^2(\sigma_\theta^2+ \sigma^2_{\delta 1})^{-1}$, that depends on $\sigma^2_{\delta 1}$ which is the error characteristics of system x_1. Thus the error contribution of the system assumed as reference bias the average of the system under validation.

This example illustrates as for unbiased Gaussian error distributions, dual colocation may leads to a wrong error interpretation.

However, dual colocation has an easy implementation, gives a quick whole picture of the measuring system performance and it allows to explicitly defining an "error" difference giving the possibility for the computation of several score indexes then facilitating the comparison with independent validation analysis. For the above mentioned reasons it is still worth to apply dual colocation methods with respect to triple colocation strategies which will be described in the next section. The score and erro indexes used in the dual colocation mode are listed in the next section below.

8.1.1 Dual colocation score and error indexes

The score indexes used in this work are here listed assuming two system under investigation which are:

x_1: the measure of the reference system and
x_2: the measure of the the system under validation,

which values are given at a space (**s**) and time (t) point, indicated by x_1(**s**,t) and x_2(**s**,t), respectively.

Score indexes

$a(x_{th})$= *Number of samples for witch* $(x_2 > x_{th}$ & $x_1 > x_{th})$: *Correct positive (Hit)*
$b(x_{th})$= *Number of samples for witch* $(x_2 > x_{th}$ & $x_1 \leq x_{th})$: *False*
$c(x_{th})$= *Number of samples for witch* $(x_2 \leq x_{th}$ & $x_1 > x_{th})$: *Miss*
$d(x_{th})$= *Number of samples for witch* $(x_2 \leq x_{th}$ & $x_1 \leq x_{th})$: *Correct negative*

(1)

POD(x_{th})	=	$a/(a+c)$	*Hit rate or Prob. of Detection*
FAR(x_{th})	=	$b/(a+b)$	*False alarm rate*
PCD(x_{th})	=	$(a+d)/n$	*Prob. Of correct detection*
CSI(x_{th})	=	$a/(a+b+c)$	*Critical success index.*

ETS (x_{th})	=	$a-h/(a+b+c+h)$	*Equitable treat score*	
HSS(x_{th})	=	$(a+d-h)/(n-h)$	*Heidke skill score*	**(2)**

Where x_{th} and n is the threshold value and the total number of samples collected in the selected space-time domain, respectively, whereas h=(a+c)(a+b)/n.

Equations for the above scores are taken from Ebert et al. (2007) and are here explained in terms of rainy variables. Every match-up duplet (x_1, x_2), obtained after performing a colocation strategy as described in section 8.3, can be classified as a hit (*a*, observed rain correctly detected), miss (*c*, observed rain not detected), false alarm (*b*, rain detected but not observed), or correct null (*d*, no rain observed nor detected) event. The sum $a+b+c+d$ is equal to the sample size n. The accuracy score or probability of correct detection is defined as $(a+ d)/n$, and it indicates the fraction of total sample that has been correctly identified as rainy or non-rainy. The probability of detection, POD gives the fraction of rain occurrences that was correctly detected, while the false alarm ratio, FAR, measures the fraction of rain detections that was actually false alarms. By considering the number of hits that could be expected due purely to random chance, given by h, the HSS score is defined as the fraction of correctly detected match-ups (as rainy or non-rainy) but after eliminating the fraction correctly identified due to random chance. Similarly to this, the ETS indicates the fraction of correctly detected match-ups (as rainy), adjusted for the number of hits that could be expected due purely to random chance. ETS is more severe than HSS since it does not take into consideration the corrected negatives. For further references see Cimini, et al., 2013.

Error indexes			
δ_x	=	x_2-x_1	*Error difference*
STD	=	$<(\delta_x-ME)^2>$	*Standard deviation error*
ME	=	$<\delta_x>$	*Mean error*
MAE	=	$<\lvert\delta_x\rvert>$	*Mean absolute error*
RMS	=	$[<(\delta_x)^2>]^{0.5}$	*Root mean square error*
RMSN	=	$[<(\delta_x/x_1)^2>]^{0.5}$	*Normalized RMS error*
FSE	=	RMS/$<x_1>$	*Fractional standard error*
AMB	=	$<x_2/x_1>$	Average *Multiplicative bias*
ARB	=	$<x_2>/<x_1>$	*Average ratio bias*
CCO	=	$[<(x_1-<x_1>)(x_2-<x_2>)>]/[<(x_1-<x_1>)^2><(x_2-<x_2>)^2>]^{0.5}$	*Correlation coefficient*
LRC	=	$(<x_2 \bullet x_1>)/(<x_2 \bullet x_2>)$	*Linear correlation coefficient.*
STD30	=	*100*Number of x_2 samples within +-30% of x_1 values/n*	*30% Confidence interval*

(3)

8.2 TRIPLE COLOCATION

One possible method for quantifying measurement errors without relying on the quality of reference data sets might be the so-called ***triple collocation method***. This is a method for estimating the random errors of three collocated data sets, which can be assumed to represent the same physical parameter while simultaneously solving for systematic differences. The method assumes independent (uncorrelated) error structures, which means that the errors must not have the same origin. This is given when using, for example, any combination of in situ measurements, active or passive satellite observations, and numerical weather prediction models, provided that the model is not driven by one of the others. Several studies highlighted the high potential of the triple collocation in becoming a standard procedure in a comprehensive satellite products validation process (*Dorigo et al., 2010; Miralles et al., 2010; Scipal et al., 2008; Stoffelen, 1998*). Nevertheless, studies also showed that the result is highly sensitive to its input configuration, including different scales and represented physical quantities of the sources, the use of absolute values or anomalies, the time span under observation, and the available number of measurement triplets (*Loew and Schlenz, 2011; Zwieback et al., 2012*).

Large scale differences among the three input data sets might introduce errors caused by the spatial and temporal variability of the input quantities considered, leading to a mismatch in the spatial or temporal representativeness. These errors, called error of representativeness, are reflected in the triple collocation theory and they can take a high fraction of the overall error, leading to an overestimation of the actual inherent sensor error (*Miralles et al., 2010*). However the estimation of the representativeness error results very difficult even though some efforts have been done in literature (Vogelzang et al., 2011).

The triple colocation allows calculating two set of important information of given measurement systems: i) RMSE of measurement estimates, ii) system calibration constants. For working, the triple collocation requires at

least three measurements plus several hypothesis on the error's structure that characterize each measuring system. However, the number of the considered measuring systems can be greater that three at the cost of increased complexity. The formulas in the next section give a synthetic overview of the triple collocation method whereas details on the formula's derivation are given in the ***appendix A***.

An important aspect to consider is related to the scales (spatial or temporal) of the derived product errors. Since the triple collocation requires as input a set collocated matchups, a rescaling procedure is needed to align the three data sources at the same common scale and then to form the matched triplets. Given its easy practical implementation, the scale resampling is usually reduced to an up-scaling procedure. In the triple collocation the data up-scaling is aimed to matching the two observations at smaller scales on the larger and common scale domain of the third observation's system.

8.2.1 Fields of applications

The triple collocation was introduced for the first time by *Stoffelen, 1998* for wind and stress comparisons (*Portabella and Stoffelen, 2009*; *Vogelzang et al., 2011*). Other area of applications included wave height comparison (*Caires and Sterl, 2003*; *Janssen et al., 2007*), for sea surface temperature (SST) (*O'Carroll et al., 2008*), soil moisture (*Scipal et al., 2010*), ice drift (*Hwang and Lavergne, 2010*), precipitation analyses (*Roebeling et al., 2011*). Recently, an interesting wok of *Zwieback et al., 2012* gives the basis for the generalization to an arbitrary number of input sources based on simulated datasets. Note that so far, only the work of *Roebeling et al., 2012* focused on precipitation analyses. In that work a simplified triple colocation model was assumed and data where analyzed on a monthly basis.

8.2.2 Definitions

In the triple colocation a given observed quantity, *x,* from the measuring system *i*-th, (x_i) is though to be a linear combination of the true variable "θ", a random error "δ_i" and its bias "b_i".

$$x_i = s_i\left(\theta + b_i + \delta_i\right) \quad \textbf{(4)}$$

where:

i = *1, 2, 3. Index that identifies one of the three measuring systems.*
x_i = *measured variable from the i-th observing system.*
θ = *true (unknow) variable.*
δ_i = *zero mean random error associated to the i-th measurement system.*
b_i = *error bias (also called average error or accuracy) of i-th measurement system.*
s_i = *scaling factor (also trend of gain factor) of i-th measurement system.*

The final goal of the triple collocation is to estimate the calibration constants s_i and b_i and then error variance of δ_i. Basically, we are assuming that each estimate x_i, is the result of a linear transformation of the true variable θ plus the random error. Note that eq. (4) describes 3 equations whereas there are 9 unknowns: s_i, b_i, δ_i with i=1, 2, 3. However, using first and second mixed moments of eq. (4) and making some key hypothesis, which are described in the appendix, we can get 9 equations that allow solving the undersized problem. It is worth noting that θ is an implicit unknown variable but it is eliminated considering the difference among couple of equations of (4): x_1-x_2, x_2-x_3 and x_1-x_3 (see appendix A for details).

8.2.3 Root mean square error and calibration constants

The solution of the triple colocation is given below when ***system 1 is taken as reference for calibration*** and ***system 3 is taken as reference for the scale of analysis*** (see appendix A for details).

Calibration coefficients:

$$b_i = \frac{M_{xi}}{s_i} - M_{x1} \quad \textbf{(5)}$$

$$\begin{cases} s_2 = \dfrac{C_{x23}}{C_{x13}} \\ s_3 = \dfrac{C_{x23}}{C_{x12} - s_2 M_{\delta 12}} \end{cases} \quad (6)$$

where $M_{xi} = \langle x_i \rangle$ and system 1 is taken as reference system which leads to $b_1=0$ and $s_1=1$ and $M_{x1}=b_1 \bullet s_1=0$. $C_{xi,j}$ is the covariance of variables x_i and x_j whereas $M_{\delta 1,2}$ is the mixed moment $\langle \delta_1 \bullet \delta_2 \rangle$, which is the representative error, see appendinx A for details. The symbol $\langle \bullet \rangle$ represents the average operator in space and time.

Error variances:

$$\begin{cases} \sigma_{\delta 1}^2 = \sigma_{x1}^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \\ \sigma_{\delta 2}^2 = \sigma_{x2}^2 \left(\dfrac{C_{x13}}{C_{x23}} \right)^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \\ \sigma_{\delta 3}^2 = \sigma_{x3}^2 \left(\dfrac{C_{x12}}{C_{x23}} - \dfrac{M_{\delta 12}}{C_{x13}} \right)^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \end{cases} \quad (7)$$

where $\sigma_{\delta_i}^2$ and $\sigma_{x_i}^2$ is the variance of the error δ_i and the measured variable x_i, respectively. Eq.s (7) is rarely used in literature but we prefer to use it because the dependencies of error variance and measurements covariances of all systems are explicit and this facilitate the interpretation of the error variances in terms of the correlations among the available observations. In eq. (7), the contribution of the representativeness error is took into account through the term $M_{\delta 1,2}$. Note that $M_{\delta 1,2}$ sums to error variances $\sigma_{\delta_1}^2, \sigma_{\delta_2}^2$ of systems 1 and 2 and subtract, in some way to the error variance of system 3: $\sigma_{\delta_3}^2$. This is a direct consequence of the choice we made to consider the system 3 as the reference for scale of analysis, which implies that systems 3 describe larger scales of systems 1 and 2 and for this reason the latters pay for an additional error term $M_{\delta 1,2}$ when up-scaled at the scale of system 3. See next section and appendix A for additional details.

RMSE:
Once the error variances of the tree measuring systems is found using for example eq. (7), the RMSE can be calculated using the following formula:

$$RMSE = \sqrt{\sigma_{\varepsilon i}^2 + b_i^2} \quad (8)$$

Note that we assumed that the biases, b_i, are only referred to the error and not to the true value.

8.2.4 Representativeness errors

A representativeness error also called discretization error comes into play whenever a spatial or temporal sampling is applied to observe a given geophysical variable. The error of representativeness captures the difference between the value of the variable on the space–time scale on which it is actually measured and its value on the space–time scale on which we wish to analyze it. Thus, to define representativeness error it is important to introduce the concept of "***scale of analysis***". When we are dealing with the comparison of two observations we have to choose the scale (spatial or temporal) at which to perform the comparison between the two sources of information. Usually, one of the two observation systems is chosen as reference, which implicitly means that it is regarded as the true scale of variation. The scale of observation of such reference system is then considered as scale of analysis. Thus, all random errors components must be referred to the scale of analysis instead to its natural scale.

This implies that the random error of the *i*-th observing system, $\boldsymbol{\delta}_i$ is scale dependent and it is given by the sum of two terms:

$$\delta_i = \delta_{Mi} + \delta_{Ri} \quad \textbf{(9)}$$

where:

$\boldsymbol{\delta}_{iM}$: *Measurement error of i-th measuring system.*

$\boldsymbol{\delta}_{iR}$: *Representativeness error of i-th measuring system.*

The component $\boldsymbol{\delta}_{iM}$ depends from the instrument and algorithms used to obtain the geophysical information desired while $\boldsymbol{\varepsilon}_{iR}$ uniquely depends from the scale of analysis compared to the scale of the *i*-th observing system.

The term $\boldsymbol{\delta}_R$ can be lowered increasing the spatial and/or temporal scale resolution of analysis. Of course, from a practical point of view, this choice depends from the availability of high resolution measuring systems, which is not always practicable.

It is worth noting as the representativeness error depends by the optimal combination of the size of the scale of analysis compared to the variability of the quantity under investigation. For example, higher sampling rates do not have a sensible impact on the representativeness error when observing slowly variable quantities.

In the triple colocation equation, the representativeness error is not explicitly took into account considering directly $\boldsymbol{\delta}_R$. Rather, the representativeness error is described through mixed moment $M\delta_{ij}$ with i=1 and j=2. Note that while it is reasonable to assume that the measuremets errors $\boldsymbol{\delta}_M$ in measurements made by totally independent techniques and systems will be truly independent, the same might be not valid for the representativeness error $\boldsymbol{\delta}_R$. This directly translate into $M\delta_{Mij}$=0 when i and j are two independent systems. On the other hand, $M\delta_{Rij}$=0 when systems i and j are two systems which do not share a significant common scale of observation. When systems i and j share a large enough portion of spatial or temporal scale, $M\delta_{Rij}$ might be different from zero. Note that in general $M\delta_{ij} = M\delta_{Rij}$ only if the terms $\boldsymbol{\delta}_{Mi}$ and $\boldsymbol{\delta}_{Rj}$ and the terms $\boldsymbol{\delta}_{Mi}$ and $\boldsymbol{\delta}_{Mj}$ are not correlated which means $<\delta_{Mi}\ \delta_{Rj}>=0$ and $<\delta_{Mi}\ \delta_{Mj}>=0$ as usually largely verified:

$$M_{\delta ij} =< \delta_i \delta_j >=< (\delta_{Mi} + \delta_{Ri}) \cdot (\delta_{Mj} + \delta_{Rj}) >=< \delta_{Mi}\delta_{Mj} > + < \delta_{Ri}\delta_{Rj} >= M_{\delta Rij} \quad \textbf{(10)}$$

Suppose now that systems i and j share a common scale but their resolution is different. Let say that system "*j*" has a higher resolution (i.e. it resolves a finer scale) than system "*i*". Thus, we can think that the system with the coarser resolution (i.e. system "*i*" in our example) has a larger representative error than system j when compared to a third (larger) reference scale of analysis. This means that $\delta_{Ri}=\delta_R+\Delta\delta_R$ and $\delta_{Rj}=\delta_R$. The term δ_R represents the common part of the representativeness error shared by the systems "*i*" and "*j*" whereas $\Delta\delta_R$ is the additional error contribution of the system "*i*" due to its coarser resolution. Under the aforementioned treatment and using the hypothesis 1 ($<\delta>=0$ see appendix A) in eq. (10), it rewrites into (11) demonstrating the direct link between the terms $M\delta_{ij}$ and the representativeness error variance $\sigma_R^2=<\delta_R^2>$:

$$M_{\delta ij} = M_{\delta Rij} =< (\delta_R + \Delta\delta_R) \cdot (\delta_R) >=< \delta_R^2 > + \Delta\delta_R < \delta_R >=< \delta_R^2 > \quad \textbf{(11)}$$

It should be noted that, if the reference scale of analysis is taken coincident with that of one of the two measuring systems "*i*" or "*j*", then $M\delta_{ij}$=0. This is because at the reference scale of analysis, the error of representativeness, δ_R is zero and eq. (11) becomes $M\delta_{ij}= <\delta_{Mi}\bullet(\delta_{Mj}+\delta_{Rj})> = <\delta_{Mi}\ \delta_{Mj}> + <\delta_{Mi}\ \delta_{Rj}> = 0$. Note that measurements errors δ_{Mi}, δ_{Mj} are likely uncorrelated if they are supposed to come from different measuring systems while δ_{Mi} and δ_{Rj} are independent following their definition.

8.2.5 Considerations

Consideration 1: representativeness error

The term $M\delta_{1,2}$ in eq. (7) that describes the error variances of the three measuring systems, adds to $\sigma^2\delta_1$ and $\sigma^2\delta_2$ and in some way subtracts to $\sigma^2\delta_3$. This is a direct consequence of hypothesis HPs 3-5 (see appendix A) that imply to assume as reference scale of analysis for the comparison that of the system number 3. Thus, all error variances in (7) are referred to the scale of system 3, which is assumed larger then that of systems 1 and 2. As a consequence

systems 1 and 2 at finer scale than system 3 have an additional term, $+M\delta_{1,2}$, which quantifies the representativeness error when the scales of the systems 1 and 2 are matched up at the larger scale of system 3.

On the contrary, system 3 has not representative error since its sale of observation coincides with that taken as reference scale of analysis. Then the term $M\delta_{1,2}$, which is characteristic of systems 1 and 2 when compared at scale of system 3, is not seen by system 3 and for this reason it is subtracted.

If we make another choice, let say fix the scale of system 1 as reference scale of analysis, the terms $M\delta_{1,2}$, $M\delta_{1,3}$ would be zero because the representative error of system 1 at its scale is zero, and $M\delta_{2,3}$ will add to systems 2 and 3 and subtract to system 1 for the same reasons explained before.

Consideration 2: Covariance calculations

The calculation of the error variances requires the computation of the covariances between couples of the considered variables. However, the variables are sampled in time and space so that they match on the larger scale common to all available observations. Additionally, without any specific rule, all matched-up samples are considered all together for the calculation of the covariances. This means that we are implicitly assuming that the phenomena that we are observing, is second order stationary in time and space. This may be in general not true for a geophysical variable as the rain rate but it is surely not true when we consider large spatial and temporal domains. An alternative way might be to fix the pixel position and then calculate the temporal covariances for each position and couple of measuring systems. The final output would be, in this case, an error 2D map. In this case we need to take care of stationarity in time domain only. Of course, time and space domain are related each other and this option is more a practical implementation strategy than a solid theoretical approach. Additionally this approach can be implemented after interpolating the three source of information on a same grid, which must not vary in time. This might be not always verified when considering polar orbiting systems. A way out to solve this issue might be to apply some kind of interpolations adding an additional source of uncertainty which would be the interpolation error.

Another aspect to consider when calculating the time covariances, is that the scaling and bias factor will result adjusted in space for each pixel position. This is a paradox since the calibration parameters should not vary in space since they are, in principle an instrumental property. On the other hand this mean that we are continuously calibrating the three systems in each grid position then increasing the error's performances.

8.3 DATA SET AND COLOCATION STRATEGIES

8.3.1 Selected data set and time period for the comparison

The data products selected for the comparison are the near surface rain rate in [mm/h] obtained from a prototype of H-SAF PR-OBS3, TRMM precipitation radar ***2A25 v7*** and TRMM microwave radiometer ***2A12 v7,*** products indicated respectively by:

- R_{H03}: near surface rain from **H03** [mm/h]
- R_{TPR}: near surface rain from **TPR** [mm/h]
- R_{TMI}: near surface rain from **TMI** [mm/h]

The algorithm's background for of the above mentioned rain (R) estimations are described in section 6. The algorithm version of the H03 prototype considered in this study is obtained from a post-processed dataset and it ingests the latest upgrades of H02 product (i.e. an upgrade of the Neural Network Precipitation Retrieval over the MSG full disk) but it does not include the H01 estimates. This is mainly due to practical reasons. In fact, H01 has not followed the same algorithmic temporal evolution of H02 and its inclusion at this stage, would have risked corrupting the final H03 rain estimates. For this reason the results that follows have not to be considered as the evaluation of the state-of the art of H03 developments. However, a qualitative indication on the impact that would have been brought by the inclusion of H01 availability is shown in section 8.3.5

While H03 provides surface estimates of rain rates every 15 min on a fixed reference grid exploiting the SEVIRI geostationary orbit, TMI and TPR fly on the TRMM polar orbiting platform, providing rain rate estimations on irregular overpasses. Note that for this study H03 data have been reprocessed to timely include the SSMIS overpasses which are ingested *a posteriori* in the operational version of the H03 product. *Figure 1* shows an example of the TPR and TMI overpasses on the H03 grid within the selected target area. The period of analysis is of ***12 months from June 2012 and May 2013.*** The overall H03, TPR, TMI data storage occupancy is of 18 GB, 45 GB and 37 GB, respectively.

It should be noted that H03, TMI and TPR rain products have additional information about quality and rain flags. These flags are collected together with the rain rate values in order to partition the collected dataset as a function of rain regimes and or data quality. The selected ***data flags*** are:

- TMI freezing level estimation: expressed in [m]
- TMI quality flag: 0= good quality; 1= medium quality; 2=bad quality.
- TPR rain flag 16 bit code indicating various rain regimes and total path attenuation.
- H03 quality flag [0, 100] % values where 100 is the best data quality
- TMI surface type 10= ocean; 20=land; 30=coast.

8.3.2 Target area

The selected target area (see *Figure 1*) lies between longitude **[24W, 35E]** and latitude **[26N, 37N]**. The selected area identifies the North Africa and Mediterranean region. The choice of this area is a compromise to guarantees data intersections between the HSAF products used for the analysis, which nominal coverage is within Lat/Lon [25N,75N], [25W,45E] and the swats of TRMM products from PR and TMI sensors which are selected for the comparison. The former extend its swat up to 37N while the latter up to 38N. Thus the simultaneous colocation of the three data sources is limited in latitude up to 37 N.

8.3.3 Data spatial matching

The three source of rain rate considered, namely, H03, TPR and TMI are spatially aligned before calculating the error score in a dual or triple colocation mode. To maintain consistency between the two methodologies of data comparison, we up-scaled the available measurements to the source with the coarser spatial resolution. In our case, the IFOV of H03, TPR and TMI is approximately 3.5×3.5 km^2, 5×5 km^2 and 12×12 km^2. This has lead to upscale H03 and TPR on the TMI grid, which means that for each TMI grid point we (unconditionally) averaged all values of H03 and TPR within a circle of 12 km diameter for the selected TMI grid point. The term "unconditionally" refers to the fact that we considered all rain values i.e. rain rate greater or equal than zero within the visited TMI IFOV. For this reason, hereafter the reader should keep in mind that the results that follows, refers to the spatial scale of 12×12 km^2.

Lastly, note that as done for the rain rates we up-scaled the quality flags associated to TPR and H03 to the TMI IFOVs as well. This cannot be accomplished calculating the mean flag value because this would lead, in some cases, to up-scaled flags that do not have correspondence with the code which univocally identify the flag meaning. For example, if one TMI IFOV includes 2 TPR pixels with rain flags respectively equal to 19 and 35 their average would be 27 with could codify another rain class or could be not associated at any TPR flag. For this reason to upscale the rainy and quality flag we used the mode operator (i.e. the most frequent) with respect to the average. Note that, on average for the case analyzed, 4 TPR and H03 IFOVs are within a TMI IFOV.

8.3.4 Data temporal matching

Data temporal matching of geostationary and polar orbiting measuring systems can be something complicated over large spatial domains (*Turk et al., 2002*). Fortunately, in our case, we verified that the time needed for TRMM overpasses, over the selected target area, never exceed 15 min, which is the time sampling of H03. Thus, the temporal matching rule we adopted was to temporarily collocate all the TRMM overpasses over the target area closest to the H03 acquisition times. For the TRMM acquisition time we considered the instant $t=t_{in}+(t_{end}-t_{in})/2$ where t_{in} and t_{end} is the TRMM starting and ending overpass instants over the target area.

After the spatial and temporal matching the ***number of collected samples*** distributes as follows:

- 2283 TRMM orbits matchups
- 29723422 total unfiltered IFOVs matchups where ($R_{TPR}\geq 0$ & $R_{H03}\geq 0$ & $R_{TMI}\geq 0$§)
- 120780 rainy filtered IFOVs where ($R_{TPR}>0$ & $R_{H03}>0$ & $R_{TMI}>0$)

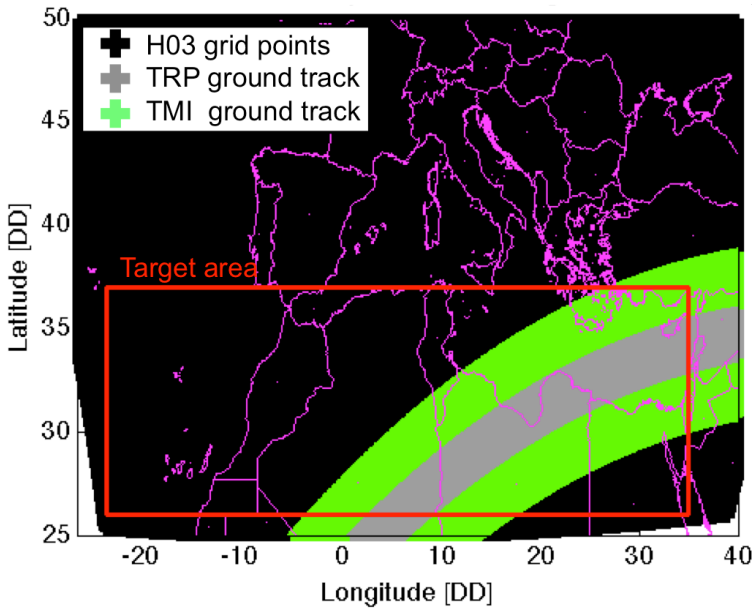

Figure 1: *Target area (in red) Lat/Lon [25N,75N], [25W,45E] and an example of TMI (gray) and TPR (green) grid points for one of the selected overpasses. The grid points of H03 are in black.*

8.3.5 Impact of H01 and H02 availability on prototype H03

In order to test the potential impact of the ingestion of H01 and H02 rain products in the generation of H03 rain retrievals, without digging into the algorithm aspects of H01, H02 and H03, we carried out a statistical analysis on the potential availability and temporal frequency of H01 and H02 for the analyzed dataset. To this aim, we grouped the available overpasses from DMSP-16-17-18 (i.e. those that would have been used to generate H01) and NOAA 18-19, METOP-A (i.e. those used in the H02 generation), within 24 hours before each TRMM overpass we considered in the comparison with the H03 prototype in the 12 months dataset form June 2012 to May 2013. The area considered for this analysis is the nominal HSAF area and not the target area in *Figure 1*. A rigorous analysis should be done considering the H01 and H02 availability with respect to the proper target area but more importantly, it should consider the spatial windows in the target area used to build-up the rain rate histograms for the implementation of the *BT* [K] vs. R_{H01}, R_{H02}, [mm/h] probability matching.

Keeping in mind the limitations above mentioned, we found that the availability of H01 and H02, considering all 24 hour-time windows before each considered TRMM overpass, if of 20% and 80% respectively. Thus, potentially,

the use of H01 in our dataset would have had a smaller impact, respect to H02, when only time availability is considered. A more detailed analysis is shown in ***Figure 2*** where the average number of available H01 and H02 retrievals is shown as a function of temporal aging from each considered TRMM overpass. In average we have we have ~100 maps from H02 against ~20 maps from H01 that contribute to the definition of the rain histogram needed to generate an instant map of H03. As last analysis we calculated the distribution of time intervals from each considered TRMM overpass and the nearest H01, H02 available retrieval. ***Figure 3*** shows the result of this analysis. The higher potential availability of H02 is confirmed for all 10min binned time intervals.

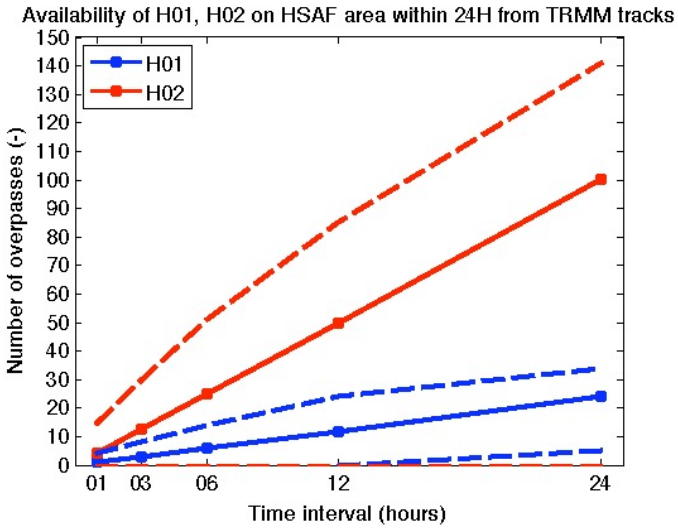

Figure 2*: Number of overpasses as a function of separation time between each considered TRMM overpass and DMSP-16-17-18 (red curve), NOAA 18-19, METOP-A (blue curve) overpasses. Dotted lines indicate min/max values whereas solid lines are average values.*

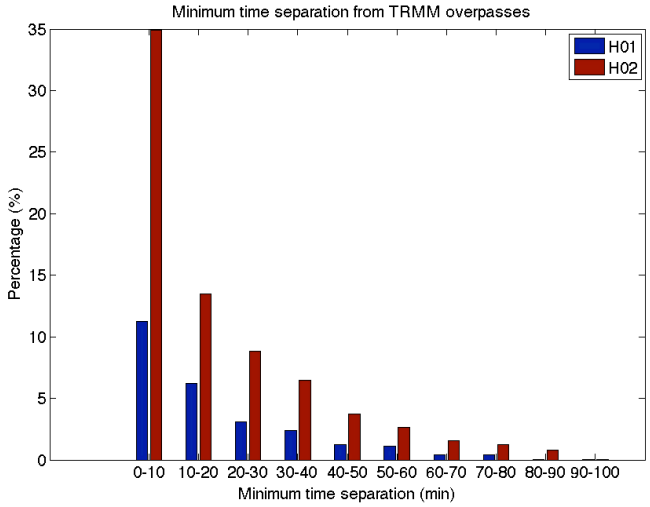

Figure 3*: Minimum time separation between each considered TRMM overpass and DMSP-16-17-18 (red bins), NOAA 18-19, METOP-A (blue bins) overpasses.*

9 RESULTS

This section shows the results of the analysis performed on 12 months collected data from June 2012 and May 2013 after up-scaling the H03 and TPR sources on the 12×12 km^2 TMI IFOV. The section divides in two parts: qualitative and quantitative results.

9.1 QUALITATIVE RESULTS

The qualitative performances of three rain products H03, TPR and TMI are firstly analyzed by a visual inspection of several case studies, which are shown in *Figure 4* by raw and by sensor from left to right, respectively. In this figure the data source are showed at their native resolution and the comparison among rain maps can be done in the

areas covered by the TPR ground track, which is grey color-coded. The list of case studies, TRMM track number and UTC hours are also reported in Table 1 for completeness.

The first thing to note is that TMI does not always provide reliable estimates of rain rate over African land (panel 1,3 in *Figure 4*). This has been explained in the TMI GPROOF algorithm section and was explained by the fact that over land, only scattering frequencies are useful by GPROOF for rain estimation, which yields too little information to use a Bayesian profile selection technique because of the high and variable microwave emissivity of land surfaces. The second thing to note is the spatial extension of precipitation of H03. Most of times H03 rain areas cover a larger domain of that covered by TPR. On the contrarily the H03 rain areas seem to be more homogeneous with respect to those covered by TMI. The spatial variability of rain precipitation appear more smoothed for H03 in comparison to that of TMI and TPR, the latter revealing a more complex rainy cell structure that what is seen from H03 and TMI. This fact lead to a underestimation of rain rate form H03 and as a consequence the hit rates defined in eq. (2) tend to reduce. Thus, high false alarms are expected. Overall we note a better consistency between TMI vs. TPR that H03 vs. TMI or H03 vs. TPR.

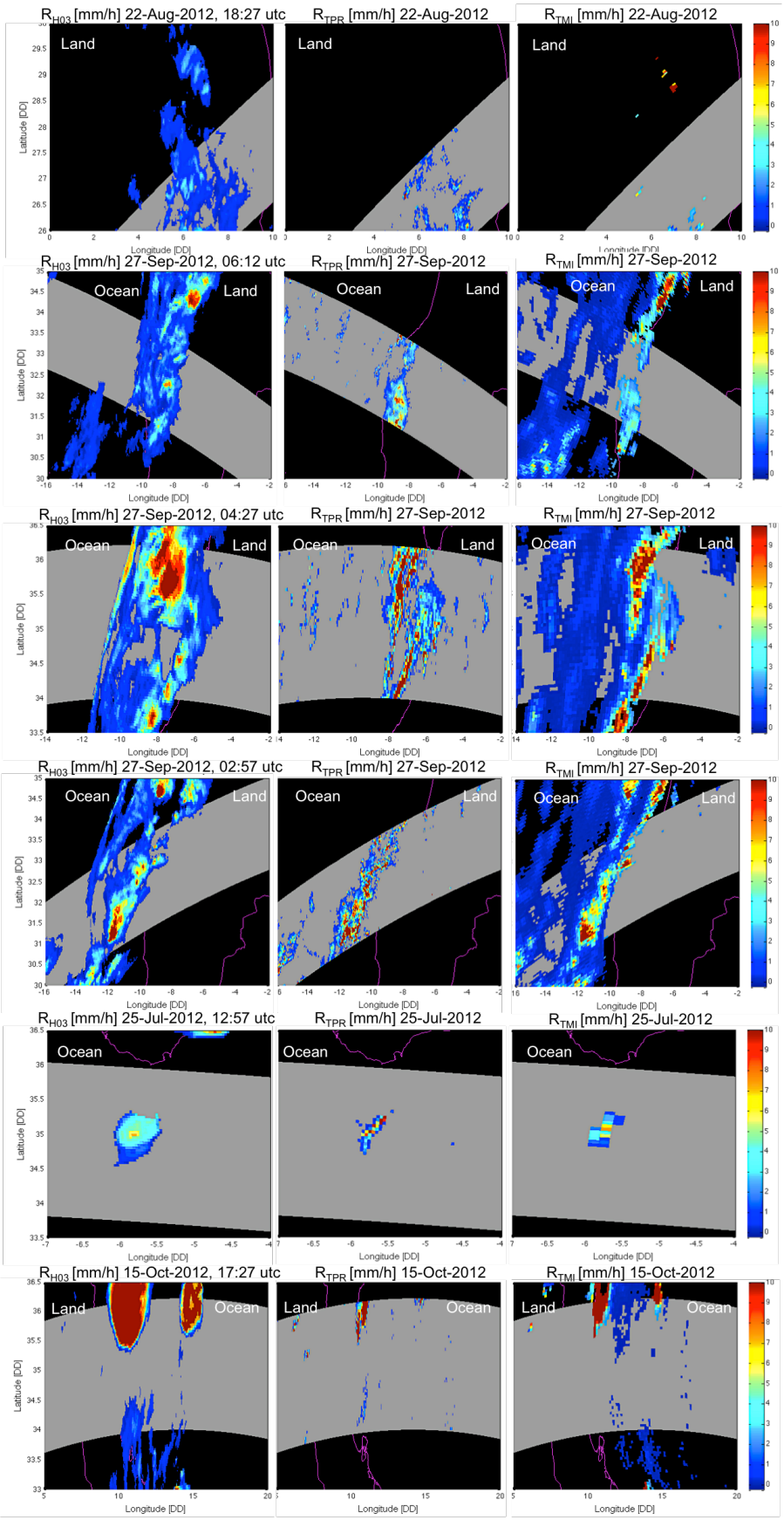

Figure 4: *Case studies for qualitative comparisons ordered by rows as indicated in each panel's title. From left to right columns: H03, TPR and TMI rain rate estimates in [mm/h], respectively. All data shown are the native spatial resolution and at approximately at the same time (i.e. with a maximum time difference of 15 min). Rain rate values are color coded in the range [0, 10] mm/h. The gray band represent the TPR ground track that is the spatial domain common to the three data sources.*

Date [MM DD YYYY]	H03 hour [utc]	TRMM hours [utc]	TRMM track number	Lon; Lal [Decimal Deg]
Jul. 25th 2012	12:57	13:01 – 13:02	83691	[-9,-4];[33,37]
Aug. 22th 2012	18:27	18:21 – 18:23	84131	[0,10];[26,30]
Sep. 27th 2012	06:12	06:06 – 06:09	84684	[-16,-2];[30,35]
Sep. 27th 2012	04:27	04:28 – 04:31	84683	[-16,-2];[30,35]
Sep. 27th 2012	02:57	02:50 – 02:54	84682	[-16,-2];[30,35]
Oct. 15th 2012	17:27	17:22 – 17:36	84972	[5,20];[33,37]

Table 1: List of case studies for qualitative comparisons.

9.2 QUANTITATIVE RESULTS

In this section the quantitative results are shown. Firstly, dual colocation methodology is applied to the couples of data FOVs, that is: H03 vs TPR, H03 vs TMI and TMI vs TPR. Even though we do not know *a priori* which of three systems perform better than the others, we refer to the TPR estimates as reference term of the comparison. For what follows we considered all data in the 12 months from June 2012 to May 2013 selecting all samples where:

Data filtering rule 1: $R_{H03}\geq 0$ & $R_{TMI}\geq 0$ & $R_{TPR}\geq 0$ & *H03_rain quality>40% & TMI_rain_quality=0 (i.e. good) & TPR_rain_flag = rain certain or no rain.*

Note that we selected all samples, which have a H03 quality greater than 40 %. It has been fond that this choice is a good compromise between the number of available samples after filtering and the correlation found between TMI and H03 rain retrievals. Additionally, it is required by rule 1 that the R_{TMI} has to be the highest possible quality (i.e. *TMI_rain_quality=0*) and, at the same time, the values of R_{TPR} have to be "certainly rain" or "no rain" (i.e. excluding cases of "possible rain" or in other words uncertain radar estimation of near surface rain estimations).

9.2.1 Dual colocation skill scores

The formal definitions of the skill scores is given by eq. (1) and (2). In those equations we chosen $x_1=R_{TPR}$ (reference system) and $x_2=R_{H03}$ (the system under validation). The definitions of hit, false, miss and correct negatives given in eq. (1) is actually here a little bit modified. Since we filtered data using rule 1 before mentioned, we are actually discarding all samples for witch TPR rain flag gives possible rain as indicated. This means to consider only the samples where $R_{TPR}=0$ or it is certain rain, i.e. $R_{TPR}\neq 0$ using the TPR rain flag code. In addition to definitions in (1) we added, only over ocean, the condition of $R_{TMI}>R_{th}$ and $R_{TMI}\leq R_{th}$ to define a rainy event and no rainy event, respectively. The rain threshold $x_{th}=R_{th}$ is left variable between 0 and 10 mm/h.

The values of POD, FAR CSI, PCD ETS and HSS defined in eq. (2) are shown in *Figure 5* where left, middle and right panel refer to result over all, land and ocean samples, respectively. While POD, ETS and CSI fall down as R_{th} increases; the FAR remain pretty stable on values of approximately of 80% and over ocean even increases up to 98 %. The underestimation of R_{H03}, already noted by the qualitative analysis of *Figure 4*, might be responsible of the increment of the miss rate (cases where rain is observed but not detected i.e.: $R_{TPR}>R_{th}$ & $R_{H03}\leq R_{th}$) and the decrement of hit rate (cases where rain is observed and detected i.e.: $R_{TPR}>R_{th}$ & $R_{H03}>R_{th}$) which traduce into a decrement of POD and stable high FAR. Over ocean we have some sensitivity of FAR to the threshold, R_{th}, and this indicates a rapid decrement of the hit rates as R_{th} increases.

The differences among land, ocean and all samples partition is better appreciated in *Figure 6* in terms of FAR and POD only. Higher POD corresponds to lower FAR and in this case is difficult to drawn conclusions. If we fix $R_{th}=0$ we can say that highest POD (50%) are over land while lowest FAR are over ocean (70%). Eventually, *Figure 7* shows POD and FAR for various H03 quality thresholds. While the H03 quality does not sensibly affect the FARs it has a detrimental impact on the POD, or in other words on the miss rates that strongly increase as the quality increases. Thus our conclusion is that the quality index is not including any quality on the detection performance of H03 product.

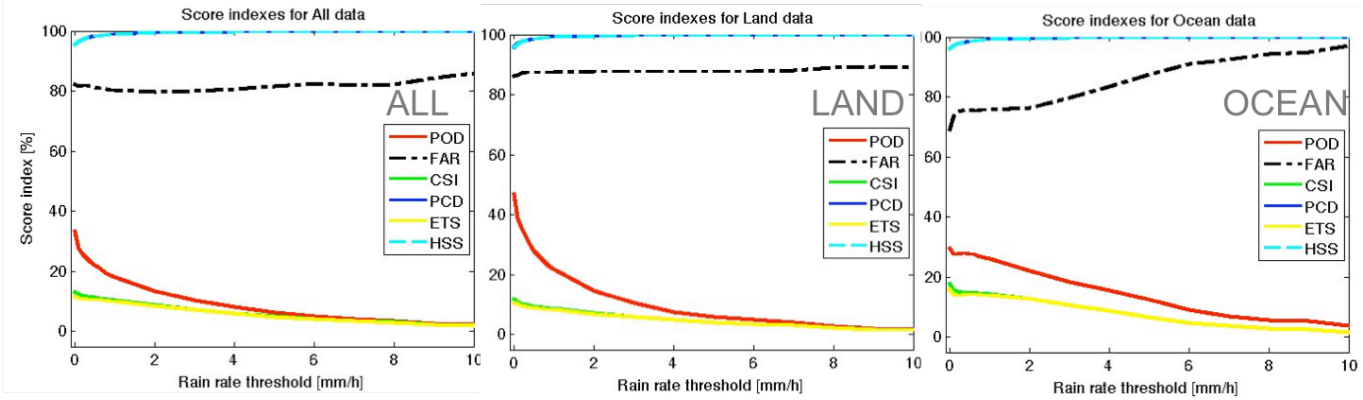

Figure 5: *skill scores defined in eq. (2) for all samples (left), land (middle) and ocean (right) panels for the whole 12 month period analyzed from June 2012 to May 2013 over the target area of* ***Figure 1****.*

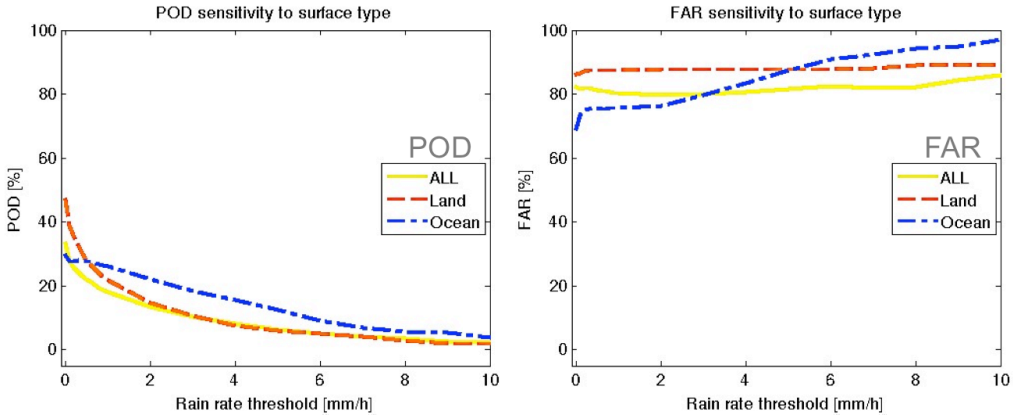

Figure 6: *POD (left) and FAR (right) partitioned into all, land and ocean samples. for the whole 12 month period analyzed from June 2012 to May 2013 over the target area of* ***Figure 1****.*

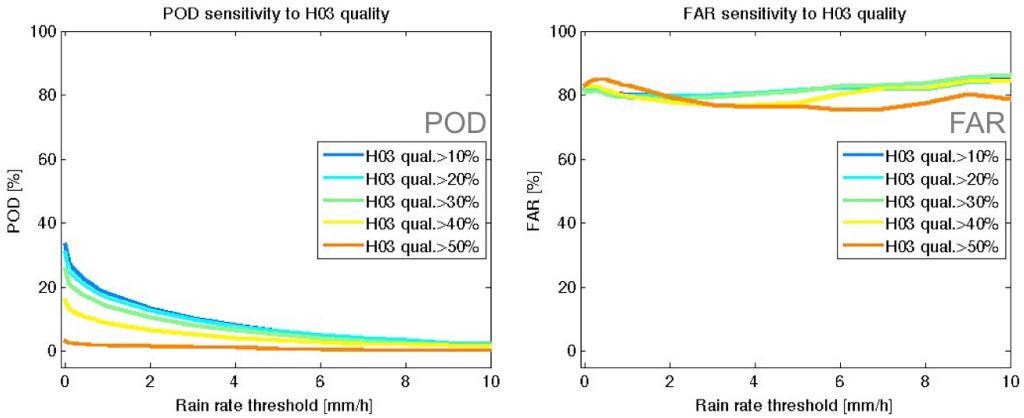

Figure 7: *As in figure 4 but for all samples as a function of H03 quality thresholds indicated in the panel's legend.*

9.2.2 Multi categorical tables

Multi categorical tables show the percentage of samples of data source n.1 that lies in a given category when, simultaneously, data source n.2 lie in another category. We selected four rain categories i.e: [0, 0.1], [0.1, 1], [1, 10] and [>10] mm/h and we build up the multi categorical tables: *Table 2*, *Table 3* and *Table 4* which refer to the comparisons between TMI vs. H03, TPR vs. H03 and TMI vs. TPR, respectively. In these tables the last column and raw list the number of samples in each category for the two data sources considered in the table.

We observe a pronounced underestimation of R_{H03}. For example when R_{TMI} is within (1,10] mm/h, 55.82% of R_{H03} estimates are within [0, 0.1] mm/h. The comparison with the TPR is even slightly worst. On the contrarily, in *Table 4* the comparison between TPR and TMI is more consistent. In this case for example, when R_{TPR} is within (1,10] mm/h, 48% of R_{TMI} are in the same category and just the 6.35% is within [0, 0.1] mm/h.

OCEAN		TMI (reference)				
Rain step [mm/h]		[0,0.1]	(0.1,1]	(1,10]	(>10)	No. H03
H03	[0,0.1]	98.37	86.79	55.82	15.02	11436000
	(0.1,1]	1.33	9.28	22.27	23.43	209170
	(1,10]	0.31	3.88	21.36	58.94	70171
	(>10)	0.00	0.06	0.55	2.61	801
No. TMI		11178000	454490	82919	1225	11717000

Table 2: *Multi categorical table TMI vs. H03. Values in the 4 x 4 inner matrix are in [%].*

OCEAN		TPR (reference)				
Rain step [mm/h]		[0,0.1]	(0.1,1]	(1,10]	(>10)	No. H03
H03	[0,0.1]	98.14	85.09	67.53	42.65	11436000
	(0.1,1]	1.48	10.15	17.41	22.79	209170
	(1,10]	0.37	4.68	14.73	33.13	70171
	(>10)	0.00	0.08	0.33	1.44	801
No. TPR		11409000	188870	11257	6197	11717000

Table 3: *Multi categorical table TPR vs. H03. Values in the 4 x 4 inner matrix are in [%].*

OCEAN		TPR (reference)				
Rain step [mm/h]		[0,0.1]	(0.1,1]	(1,10]	(>10)	No. TMI
TMI	[0,0.1]	97.31	36.69	6.35	0.32	11178000
	(0.1,1]	2.63	54.39	45.04	8.15	454490
	(1,10]	0.06	8.92	48.04	82.15	82919
	(>10)	0.00	0.00	0.57	9.38	1225
No. TPR		11409000	188870	11257	6197	11717000

Table 4: *Multi categorical table TMI vs. TPR. Values in the 4 x 4 inner matrix are in [%].*

9.2.3 Dual colocation error scores

For the calculation of the error scores we excluded from the analysis all no rainy data, which means adopting a new filtering rule:

Data filtering rule 2: $R_{H03}>0$ & $R_{TMI}>0$ & $R_{TPR}>0$ & *H03 rain quality>40% & TMI rain quality=0 (i.e. good) & TPR_rain_flag = rain certain or no rain*

With respect to the filtering rule 1, used for the calculation of the skill scores, we required all rain rate estimates to be strictly greater than zero rather then greater or equal to zero. Partition with respect land, ocean is performed using the TMI surface type flag whereas the stratiform, convective rain regimes are discerned using the TPR rain flag once up-scaled at the TMI IFOVs.

The choice to exclude the rain values equal to zero in rule 2 is because we are thinking the whole rain process as the product of two independent processes: the intermittence process (i.e. rain, no rain process) and the rain evolution process. The evaluation of the intermittence process is in some way accomplished by the analysis of skill scores, previously shown, while the evaluation of the performance of the rain evolution process is carried out in this section considering only the values during precipitating periods.

Firstly, the performance evaluation of the R_{H03}, R_{TPR} R_{TMI} is shown in terms of their probability density function (PDF_{H03}, PDF_{TPR}, PDF_{TMI}) in ***Figure 8***. From this figure, as expected for the reasons explained before, it is evident the anomaly of the PDF_{TMI} for land scenarios (bottom panel cyan curve) with respect to the PDF_{H03} and PDF_{TPR}. For this reason, hereafter, the results that will be shown in the next and that include TMI rain estimates, are referred to ocean case only. Looking at PDFs, it is interesting to note the left end tail of the PDFs for values less then 0.1 mm/h where a peak is noted. This peak is more pronounced for PDF_{H03}, and it is due to the up-scaling procedure we applied to adapt the finer spatial scale of TPR and H03 to the larger one of TMI. Thus, for example, if the average of several H03-IFOVs onto the closest TMI-IFOV includes a lot of no rainy of low rain H03-IFOV, the resulting average H03-IFOV may be a very low value (i.e. a value lower that the rain product sensibility which is 0.1 mm/h for R_{H03}). In ***Figure 8*** the PDF_{H03} is more peaked than PDF_{TPR} for values less than 0.1 mm/h and, for what explained before, this is due to the fact that R_{H03} miss more rain events than R_{TPR} (compare the peak levels of PDF_{H03} and PDF_{TPR} for values less than 0.1 mm/h). Looking at the stratiform, convective PDFs over ocean we note a good agreement for stratiform rain while for convective rain both PDF_{TMI} and PDF_{H03} are not in agreement with PDF_{TPR}.

The second test benchmark is the monthly analysis. In this case for each month we calculate the RMS of the error differences R_{H03}-R_{TMI} and R_{H03}-R_{TPR}, respectively indicated by $RMS_{H03\text{-}TMI}$ and $RMS_{H03\text{-}TPR}$. The values of such monthly RMS are in ***Figure 9*** where the average R_{TPR} values are shown as well by black/yellow markers. From this figure, $RMS_{H03\text{-}TPR}$ (red)> $RMS_{H03\text{-}TMI}$ (black), indicating a better agreement of R_{H03} with R_{TMI} than with R_{TPR}. However, the difference between $RMS_{H03\text{-}TPR}$ and $RMS_{H03\text{-}TMI}$ is very small during stratiform events over ocean or for low rain rate. Values of RMS, ME, CCO as defined in eq (3), for ocean stratiform case are listed in ***Table 5***.

The last validation test is represented by the two-dimensional (2D) histograms. In this case we considered duplet of rain IFOVs and binned them for class of rain rate. For each class we calculate the probability of the rainy IFOV in each class with respect to the total number of duplet. The result is shown in ***Figure 10*** in the ocean case. The best visual correlation is obtained in the ocean-stratiform case when R_{TPR} and R_{TMI}. In all other cases the 2D-histograms are very scattered from the bisector line. All the scores defined in eq. (3), divided by seasons for ocean convective and stratiform rain regimes, are listed in ***Table 6***,

Table 7, ***Table 8*** where for the error is defined as R_{H03}-R_{TMI}, R_{H03}-R_{TPR}, and R_{TMI}- R_{TPR}, respectively. These tables confirm what previously noted through the multi-categorical tables, PDFs and average monthly trends, that is, better performance of H03 over ocean stratiform than convective rain regimes, a better agreement of H03 with TMI than TPR with low correlation that not exceed 0.31. It is worth noting that the values in table 8 (i.e. TRMM comparisons) can be taken as reference for the other two tables. In fact, it seems reasonable to have better performance when comparing two microwave-based estimates which are collocated on the same platform than the H03 blended product. Thus, comparing the score in table 6 and 7 with those in table 8 we have, for the prototype H03, an increment of 40% - 50% of RMSE and approximately 50 decrement for correlation coefficient.

Note that the error score in the tables just introduced, are obtained selecting rain rates values strictly larger than zero. A different way to present the results is to select only those rain rates that exceed a variable threshold (>0) and calculate the error score for that threshold. Since the choice of a threshold would be arbitrary we preferred to vary the rain threshold and show the variations of the error scores as a function of the varying threshold. This is shown in ***Figure 11*** only for RMS, RMSN and CCO for ease of reading. From this figure, while the RMSN and CCO decreases, the RMS increases. Thus, in our case, we do not have a benefit in increasing the rain threshold for error evaluation because the RMS and CCO respectively increase and decreases.

	R_{H03}-R_{TPR} [mm/h]	R_{H03}-R_{TMI} [mm/h]	R_{TMI}-R_{TPR} [mm/h]
ME	-0.70	-0.70	-0.33
RMS	2.84	2.84	2.00
CCO	0.26	0.36	0.67

Table 5*: Mean Error, Root Mean Square and correlation Coefficient, for the monthly averages shown in* ***Figure 9*** in the ocean-stratiform case.

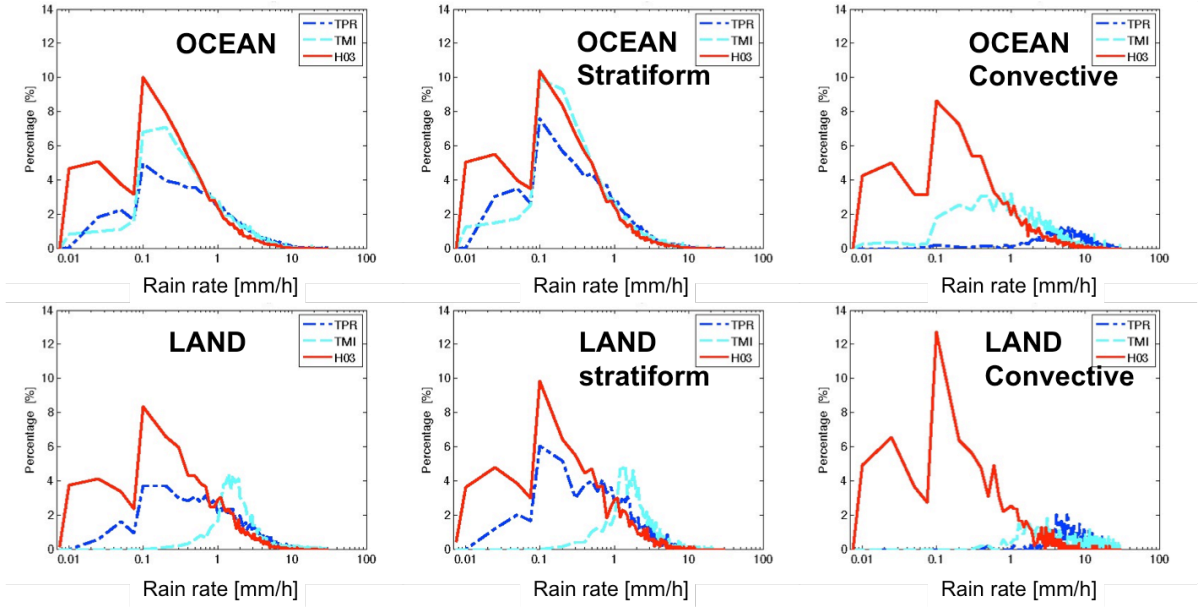

Figure 8: *Probability density functions partitioned into samples of ocean (top), Land (bottom), All (left), stratiform (middle) and convective (right) for the whole 12 month period analyzed from June 2012 to May 2013 over the target area of Figure 1.*

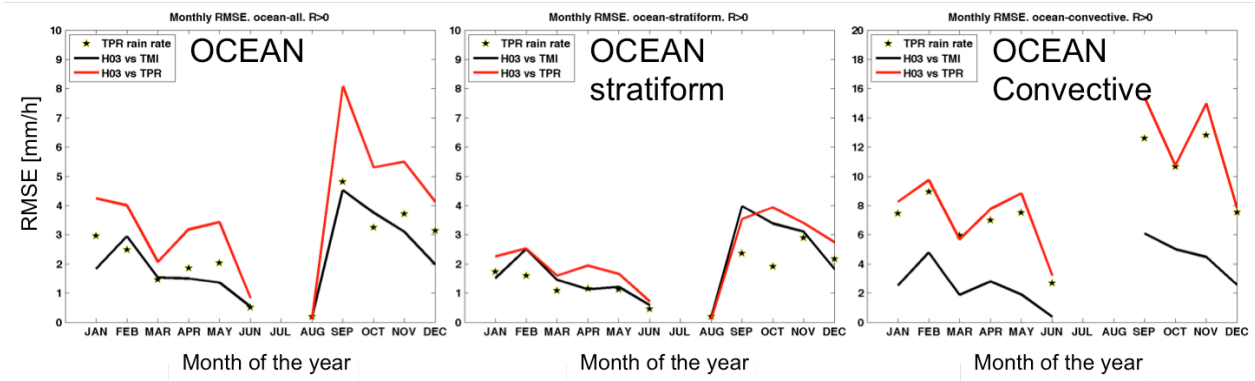

Figure 9: *Conditional (R>0) monthly RMSE of TPR vs. H03 (red), TMI vs. H03 (black). Black stars refer to values of* R_{TPR}.

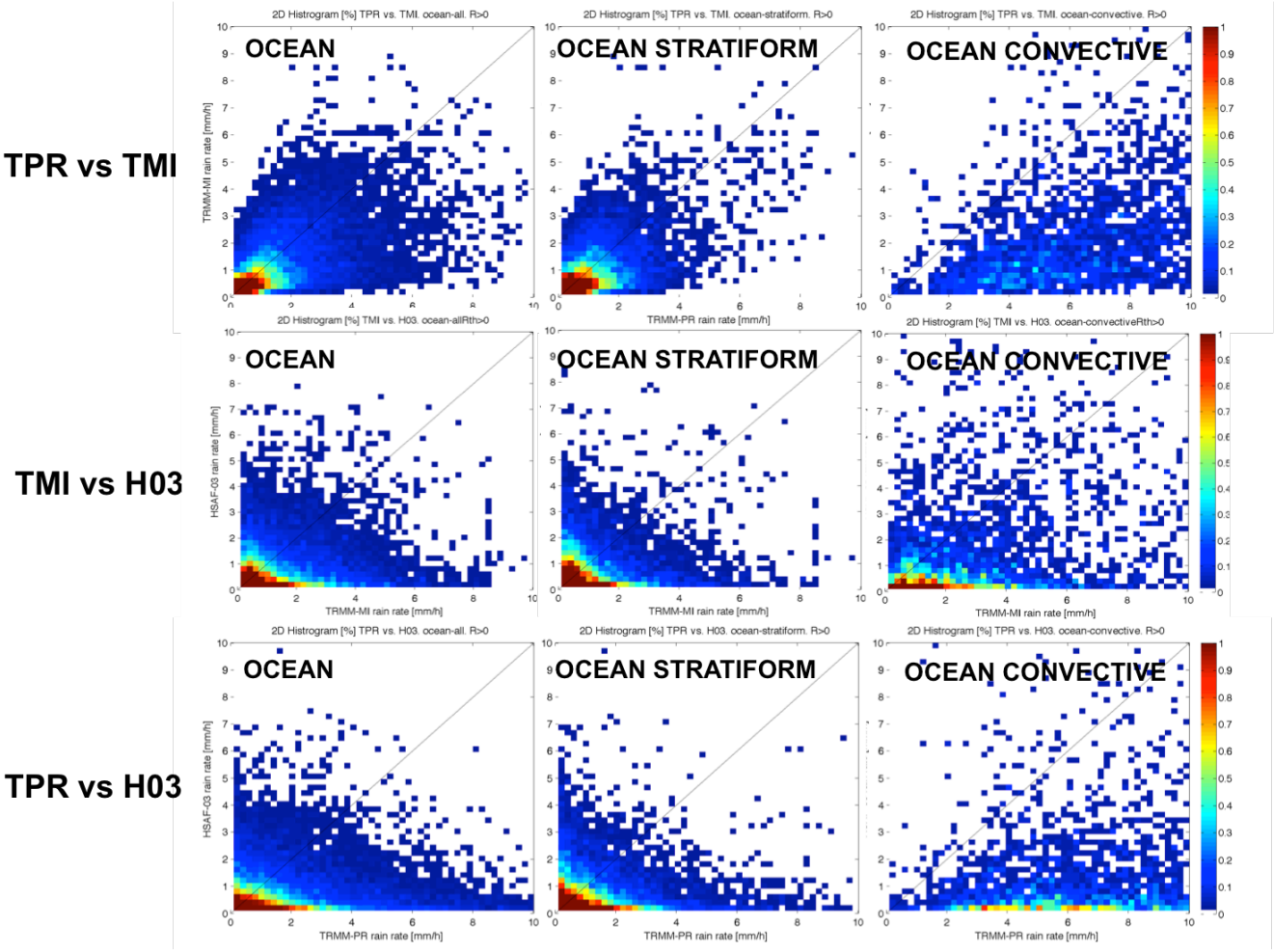

Figure 10: *two-dimensional histograms of samples as indicated in the left legend and the titles inside the panels for the whole 12 month period analyzed from June 2012 to May 2013 over the target area of Figure 1. Colors referrers to the probability (red=1, withe=0) to find (x,y) samples.*

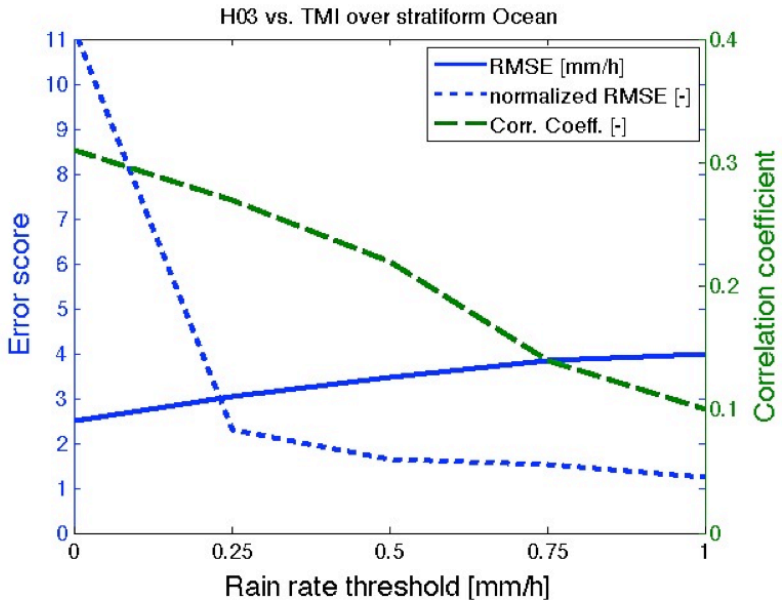

Figure 11: *Error scores between H03 and TMI rain estimates: root mean square error RMSE (solid blue), Normalized RMSE (dashed blue) and correlation coefficient (green) as a function of rain threshold in [mm/h]. Values refer to the ocean-stratiform case and for the period from June 2012 to May 2013 over the target area of Figure 1.*

H03 vs TMI (OCEAN stratiform)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	2.49	2.10	1.28	0.59	3.29
RMSN	11.27	10.74	12.39	3.89	11.22
STD	2.46	1.98	1.27	0.57	3.28
FSE	2.01	2.97	1.87	1.25	1.57
ME	-0.37	-0.69	-0.15	0.15	-0.24
MAE	1.44	1.24	0.79	0.44	2.07
AMB	2.93	2.54	3.15	2.83	3.15
ARB	0.77	0.51	0.82	1.49	0.90
CCO	0.31	0.17	0.19	-0.18	0.27
LRC	0.43	0.23	0.42	0.65	0.51
STD30%	14.84	12.23	15.63	11.92	16.86
Ns.	13160	4709	3128	193	5130

H03 vs TMI (OCEAN convective)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	4.03	3.24	2.25	0.39	4.97
RMSN	3.45	3.94	2.18	0.71	3.37
STD	3.71	2.59	1.95	0.24	4.77
FSE	2.38	4.15	2.40	3.43	1.83
ME	-1.57	-1.94	-1.13	-0.32	-1.41
MAE	2.58	2.17	1.59	0.32	3.27
AMB	1.08	0.74	0.98	0.32	1.40
ARB	0.52	0.29	0.45	0.26	0.66
CCO	0.22	0.28	0.09	-0.15	0.14
LRC	0.34	0.20	0.29	0.21	0.40
STD30%	14.99	12.00	15.52	0.00	17.36
Ns.	2956	1133	451	7	1365

Table 6*: Error score as in eq. (3), for R_{H03}- R_{TMI} in [mm/h]. Ns=Number of samples. Seasonal analysis is shown and months are grouped as indicated by in the third-sixth columns of each table.*

H03 vs TPR (OCEAN stratiform)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	2.84	2.58	1.74	0.71	3.57
RMSN	10.27	5.35	7.66	7.90	14.36
STD	2.75	2.30	1.68	0.71	3.54
FSE	2.29	3.65	2.54	1.50	1.71
ME	-0.70	-1.16	-0.43	0.01	-0.47
MAE	1.71	1.60	1.01	0.52	2.27
AMB	2.82	1.66	2.63	3.96	3.97
ARB	0.64	0.38	0.61	1.03	0.82
CCO	0.26	0.05	0.21	-0.10	0.26
LRC	0.36	0.17	0.28	0.40	0.45
STD30%	13.73	11.60	13.08	12.44	16.14
Ns.	13160	4709	3128	193	5130

H03 vs TPR (OCEAN convective)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	11.25	8.43	7.93	3.23	13.93
RMSN	0.95	0.97	0.86	0.85	0.96
STD	7.91	4.57	5.09	2.10	10.27
FSE	6.64	10.82	8.45	28.20	5.14
ME	-8.00	-7.09	-6.08	-2.58	-9.42
MAE	8.24	7.09	6.11	2.58	9.92
AMB	0.25	0.14	0.19	0.20	0.36
ARB	0.17	0.10	0.13	0.04	0.22
CCO	0.17	0.12	0.09	0.55	0.07
LRC	0.13	0.08	0.09	0.03	0.14
STD30%	4.70	1.06	2.44	14.29	8.42
Ns.	2956	1133	451	7	1365

Table 7*: Error score as defined in eq. (3), for error difference R_{H03}- R_{TPR} in [mm/h]. Ns=Number of samples. Seasonal analysis is shown and months are grouped as indicated by in the third-sixth columns of each table.*

TMI vs TPR (OCEAN stratiform)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	2.00	1.98	1.35	0.48	2.35
RMSN	2.76	2.27	3.13	1.83	2.95
STD	1.97	1.93	1.33	0.46	2.33
FSE	1.24	1.43	1.62	1.51	1.00
ME	-0.33	-0.47	-0.28	-0.14	-0.23
MAE	1.13	1.17	0.72	0.30	1.39
AMB	1.55	1.22	1.65	1.56	1.79
ARB	0.83	0.75	0.75	0.69	0.91
CCO	0.67	0.56	0.57	0.51	0.70
LRC	0.68	0.60	0.49	0.46	0.74
STD30%	24.57	22.96	23.53	25.39	26.67
Ns.	13160	4709	3128	193	5130

TMI vs TPR (OCEAN convective)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
RMS	9.37	6.76	6.55	2.92	11.74
RMSN	0.74	0.72	0.71	0.91	0.76
STD	6.81	4.39	4.29	2.01	8.59
FSE	2.87	2.49	3.17	6.70	2.85
ME	-6.43	-5.15	-4.95	-2.26	-8.00
MAE	6.68	5.42	4.97	2.32	8.32
AMB	0.39	0.37	0.34	0.70	0.42
ARB	0.34	0.35	0.29	0.16	0.34
CCO	0.52	0.37	0.59	0.63	0.52
LRC	0.29	0.31	0.26	0.13	0.29
STD30%	8.83	8.91	7.10	0.00	9.38
Ns.	2956	1133	451	7	1365

Table 8*: Error score as defined in eq. (3), for error difference R_{TMI}- R_{TPR} in [mm/h]. Ns=Number of samples. Seasonal analysis is shown and months are grouped as indicated by in the third-sixth columns of each table.*

9.2.4 *Spatial correlation analysis*

The spatial correlation of rain rate describes how correlation varies in space. We calculate the spatial correlation considering only rain rate strictly greater than zero to be consistent with the analysis carried out in the previous section. There are various techniques to calculate the spatial correlation form Fourier inversion of power spectral density to convolutional approaches. These methods require the data to be in a gridded matrix form. Since, in our analysis we excluded the non-precipitating values, our data have some spatial discontinuities and are not anymore in a matrix format and for this reason the methods before mentioned cannot be applied anymore. Thus we calculate the correlation function through the calculation of the variogram as in *Montopoli et al., 2012*. The isotropic spatial correlation functions for R_{H03}, R_{TPR}, R_{TMI} is shown in ***Figure 12*** for ocean scenarios, ocean-stratiform, ocean-convective rain regimes. As expected for convective rain regimes the correlation fall more quickly that in the other cases. Additionally in all panel's plots the correlation of R_{H03} is higher over the [0, 50] km scales while the R_{TMI} and R_{TPR} are more in agreement each other. The higher correlation of R_{H03} is indicating that it is more spatially homogeneous (i.e. smoothed) than R_{TMI} and $R_{TPR.}$ The same conclusion can be also achieved by the visual inspection of case studies in *Figure 4.* The length of correlation $L_{H03,}$ $L_{TMI,}$ L_{TPR} in [km], extracted by ***Figure 12*** is (44, 25, 19) for the ocean-stratiform case and (17, 7.5, 1.5) for the ocean-convective case. It is worth noting that higher spatial correlations of H03 may lead to overestimations in rain spatial and temporal accumulation. The explanation of that is because higher correlation in space may imply a higher persistence of rain precipitation in time. Thus, the spatial correlation can be considered as an alternative tool, with respect to considering ground reference values, to drive algorithm settings in order to quickly calibrate H03 rain accumulations.

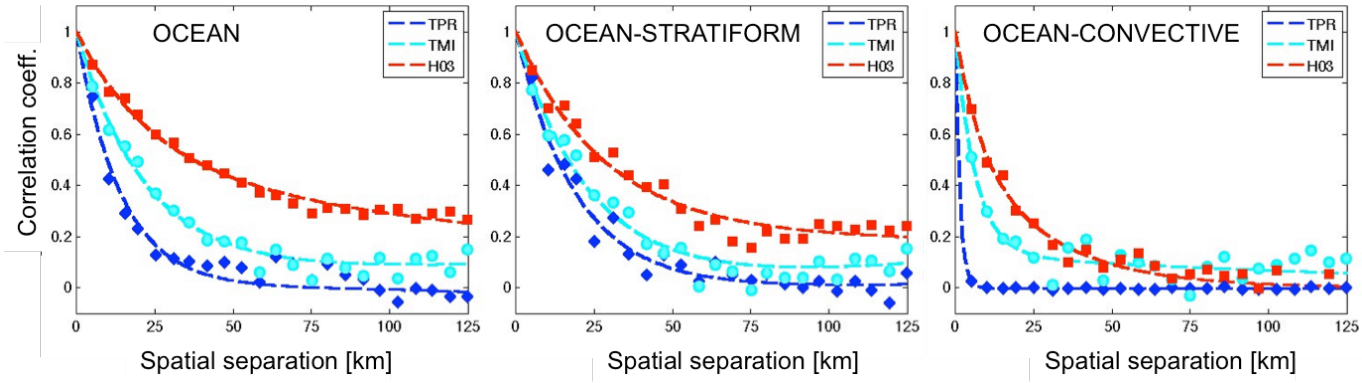

Figure 12: *isotropic spatial correlation function calculated for ocean, ocean stratiform and Ocean convective cases.*

9.2.5 *Triple colocation scores*

To obtain the triple colocation error score we applied eq.s (6)-(8) assuming systems 1,2,3 as TRP, H03 and TMI, respectively. TPR is assumed as reference for the calibration (i.e. b_1=b_{TMI}=0 and s_1=s_{TMI}=1) whereas TMI is the reference for scale of analysis which means that IFOVs of R_{TPR} and R_{H03} are averaged onto the TMI ~12 x 12 km^2 IFOVs. Note that the choice of the reference calibration system is arbitrary even though the other alternative options do not change the final results (not shown). As stated in the triple colocation system equation (7) the error variance depends from the representativeness error, which is described, in the triple colocation system by the mixed moment between errors of system 1 and 2 which is in our case $M_{\delta 1,2}$= $M_{\delta TPR,TMI}$. While it is difficult to exactly determinate the $M_{\delta 1,2}$ term, it is easy to find, from (7), its range of variation by imposing that $\sigma_{\delta_i}^2>0$. ***Figure 13*** shows the RMSE obtained after running the triple colocation as a function of $M_{\delta 1,2}$ for the ocean stratiform case. As evident from this figure, the $M_{\delta 1,2}$ has a small effect of $RMSE_{H03}$ whereas, at larger $M_{\delta 1,2}$ TMI performs better than TPR and the opposite is true for $M_{\delta 1,2}$ < 0.4. Thus, at the scale of TMI, TMI perform best. On the contrarily if TRP had a smaller representativeness error it would perform best. It is interesting how the application of the triple colocation penalizes H03 product ($RMSE_{H03}$~5.5 when $M_{\delta 1,2}$=0) whereas $RMSE_{TMI}$~2 and $RMSE_{TPR}$~1.6 mm/h. This is because, as evident in ***Figure 10***, the covariance ($C_{ij}=\rho_{ij}\bullet\sigma i\bullet\sigma j$ where ρ=correlation coefficient and σ=standard deviaton), between TPR and H03 is very low as well as that between H03 and TMI while the covariance between TPR and TMI is relatively high. In eq (7), for $\sigma\delta 2^2$, it reads C_{12} and C_{23} are low and C_{13} is high, respectively. Eventually, ***Table 9*** list the scores obtained by the triple colocation when ocean stratiform and convective rain regimes are selected. These tables also list the calibration coefficients of eq. s (5)-(6).

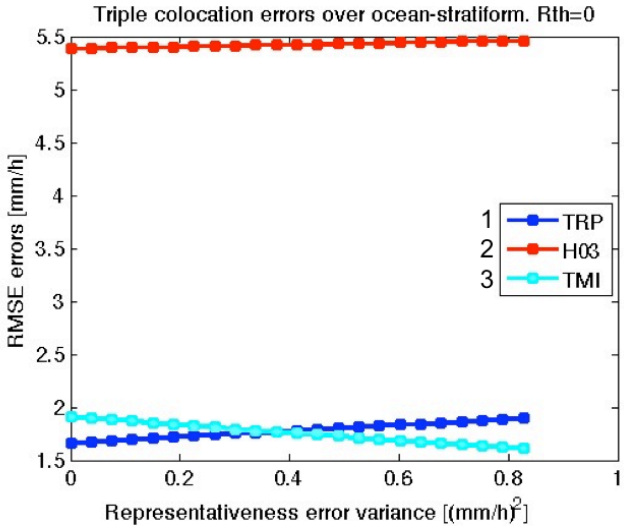

Figure 13: *RMSE as a function of representativeness error variance: term $M\delta_{12}$ in eq. (7).*

TC (OCEAN stratiform)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
s_{TPR}	1.00	1.00	1.00	1.00	1.00
s_{H03}	0.34	0.13	0.19	-0.29	0.31
s_{TMI}	1.02	3.20	0.58	1.09	0.92
b_{TPR}	0.00	0.00	0.00	0.00	0.00
b_{H03}	3.64	5.40	3.54	-1.61	6.71
b_{TMI}	1.58	0.44	1.45	0.29	2.54
$\sigma^2\delta_{TPR}$	2.77	3.93	0.92	0.20	3.24
$\sigma^2\delta_{H03}$	27.48	47.83	22.42	2.05	62.72
$\sigma^2\delta_{TMI}$	1.07	-0.36	1.63	0.00	2.57
RMS_{TPR}	1.66	1.98	0.96	0.45	1.80
RMS_{H03}	5.39	6.95	4.78	1.51	8.19
RMS_{TMI}	1.92	1.25	1.53	0.32	2.83
Ns.	13160	4709	3128	193	5130

TC (OCEAN convective)					
Scores	**All year**	**DJF**	**MAM**	**JJA**	**SON**
s_{TPR}	1.00	1.00	1.00	1.00	1.00
s_{H03}	0.14	0.16	0.04	-0.01	0.09
s_{TMI}	0.55	1.38	0.37	-0.03	0.81
b_{TPR}	0.00	0.00	0.00	0.00	0.00
b_{H03}	12.49	4.79	26.75	-13.44	30.97
b_{TMI}	5.92	1.96	5.64	-16.18	5.07
$\sigma^2\delta_{TPR}$	37.81	17.76	11.85	15.42	73.94
$\sigma^2\delta_{H03}$	319.64	32.41	1004.90	94.39	1359.30
$\sigma^2\delta_{TMI}$	11.60	0.51	7.86	76.25	-1.51
RMS_{TPR}	6.15	4.21	3.44	3.93	8.60
RMS_{H03}	17.96	5.75	31.71	9.72	36.97
RMS_{TMI}	4.72	2.81	3.48	8.74	3.93
Ns.	2956	1133	451	7	1365

Table 9*: triple colocation score in case of Ocean stratiform and Ocean convective rain regimes. The Symbol "s_x", "b_x" σ^2_x and RMS_x stands for scaling factor, bias, error variance and root mean square error of system x= TPR or TMI or H03, respectively. Values in the table refer to the case where $M\delta_{12}$* in eq. (7) is put to zero.

10 CONCLUSIONS

This work analyzed one year of space-time collocated data of rain rate retrievals from June 2012 to May 3013 derived by a prototype version H03, TRMM TMI (2A12 v7) and TPR (2A25 v7) products. The selected target area is the North African and Mediterranean region. This choice is because the next future extension of HSAF products over the MSG full disk (mainly the African territory) needs a benchmark for tracking the product's performance evaluation. The different temporal upgrades of H01 and H02 leaded to consider a prototype H03 instead of the "state of the art" H03. As a consequence the H03 prototype we considered ingests only H02 ignoring H01. Even though this surely have an impact on the results of this work, a preliminarily statistical analysis of potential availability of H01 showed that the number of ignored H01 rain rate retrievals weight approximately 20%.

Keeping in mind what just mentioned, the results highlight an overall rain underestimation of H03 product more pronounced for convective cases. Dual colocation error score of H03 are more in agreement with TMI rain rate retrievals that with those from TPR. This is some way expected since H03 ingest microwave retrievals, which are more consistent with TMI retrieval sharing the same principle of measurements.

Spatial correlation analysis evidences too an higher correlation of R_{H03} with respect to R_{TMI} and R_{TPR} indicating as the R_{H03} is more spatially smoothed that other reference retrievals. This is also confirmed by visual inspection of case studies. The skill scores of H03 evidenced very high false alarms (~80%). However, the visual inspection of several case studies has showed as an important contribution to high false alarms comes from areas where R_{H03}

correctly detect precipitation systems even though the detected contours of single cells might differ with those detected by the reference system. Additional the rain underestimation of H03 also contribute to the invariant false alarm rates for variable rain thresholds.

Eventually, triple colocation methodology shows scarce dependency of H03 RMS to the representativeness error as opposed to what happens for TMI and TPR. Indeed at the scale of TMI, It seems to perform better that TPR but the opposite would be true at the scale of TPR.

11 APPENDIX A: Triple collocation system

11.1 Problem definition

Suppose you have a measured variable x_i that is linearly related with the true variable θ plus the error ε_i of the measuring system i-th (i=1,2,3) through an error bias b_i and a scaling factor s_i:

$$x_i = s_i\left(\theta + b_i + \delta_i\right) \tag{A1}$$

which is equivalent to:

$$x_i = s_i\theta + \beta_i + \varepsilon_i \tag{A2}$$

When the $s_i \bullet b_i = \beta_i$ and $s_i \bullet \delta_i = \varepsilon_i$ The two equations are formally the same even though they lead to not identical results. However, the first one equation is more correct since in the reality when the true variable is zero the measuring instrument still scale the observation following its calibration scale factors s_i. In that case the instrument just scale the noise component. For this reason, for what follows, we referrer to equation (A1).

The goal of the triple collocation is to calculate the error variance terms of δ_i. To this end, the first step is the calculation of the calibration constants s_i, b_i.

11.2 Definitions

i = *1, 2, 3. Index that identifies a measuring system.*
x_i = *Measured variable from the i-th observing system.*
θ = *True (unknow) variable.*
δ_i = *Random error of i-th measurement system.*
b_i = *Error bias (also called average error or accuracy) of i-th measurement system.*
s_i = *Scaling factor (also trend of gain factor) of i-th measurement system.*

M_{xi} = *1st order moment of variables x_i (the mean).*
$M_{xi,j}$ = *2nd order mixed moment of variables x_i, x_j.*
$C_{xi,j}$ = *Covariance of variables x_i, x_j.*
$\rho_{xi,j}$ = *Correlation coefficient between variables x_i, x_j.*
σ_{xi} = *Standard deviation of variable x_i.*
$<\bullet>$ = *spatial-temporal average.*

$M_{xi} = <x_i>$:
$M_{xii} = <x_i^2>$:
$C_{xij} = M_{xij} - M_{xi}\,M_{xj}$
$\rho_{xi,j} = C_{xij} \bullet \sigma_{xi}^{-1} \bullet \sigma_{xj}^{-1}$
$\sigma_{xi} = M_{xii} - M_{xi}^2$

11.3 Hypothesis

Some hypothesis (HPs) that greatly simplifies the triple collocation theory has to be done to allow a practical implementation of the method. Those hypotheses concern the structure of δ_i, its cross auto-correlations and the cross correlation between "δ_i" and the true value "θ".

HP1: $M_{\delta i} = 0$

HP2: $M_{\delta i,\theta} = 0$

HP3: $M_{\delta 1,3} = 0$

HP4: $M_{\delta 2,3} = 0$

HP5: $M_{\delta 1,2} \neq 0$

HP6: $a_1 = 1; \quad b_1 = 0$

HP1-2: In particular the random errors are hypnotized to have zero-mean Gaussian distribution and they are required to be uncorrelated from the true value "θ" (HP2).

HP3-4: Additionally, the correlation among errors from different systems may be zero. The assumption of uncorrelated errors is by no means evident, and should, if possible, be tested.

HP5: A positive error correlation term can be introduced in the triple collocation technique to include the error of representativeness associated to the different sensors' resolutions as in HP5. The representativeness error is discussed in more details in the main text. Now is important to remember that the quantities $M_{\delta i,j}$ are something related to the representativeness error (also called discretization error) of systems "i" and "j" with respect to a fixed reference scale. $M_{\delta 1,2} \neq 0$ for example, describe a representativeness error of systems 1 and 2 when the scale of system 3 is the reference scale for the comparison. Note that the errors terms δ_i's might be thought as the sum of two contributions: a measurement error and a representativeness error. The former is typically an instrumental error, which can be sensor dependent whereas the latter is a scale dependent error contribution. Observations at large scales miss information at finer scales and this is took into account trough a representativeness error.

HP6: Eventually, to be able to calculate the calibration constant in (A1) at least one of the three measuring systems have to be taken as reference for the calibration of the other twos. We arbitrarily choose system number 1 as the reference system. We have the freedom to do this because results on the errors do not depend on this choice. In the description that follows the system 1 is taken as reference. This implies to put bias b_1 and gain factor s_1 equal to zero and equals to unity.

11.4 Calibration constants

First let's express the moments of order of 1 for i-th system:

$$M_{xi} = < x_i > = s_i \left(M_\theta + b_i + M_{\delta i} \right) \tag{A3}$$

To calculate the calibration constants, one of the three systems has to be chosen as reference. However, this choice does not affect the final result. We chose system k-th (with k=1 or 2 or 3) as reference calibration system, which implies to assume s_k=1 and b_k=0 and as a consequence $M\theta = <\theta> = M_{xk}$. Note also that also that $M\delta i$=0,under HP1. Then (A3) transforms into:

$$b_i = \frac{M_{xi}}{s_i} - M_{xk} \tag{A4}$$

Second step is calculate the second order moments $M_{xii} = <x_i^2>$:

$$M_{xii} = < x_i^2 > = s_i^2 \left(M_{\theta\theta} + M_{\delta ii} + b_i^2 + 2 b_i M_\theta \right) \tag{A5}$$

Using $M\theta = M_{xk}$ and substituting (A4) into (A5) we obtain:

$$M_{xii} = < x_i^2 > = s_i^2 \left(M_{\theta\theta} + M_{\delta ii} + \frac{M_{xi}^2}{s_i^2} - M_{xk}^2 \right) \tag{A6}$$

The latter equation can be expressed in terms of error variance $M\delta ii$, recognizing that $\sigma^2_{xi} = M_{xii} - (M_{xi})^2$:

$$M_{\delta ii} = < \delta_i^2 > = \frac{\sigma_{xi}^2}{s_i^2} - \left(M_{\theta\theta} - M_{xk}^2 \right) \tag{A7}$$

The term $M_{\delta ii}$ in (A7) is our target quantity since it is the error variance $\sigma_{\delta_i}^2$. To fully resolve (A7) we have to calculate s_i and $M_{\theta\theta}$. To this aim we make use of mixed moments:

$$M_{xij} =< x_i x_j >= s_i s_j \left(M_{\theta\theta} + b_j M_\theta + b_i M_\theta + b_i b_j + M_{\delta ij} \right) \quad \textbf{(A8)}$$

Substituting (A4) into (A8) we obtain:

$$M_{xij} =< x_i xj >= s_i s_j \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta ij} + \frac{M_{xi} M_{xj}}{s_i s_j} \right) \quad \textbf{(A9)}$$

And recognizing that the covariance of measurements x_i and x_j is $C_{xij}=M_{xij}-M_{xi}\bullet M_{xj}$ we can rewrite (A9) as:

$$C_{xij} = s_i s_j \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta ij} \right) \quad \textbf{(A10)}$$

Eq. (A10) when written for the three systems produces:

$$\begin{cases} C_{xij} = s_i s_j \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta ij} \right) \\ C_{xik} = s_i s_k \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta ik} \right) \\ C_{xjk} = s_j s_k \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta jk} \right) \end{cases} \quad \textbf{(A11)}$$

To solve (A11) we have to make some hypothesis. We fixed the system k-th as that for the reference calibration which leads to $s_k=1$. If the scale of analysis is fixed to that of system i-th then $M_{\delta ij}=M_{\delta ik}=0$ which is consistent with HP3- HP 5 Then system (A11) reduces to:

$$\begin{cases} C_{xij} = s_i s_j \left(M_{\theta\theta} - M_{xk}^2 \right) \\ C_{xik} = s_i \left(M_{\theta\theta} - M_{xk}^2 \right) \\ C_{xjk} = s_j \left(M_{\theta\theta} - M_{xk}^2 + M_{\delta jk} \right) \end{cases} \quad \textbf{(A12)}$$

The ratio between the first and second equation of (A12) gives the coefficient s_j while substituting the third equation into the first one gives the coefficient s_i:

$$\begin{cases} s_i = \dfrac{C_{xij}}{C_{xjk} - s_j M_{\delta jk}} \\ s_j = \dfrac{C_{xij}}{C_{xik}} \end{cases} \quad \textbf{(A13)}$$

(A13) gives the calibration coefficients when "k" and "i" are the systems chosen as reference for calibration and scale of analysis, respectively. Of course system i, j, k can be one of the three systems 1,2,3 with $i \neq j \neq k$.

11.5 Error variances

The last step is to calculate the error variances $\sigma_{\delta_i}^2$. This is straightforward using the third equation of (A13) into (A7) obtaining:

$$\sigma_{\delta m}^2 = \left(\frac{\sigma_{xm}^2}{s_m^2} \right) - \frac{C_{xjk}}{s_j} + M_{\delta ik} \tag{A14}$$

Eq. (A14) can be further simplified substituting the coefficient s_j from the second equation of (A13):

$$\sigma_{\delta m}^2 = \left(\frac{\sigma_{xm}^2}{s_m^2} \right) - \frac{C_{xkj} C_{xki}}{C_{xji}} + M_{\delta ki} \tag{A15}$$

Eq. (A15) is the compact general form of the triple collocation solution for systems $i, j, k, i \neq j \neq k$ ranging from 1 to 3 and where $m = i$ or j or k is just an index ranging from 1 to 3. Note that in (A15)

System k = reference system for calibration

System i = reference system for scale analysis.

Note that error system variance in (A14) has a variable term, which is the square of the scaled measurement variance, and invariant terms, which are the second and third term of the right side of (A14).

For example when system i=3 and k=1, and j= 2, we have:

$$\begin{cases} \sigma_{\delta 1}^2 = \sigma_{x1}^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \\ \sigma_{\delta 2}^2 = \sigma_{x2}^2 \left(\dfrac{C_{x13}}{C_{x23}} \right)^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \\ \sigma_{\delta 3}^2 = \sigma_{x3}^2 \left(\dfrac{C_{x12}}{C_{x23}} - \dfrac{M_{\delta 12}}{C_{x13}} \right)^2 - \dfrac{C_{x12} C_{x13}}{C_{x23}} + M_{\delta 12} \end{cases} \tag{A16}$$

The aforementioned demonstration assumes that the reference systems for calibration and scale of analysis are two different measuring systems. There is no apparent reason to make such assumption. However, if the same measuring system is taken as reference for both scale analysis and calibration we have a slightly different solution.

The error variance formulation in (A14) remains unchanged but the coefficients in (A13) slightly changes. If system k is unique reference system, the representativeness error $M\delta_{jk}=M\delta_{ik}=0$ and $M\delta_{ij}\neq 0$, the calibration coefficients in (A13) rewrites into:

$$\begin{cases} s_i = \dfrac{C_{xij}}{C_{xjk} - s_j M_{\delta ij}} \\ s_j = s_i \dfrac{C_{xjk}}{C_{xik}} \end{cases} \tag{A17}$$

Lastly, note that (A13) and (A17) are the same if representativeness error, $M\delta_{ij}$, is zero.

12 REFERENCES

General

Smith, E. A., Mugnai, A., Cooper, H. J., Tripoli, G. J., and Xiang, X.: Foundations for statistical – physical precipitation retrieval from passive microwave satellite measurements, Part I: Bright- ness temperature properties of a time dependent cloud – radiation model, J. Appl. Meteor., 31, 506–531, 1992

Calheiros, R. V. and Zawadzki, I. I.: Reflectivity rain-rate relation- ship for radar hydrology in Brazil, J. Clim. Appl. Meterol., 26, 118–132, 1987.

Stephens, G. L. and C. D. Kummerow, "The remote sensing of clouds and precipitation from space: A review," J. Atmos. Sci., vol. 64, no. 11, pp. 3742–3765, Nov. 2007.

Ebert E., Janowiak, J. E., and Kidd, C.: Comparison of Near- Real-Time Precipitation Estimates from Satellite Observations and Numerical Models, Bull. Amer. Meteor. Soc., 88, 47–64, doi:10.1175/BAMS-88-1-47, 2007.

Cimini D., F. Romano, E. Ricciardelli, F. Di Paola, M. Viggiano, F. S. Marzano, V. Colaiuda, E. Picciotti, G. Vulpiani, and V. Cuomo, (2013), Validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauge networks, Atmos. Meas. Tech., 6, 3181–3196, www.atmos-meas-tech.net/6/3181/2013/ doi:10.5194/amt-6-3181-2013

Turk F. Joseph, Elizabeth E. Ebert, Hyun-Jong Oh and Byung-Ju Sohn, Vincenzo Levizzani, Eric A. Smith, Ralph Ferraro, (2002), Validation of an operational global precipitation analysus at short tome scales, in operational Applications and Artificial Intelligence (Joint between 12th Conference on Satellite Meteorology and Oceanography and Third Conference on Artificial Intelligence Applications to Environmental Science)

Montopoli, M., Pierdicca, N., & Marzano, F. S. (2012). Spectral Downscaling of Integrated Water Vapor Fields From Satellite Infrared Observations. IEEE T GEOSCI REMOTE, 50(2), 415-428. doi:10.1109/TGRS.2011.2161996

Montopoli, (2014), "Inventory and description of precipitation ground data and TRMM products available on the Euro-African region, "Visiting Scientist report n.1 ,February, 2014.

HSAF product description

Mugnai, A., Casella, D., Cattani, E., Dietrich, S., Laviola, S., Lev- izzani, V., Panegrossi, G., Petracca, M., Sano', P., Di Paola, F., Biron, D., De Leonibus, L., Melfi, D., Rosci, P., Vocino, A., Za- uli, F., Puca, S., Rinollo, A., Milani, L., Porcu', F., and Gattari, Nat. Hazards Earth Syst. Sci., 13, 1959–1981,2013www.nat-hazards-earth-syst-sci.net/13/1959/2013/ doi:10.5194/nhess-13-1959-2013.

Puca, S., Baguis, P., Campione, E., Ertu ̈rk, A., Gabellani, S., Iwan ́ski, R., Juras ̆ek, M., Kan ̆a ́k, J., Kere ́nyi, J., Koshinchanov, G., Kozinarova, G., Krahe, P., Łapeta, B., La ́bo ́, E., Milani, L., Okon,O.,O ̈ztopal,A.,Pagliara,P.,Pignone,F.,Porcu`,F.,Rachi- mow, C., Rebora, N., Rinollo, A., Roulin, E., So ̈nmez, I ̇., Toni- azzo, A., Vulpiani, G., Biron, D., Casella, D., Cattani, E., Diet- rich, S., Laviola, S., Levizzani, V., Melfi, D., Mugnai, A., Pane- grossi, G., Petracca, M., Sano`, P., Zauli, F., Rosci, P., and De Leonibus, L.: The validation service of the Hydrological SAF geostationary and polar satellite precipitation products, Nat. Hazards Earth Syst. Sci., 14, 871–889, (2014), www.nat-hazards-earth-syst-sci.net/14/871/2014/ doi:10.5194/nhess-14-871-2014.

Mugnai, A., Smith, E. A., Tripoli, G. J., Bizzarri, B., Casella, D., Dietrich, S., Di Paola, F., Panegrossi, G., and Sano', P.: CDRD and PNPR satellite passive microwave precipitation retrieval algorithms: EuroTRMM/EURAINSAT origins and H- SAF operations, Nat. Hazards Earth Syst. Sci., 13, 887–912, doi:10.5194/nhess-13-887-2013, 2013b.

Smith, E. A., Leung, H. W.-Y., Elsner, J. B., Mehta, A. V., Tripoli, G. J., Casella, D., Dietrich, S., Mugnai, A., Panegrossi, G., and Sano', P.: Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements: Part 3 – Identification of optimal meteorological tags, Nat. Hazards Earth Syst. Sci., 13, 1185–1208, doi:10.5194/nhess-13-1185- 2013, 2013.

Sano', P., Casella, D., Mugnai, A., Schiavon, G., Smith, E. A., and Tripoli, G. J.: Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements, Part 1: Algorithm description and testing, IEEE Trans. Geosci. Remote Sens., in press, 2013.

Casella, D., Panegrossi, G., Sano', P., Mugnai, A., Smith, E. A., Tripoli, G. J., Dietrich, S., Formenton, M., Di Paola, F., Le- ung, H. W.-Y., and Mehta, A. V.: Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive mi- crowave measurements, Part 2: Overcoming database profile se- lection

ambiguity by consideration of meteorological control on microphysics, IEEE Trans. Geosci. Remote Sens., in press, 2013.

TRMM product description

Villarini, G., W.F. Krajewski, 2007: Evaluation of the Research-Version TMPA Three-Hourly 0.25°x0.25° Rainfall Estimates over Oklahoma. Geopys. Res. Lett., 34, doi:10.1029/2006GL029147.

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

Ryu, Geun-Hyeok; Sohn, Byung-Ju; Kummerow, Christian; Seo, Eun-Kyoung; Tripoli, Gregory, "Improved Goddard profiling (GPROF) database over the Korean Peninsula and its impact on TRMM TMI rainfall", Remote Sensing of the Atmosphere and Clouds III. Edited by Mukai, Sonoyo; Frouin, Robert J.; Sohn, Byung-Ju. Proceedings of the SPIE, Volume 7859, article id. 78590A, 10 pp. (2010).

Olson, W. S., C. D. Kummerow, S. Yang, G. W. Petty, W.-K. Tao, T. L. Bell, S. A. Braun, Y. Wang, S. E. Lang, D. E. Johnson, and C. Chiu, 2006: Precipitation and latent heating distributions from satellite passive microwave radiometry. Part I: Improved method and uncertainties. J. Appl. Meteor. Climatol., 45, 702–720.

Wolff David B., Brad L. Fisher, COMPARISON OF INSTANTANEOUS TRMM SATELLITE AND GROUND VALIDATION RAIN RATE ESTIMATES, VOLUME 47, pp. 2215-2237, JAMC, August 2008.

Triple collocation

Stoffelen, A., 1998, Towards the true near-surface wind speed: Error modeling and calibration using triple collocation. J. Geophys. Res. 103C3, 7755-7766.

Dorigo, W.A., K. Scipal, R.M. Parinussa, Y.Y. Liu, W. Wagner, R.A.M. de Jeu, and V. Naeimi. 2010. Error characterisation of global active and passive microwave soil moisture datasets. Hydrol. Earth Syst. Sci. 14:2605–2616. doi:10.5194/hess-14-2605-2010

Portabella, M. and A. Stoffelen, 2009, On scatterometer ocean stress. J. Atm. Ocean. Tech. 26, 368-382.

Vogelzang, J., A. Stoffelen, A. Verhoef, and J. Figa-Saldaña, (2011), On the quality of high-resolution scatterometer winds. J. Geophys. Res. 116, C10033, doi:10.1029/2010JC006640.

Caires, S., A. Sterl, 2003, Validation of ocean wind and wave data using triple collocation, J. Geophys. Res. 108, 3098-3113, doi:10.1029/2002JC001491.

Janssen, P.A.E.M., S. Abdalla, H. Hersbach, and J.-R.Bidlot, 2007, Error estimates of buoy, satellite, and model wave height data. J. Atm. Ocean. Tech. 24, 1665-1677.

O'Carroll, A.G., J.R. Eyre, and R.W. Saunders, 2008, Three-way error analysis between AATSR, AMSR-E, and in-situ sea surface temperature observations. J. Atm. Ocean. Tech. 25, 1197-1207.

Scipal, K., W. Dorigo, R. deJeu, 2010, Triple collocation – a new tool to determine the error structure of global soil moisture products, Proc. IGARSS 2010, page(s): 4426 - 4429, Doi: 10.1109/IGARSS.2010.5652128, http://publik.tuwien.ac.at/files/PubDat_189396.pdf.

Hwang, B.J., and T. Lavergne (2010 Sept). Validation and Comparison of OSI SAF Low and Medium Resolution and IFREMER/Cersat Sea ice drift products, EUMETSAT Associated & Visiting Scientist Activity Report, SAF/OSI/CDOP/met.no/SCI/RP/151, EUMETSAT OSI SAF – Ocean and Sea Ice Satellite Application Facility.

Roebeling, R., E. Wolters, and H. Leijnse, 2011, Triple collocation of precipitation retrievals from SEVIRI with gridded rain gauge data and weather radar observations, http://www.knmi.nl/research/climate_observations/meetings/presentations/Roebeling_20110526. pdf.

Loew, A., and F. Schlenz. 2011. A dynamic approach for evaluating coarse scale satellite soil moisture products. Hydrol. Earth Syst. Sci. 15:75–90. doi:10.5194/hess-15-75-2011

Zwieback S., K. Scipal, W. Dorigo, and W. Wagner, 2012, "Structural and statistical properties of the collocation technique for error characterization", Nonlin. Processes Geophys., 19, 69–80, 2012 www.nonlin-processes-geophys.net/19/69/2012/ doi:10.5194/npg-19-69-2012

Miralles, D.G., W.T. Crow, and M.H. Cosh. 2010. Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. J. Hydrometeorol. 11:1423–1429. doi:10.1175/2010JHM1285.1

END OF DOCUMENT