Bull. of Atmos. Sci.& Technol. (2024) 5:9
https://doi.org/10.1007/s42865-024-00073-y

RESEARCH

Anatomy of the uncertainty of satellite vicarious calibration using radiosondes: concepts and preliminary results for microwave radiometric observations

Domenico Cimini[1,2] · Vasileios Barlakas[3,4] · Fabien Carminati[5] · Francesco De Angelis[4] · Francesco Di Paola[1] · Alessandro Fassò[6] · Donatello Gallucci[1] · Sabrina Gentile[1,2] · Tim Hewison[4] · Salvatore Larosa[1] · Fabio Madonna[1,7] · Vinia Mattioli[4] · Mario Montopoli[2,8] · Filomena Romano[1] · Marco Rosoldi[1] · Mariassunta Viggiano[1] · Axel Von Engeln[4] · Elisabetta Ricciardelli[1]

Received: 31 May 2024 / Accepted: 18 July 2024
© The Author(s) 2024

Abstract
Calibration of satellite observations is crucial for ensuring the quality of retrieved products essential for meteorological and climate applications. Calibration is obtained and monitored through a cascade of stages, including postlaunch vicarious calibration/validation activities through comparison with independent reference measurements. Here, the vicarious calibration method using radiative transfer simulations based on reference radiosondes is considered in the framework of the calibration/validation activities for the Microwave Imager (MWI) and the Ice Cloud Imager (ICI) to be launched with the Second Generation of the EUMETSAT Polar System. This paper presents an overview of the uncertainty characterizing the vicarious calibration of MWI and ICI using radiosondes as performed within the EUMETSAT-funded VICIRS study. The uncertainty characterization is pursued following a metrological approach, providing a preliminary estimation of all the identified sources. The same approach is used to develop a rigorous method for estimating the number of comparison pairs (i.e., observations vs. simulations) needed to reach a certain level of accuracy in bias determination.

Keywords Satellite observations, microwave radiometry, vicarious calibration, radiosonde

1 Introduction

Radiometric observations from space are currently the backbone of satellite meteorology and climate monitoring. Radiometric sensors are particularly sensitive to calibration, due to the relatively weak signals they are intended to detect, and thus accurate methods are crucial for ensuring the quality of retrieved products essential for meteorological and climate

Extended author information available on the last page of the article

applications. For this reason, calibration is obtained and monitored through a cascade of stages, including (i) preflight (laboratory tests on the ground prior to launch), (ii) in-flight (postlaunch aboard the spacecraft), and (iii) vicarious (postlaunch intercalibration/validation activities through comparisons with independent reference measurements). Although the preflight calibration completely characterizes the radiometric sensor, frequent checks should be performed to monitor the sensor calibration in flight, using onboard high emissivity targets and cold-space background. In addition, vicarious calibration methods are used to check the calibration status and monitor any postlaunch degradation, relying on independent reference measurements external to the satellite. Generally, vicarious calibration methods use three types of reference data sources: (i) radiative transfer (RT) simulations based on atmospheric profiles from radiosonde (RS) sensors (e.g., Moradi et al. 2013) or numerical weather prediction (NWP) data (Saunders et al. 2013; Duncan et al. 2024); (ii) similar instruments aboard aerial platforms (e.g. Wilheit 2013); and (iii) similar instruments aboard other spaceborne platforms (Moradi et al. 2015), also called intercalibration. In the case of new satellite sensors, some channels may be unprecedented aboard operating platforms and thus cannot be compared with reference observations from space. Thus, their data can be compared only with similar airborne observations (Fox et al. 2017, 2024), which are expensive and seldom available, or with RT calculations from reference profiles, such as radiosonde and modeled profiles (forecast or reanalysis). To this end, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) promoted a study to characterize the vicarious calibration of microwave (MW) radiometer imagers aboard the Second Generation of the EUMETSAT Polar System (EPS-SG), which is scheduled to be launched in 2025/2026. In particular, the EPS-SG satellite B will host the Microwave Imager (MWI) and the Ice Cloud Imager (ICI): MWI operates at 18 frequencies between 18 and 183 GHz, while ICI will be the first operational sensor covering the mm/submm wavelengths from 183 to 664 GHz. The combined use of MWI and ICI radiometers will provide an unprecedented set of microwave passive measurements, from 18.7 GHz up to 664 GHz. This paper presents an overview of the uncertainty characterizing the vicarious calibration of MWI and ICI using radiosondes as performed within the EUMETSAT-funded VICIRS study (development of VIcarious Calibration tools for MWI and ICI using RadioSoundings). Characterizing uncertainty is a fundamental step to ensure that the comparison between observations and simulations is metrologically valid. The uncertainty characterization developed in the present study represents a building block of the tools that will be used for the calibration/validation (cal/val) activities for the unprecedented observations provided by MWI and ICI after their lunch in 2025/2026.

2 Methodology

The availability of measurement and collocation uncertainties is fundamental for establishing whether spaceborne observations and RT simulations agree within the uncertainty limits (Immler et al. 2010). Therefore, the analysis of the uncertainty sources is the first and most important step in making proper use of satellite observations.

2.1 Metrological approach

Although the terms "error" and "uncertainty" are often treated as synonymous, in metrology they correspond to different definitions. According to the Vocabulaire International de Métrologie (VIM), published by the Joint Committee for Guides in Metrology (JCGM 2012) of the International Bureau of Weights and Measures (BIPM), an error is defined as the measured value of a quantity minus the reference value of the same quantity. This error results from different contributions, as for instance instrumental and collocation errors, and can assume either positive or negative values. Conversely, uncertainty is defined as a non-negative parameter quantifying the statistical properties of an ensemble of errors. In brief, the term error is used for the deviation between a single value and the corresponding reference, while the term uncertainty indicates the statistical properties of these errors. Any error can have two components: random and systematic. The bias is defined as the deviation of the measurements from the reference value arising from systematic errors (Immler et al. 2010). Although corrections are often applied to compensate for the bias, the uncertainty associated to the correction is seldomly considered. Conversely, bias correction shall be accompanied with a robust analysis of the uncertainty associated with the correction (JCGM 2008). Note that systematic errors are not necessarily fixed in time, as they may depend upon operating conditions and thus change on longer scale with respect to single measurements. In such cases, the bias shall be monitored in time and bias correction and associated uncertainty be evaluated periodically. Assuming that bias correction is applied correctly, the expected value of the random error is zero. Thus, the measurement uncertainty can be expressed by a single value, the uncertainty *u*, which is the estimated standard deviation (std) of the random error (Immler et al. 2010). This uncertainty results from the contribution of all uncertainty sources evaluated through the rule of uncertainty propagation:

$$u_y = \sqrt{\sum_{n=1}^{N} \left(\frac{\delta\, y}{\delta\, x_n}\right)^2 u_n^2 + 2\sum_{m=1}^{N} \left(\sum_{n=1(n\neq m)}^{N} \frac{\delta\, y}{\delta\, x_n} \frac{\delta\, y}{\delta\, x_m}\, u_{m,n}\right)} \tag{1}$$

where $u_n^2 = u_{n,n}$ and $u_{n,m}$ $(n \neq m)$ indicate respectively the variance and covariance of the input variables x_n. Thus, a rigorous uncertainty budget requires the evaluation of correlations between the different sources. In the simplest case of uncorrelated sources, the different uncertainty contributions can be summed quadratically, obtaining the combined standard uncertainty. Finally, once all the uncertainty sources have been identified and evaluated, two independent measurements can be cross-checked for consistency, which is validated if the measurements agree within the total uncertainty

2.2 Uncertainty model

The vicarious calibration aims to cross-check two independent measurements m_1 and m_2 of the same measurand (e.g., the brightness temperature BT) with standard uncertainties u_1 and u_2, respectively. However, a trade-off is needed between the abundance of comparison pairs, on the one hand, and on the other hand, the uncertainty of the matchup, such as imperfect spatial and temporal collocation between satellite and radiosonde measurements (Ignaccolo et al., 2015, Fassò et al. 2014). Such uncertainty sources cannot be eliminated;

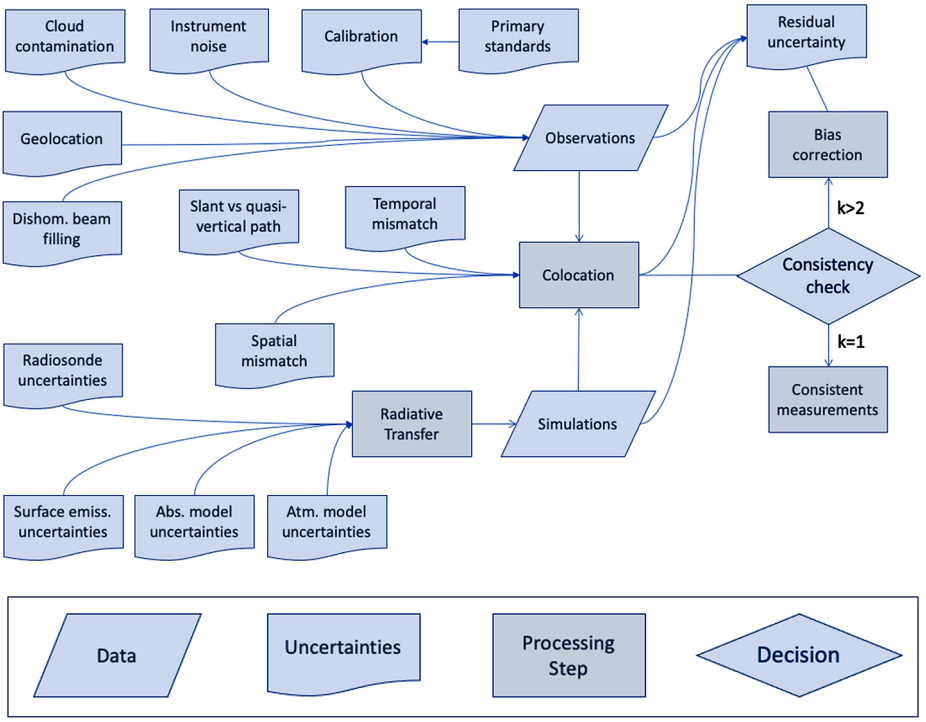

Fig. 1 Top: Uncertainty model diagram for the vicarious calibration/validation using radiosoundings. Bottom: Legends for the uncertainty model diagram blocks

therefore, the uncertainty budget of the vicarious calibration must account for not only the measurement uncertainties but also for the uncertainties related to sampling and smoothing of the inhomogeneous and variable atmospheric fields (Verhoelst et al. 2015). Calling σ the intrinsic uncertainties of the comparison (e.g. the collocation uncertainty), and assuming that m_1 and m_2 have the same expected value with normally distributed uncertainties, the probability that:

$$|m_1 - m_2| < k\sqrt{\sigma^2 + u_1^2 + u_2^2} \tag{2}$$

depends on the coverage factor k, which determines an interval about the mean value as a multiple of the standard uncertainty (Immler et al. 2010). If the results agree within $k=1$, the data are "consistent", while within $k=2$, they are "in (statistical) agreement". Conversely, if the results do not agree within $k=2$, the data are considered "significantly different", while they are considered "inconsistent" if the data do not agree within $k=3$. In the latter cases, a bias is likely present; i.e., an unaccounted systematic effect needs to be removed.

One way to visualize the uncertainty contributions is to draw a metrological uncertainty model diagram. The uncertainty model diagram consists of a visual sketch of the processing steps, describing the flow of the process, including references to calibration, uncertainty sources, and linkages to reference standards. Adopting the formalism indicated by the

GAIA-CLIM project (GAIA-CLIM 2017), an uncertainty model diagram for the vicarious calibration of satellite radiometric observations was developed within the VICIRS study, as reported in Fig. 1. The identified uncertainty sources are reviewed in the next section, including the initial evaluation performed within the VICIRS study.

Note that Eq. (2) can be inverted to estimate the number of comparison pairs (N) needed to reach the uncertainty level we aim to achieve with vicarious calibration. In fact, calling b the expected value of the bias (i.e., $b=E(m_1)-E(m_2)$) and $\widehat{b}$ its estimate ($\hat{b} = \underline{m_1} - \underline{m_2}$), the expected error for b (at the 95% confidence interval) is:

$$\left|\widehat{b} - b\right| \leq 2\frac{\sigma\,(m_1 - m_2)}{\sqrt{N}} \tag{3}$$

where $\sigma\,(m_1 - m_2)$ is the standard deviation of the comparison, which can be estimated as $\sqrt{\sigma^{\,2} + u_1^2 + u_2^2}$. Therefore, assuming we want to estimate b with an uncertainty not larger than u_b (at the 95% confidence interval), the number of comparison pairs must be:

$$N \geq \left(2\sqrt{\sigma^{\,2} + u_1^2 + u_2^2}/u_b\right)^2 \tag{4}$$

3 Results

The metrological approach and the uncertainty model outlined above have been used to list the uncertainty sources, to review their knowledge status, and to evaluate their contribution to the uncertainty of the vicarious calibration of MWI and ICI using radiosondes. Once the uncertainty sources have been identified as in Fig. 1, their quantification is likely the most difficult task. Several papers are available in the literature addressing different aspects of the problem, providing an estimate of the uncertainty, sometimes through a deep investigation, more often with just a crude guess. Thus, the relevant literature on uncertainty contributions has been reviewed within the VICIRS study. Published values were adopted for contributions already estimated and available in the open literature; for the remaining contributions, original methods are proposed to estimate their values.

3.1 Radiometer accuracy

The instruments of interest, i.e., ICI and MWI, were designed and built in response to the requirements set by the meteorological satellite user community, including radiometric accuracy (low bias) and precision (high repeatability). The instrument precision is characterized by the noise equivalent delta temperature (NEΔT). The total accuracy of a single observation is defined in the EUMETSAT end-user requirement document as the sum in quadrature of the bias (i.e., systematic uncertainty) and the NEΔT. The accuracy is obtained and maintained through preflight and in-flight calibrations, which are linked to primary or secondary metrological standards. Specifications of on-ground and in-flight instrument calibrations are reported in EUMETSAT documents (such as the EPS-SG Programme Overall Calibration and Validation Plan and the ICI and MWI Calibration and Validation Plan) and

references therein. In-flight deep space calibration data from roll maneuvers are part of cal/val activities and are used for determining important parameters for antenna pattern correction, such as spillover and near-sidelobe radiation. NEΔT depends on the channel, ranging from 0.7 to 1.3 K for MWI and from 0.8 to 2.0 K for ICI (Duncan et al. 2024).

3.2 Radiosonde measurements

Radiosonde measurements are obtained by disposable temperature and humidity sensors, whose uncertainty is characterized by the manufacturer. However, common RS archives do not provide the associated uncertainty, which is the case, for example, for the comprehensive Integrated Global Radiosonde Archive (IGRA, Durre et al., 2018). Conversely, uncertainty is characterized for each RS flight of the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN) archive. GRUAN is a high-quality low-density network that provides reference upper-air data for other more comprehensive networks. The GRUAN RS uncertainty is characterized through laboratory and inflight tests, independently of the manufacturer (Dirksen et al. 2014, 2020; Sommer et al. 2023). The resulting uncertainty is provided within the operational GRUAN standard radiosonde products. Building on the GRUAN expertise and radiosonde intercomparison experiments performed by the World Meteorological Organization (WMO), the RHARM (Radiosounding HARMonization) archive was generated to provide adjusted RS observations of temperature, humidity and wind with estimated uncertainties at ~700 stations worldwide (Madonna et al., 2022). Thus, RHARM is a lower-quality but higher-density archive with respect to GRUAN, available from 1978 to present. The radiosonde uncertainty profiles, either from the GRUAN or RHARM dataset, are propagated through the GRUAN processor (Carminati et al. 2019) to compute the uncertainty associated with the simulated radiosonde BT. Note that the GRUAN processor propagates the RS uncertainties into the BT space via perturbations in the temperature, humidity, and pressure profiles by summing and subtracting their total uncertainties, thus assuming complete correlation of the uncertainties at all levels (Carminati et al. 2019; Newman et al. 2020). This is a conservative assumption and the resulting uncertainty obtained in radiance space is therefore representative of the maximum uncertainty of the GRUAN component. The true RS uncertainty in BT space should be smaller than that calculated as such, as only a fraction of the total RS uncertainty is really correlated across the entire profile.

3.3 Collocation

One of the dominant uncertainty contributions comes from spatial/temporal collocation between satellite observations and radiosonde measurements. In fact, radiosondes provide in situ measurement profiles during the time of flight, while each satellite observation is nearly instantaneous and corresponds to the area observed within the field of view (FOV). The radiosonde drift distance typically ranges from a few km in the low troposphere to approximately 50 km in the lower stratosphere depending on parameters such as wind speed and direction, height above the surface, latitude and season (Seidel et al. 2011). Therefore, the collocation uncertainty is evaluated following the target area (TA) approach (Buehler et al. 2004), where the TA is defined as the circular area of a 50-km radius centered at the radiosonde launch site. The satellite observations falling into the TA are considered for

comparison, allowing the user to select either all the FOVs (more comprehensive and thus conservative) or just the FOVs nearest to the radiosonde flight ground track (more stringent collocation). The collocation uncertainty is then evaluated as the std of the BT corresponding to the selected FOV. Once the matchups have been identified, tests are applied to screen out cases affected by cloud contamination and radiosonde drifts larger than the TA radius.

3.4 RT parameterization

At the core of the GRUAN processor are two products of the EUMETSAT's Satellite Applications Facility for Numerical Weather Prediction (NWPSAF, https://nwp-saf.eumetsat.int/), namely, the fast radiative transfer code RTTOV (Saunders et al. 2018) and the Radiance Simulator (https://nwp-saf.eumetsat.int/site/software/radiance-simulator/). RTTOV is parameterized in the sense that the atmospheric optical depths are computed from the thermodynamics of each layer through regression. The regression is trained and tested against channel-integrated spectrally resolved line-by-line (LBL) reference calculations; thus, the regression uncertainty contributes to the total uncertainty. This uncertainty contribution is evaluated by the NWPSAF as part of the continuous development of RTTOV, and thus was evaluated also for the MWI and ICI channels in preparation for EPS-SG. The uncertainty contribution of RT parameterization is evaluated as the std of the differences between RTTOV and LBL BT computed for a dataset of 83 diverse profiles and six zenith angles, assuming constant unit surface emissivity, top-hat passbands, and MW v13 predictor coefficient files (NWPSAF 2023). The results show that the std is below 0.1 K for all the MWI and ICI channels.

3.5 LBL absorption model

The reference LBL absorption model is also affected by uncertainty, due to the computational or experimental uncertainty underlying the adopted values of spectroscopic parameters. The evaluation of the absorption model uncertainty was first assessed for downwelling radiation (Cimini et al. 2018, 2019) and recently extended to upwelling radiation from centimeter to submillimeter wavelengths (16 to 700 GHz, Gallucci et al. 2024).

The absorption model uncertainty was evaluated by propagating the covariance matrix of 135 identified dominant water vapor and oxygen spectroscopic parameters. While ozone also contributes to line absorption in this range, the uncertainty was found lower than 0.1 K for ICI/MWI channels and was thus deemed negligible. The simulated observation geometry mimics the observations from MWI and ICI, i.e., downlooking from the top of the atmosphere at 53° incident angle. The emissivity of a sea background (covering 72% of the globe) is considered, assuming typical conditions (8 m/s wind speed; 290 K sea surface temperature; 35 PSU salinity). The uncertainty was evaluated for six typical climatology conditions (tropical, midlatitude summer, midlatitude winter, subarctic summer, subarctic winter, U.S. standard).

3.6 Spectral response function

The absorption model uncertainties were estimated at the MWI and ICI channels by convolving uncertainty spectra at 50 MHz spectral resolution within a first-order approximation

to the instrument bandpass filters (i.e., a rectangular box-average). Modeling the channel spectral response function (SRF) with idealized profiles, such as rectangular bandpasses, adds uncertainty to the simulated BT. Even when SRF profiles are measured, their discretization contributes to BT uncertainty. At the time of writing, RTTOV coefficients implement the measured SRF for ICI, but not for MWI. However, Buehler et al. (2004) evaluated the impact of SRF modeling, reporting that BT differences are well below 0.1 K for either rectangular or Gaussian bandpass shapes; thus, this uncertainty is deemed negligible.

3.7 Vertical discretization

In the RT computation, the naturally continuous atmospheric profiles are represented by discrete levels, which also generates uncertainty. As such, the uncertainty decreases as the profile vertical resolution increases. GRUAN radiosonde data are provided at the original high resolution, i.e., at thousands of levels from the surface to the balloon burst height (aiming to reach pressure levels below 10 hPa). Conversely, RHARM data are available at much fewer pressure levels, providing far less detail than high resolution data. However, Buehler et al. (2004) reported that low-vertical-resolution data, as found in operational archives, are sufficient for accurately calculating satellite radiances, provided that low-resolution data are interpolated to a fine grid before calculating the radiative transfer. The 54 pressure levels used within RTTOV are deemed sufficient for accurately calculating satellite radiances and other path-integrated quantities (e.g., integrated water vapor). The vertical interpolation uncertainty was evaluated for channels of the Advanced Technology Microwave Sounder (ATMS) at the EUMETSAT NWPSAF (Hocking 2014). These results were extended within the VICIRS study to MWI and ICI channels, most of which closely match to one correspondent ATMS channel. For the remaining MWI and ICI channels, which do not closely match to one ATMS channel, the analysis extrapolated the vertical discretization uncertainty, assuming a linear relationship with atmospheric opacity. The relationship was estimated separately for the MWI and ICI channels, computing the atmospheric opacity from standard atmosphere profiles. It is recognized that this method likely overestimates the uncertainty of most opaque ICI channels and should be then further investigated.

3.8 Geolocation

The uncertainty in geolocating the instantaneous FOV also contributes to the uncertainty of the observed BT. The geolocation requirement for MWI and ICI sets in the EUMETSAT user requirements document is 2.5 km. However, the geolocation uncertainty of an operating MW conical scanner (the Special Sensor Microwave Imager/Sounder, SSMIS) was evaluated in a previous EUMETSAT-funded project (Papa et al. 2021), which reported average geolocation errors between 5.4 and 6.2 km, as estimated by looking at the boundaries of ice shelves, mountainous lakes, and sea bays. To be conservative, we assumed an average geolocation uncertainty of 6 km. The corresponding BT uncertainty was evaluated for each MWI and ICI channel using simulated test data provided by EUMETSAT, as the BT variability (std) over a 3-by-5 box (3 along-track, 5 across-track). Since the along-track and cross-track distances are ~9 and ~2 km, respectively, such a box corresponds to an area of 144 km^2, which is larger than the circle corresponding to a 6 km geolocation uncertainty (~113 km^2). The BT uncertainty was evaluated for each MWI and ICI channel as the aver-

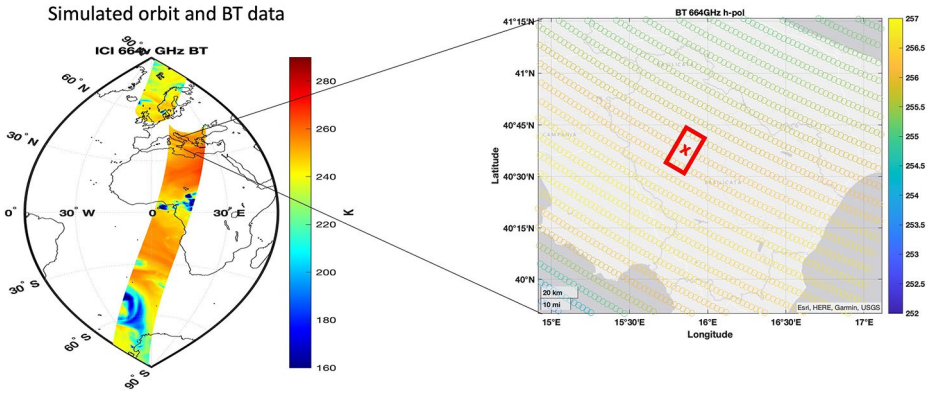

Fig. 2 One full orbit of simulated ICI test data at 664 GHz (from 2007/09/12 08:43:21 to 2007/09/12 10:22:24) and magnification of one 3-by-5 box used to estimate the contribution of geolocation uncertainty

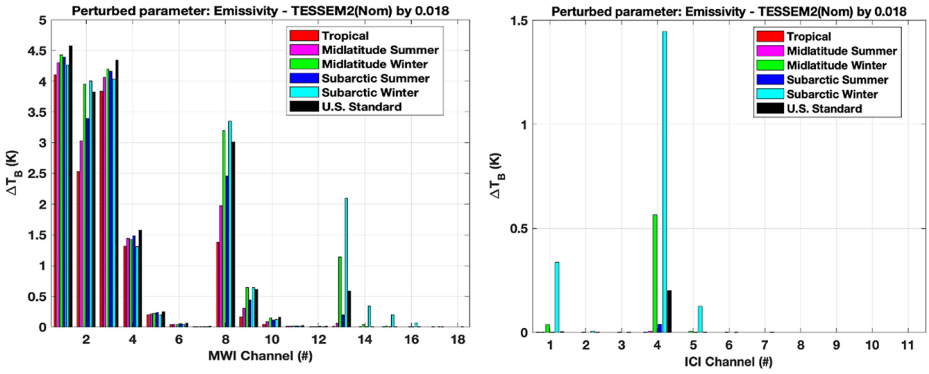

Fig. 3 Uncertainty of simulated BT for MWI (left) and ICI (right) channels due to uncertainties in sea surface emissivity estimated from Kilic et al. (2023). The bars indicate the difference between two spectra (control and control perturbed with 1-std uncertainty). Downlooking view from the top of the atmosphere at 53° incident angle. The colored bars indicate the six typical climatological conditions considered (tropical, midlatitude summer, midlatitude winter, subarctic summer, subarctic winter, U.S. standard)

age of the standard deviation across more than three million 3-by-5 boxes extracted from one entire orbit of test data (from 2007/09/12 08:43:21 to 2007/09/12 10:22:24, processed on 2022/06/13 10:44:00), as shown in Fig. 2. The estimated uncertainty depends on the channel frequency and polarization, ranging from ~0.1 to 0.9 K for MWI and from ~0.1 to 0.6 K for ICI. The values are greater for H-pol than for the corresponding V-pol channels. However, it must be noted that the simulated test data were produced using NWP model reanalysis and thus likely represent lower spatial resolution than real MWI and ICI data. This contribution will be refined once real MWI and ICI observations become available.

3.9 Surface emissivity

Surface emissivity affects the outgoing radiation from the Earth's surface and thus modulates the background radiation traveling through the atmosphere and reaching the space-

borne radiometers. To our knowledge, the contribution of surface emissivity uncertainty to the uncertainty of brightness temperature simulations has not been quantified before. Quantification of the uncertainty affecting surface emissivity modeling is available at some channels and under certain conditions, while uncertainty propagation to BT simulations is currently lacking. Therefore, a dedicated analysis was performed within the VICIRS study. Surface emissivity models, such as TELSEM2 (Wang et al. 2017) and SURFEM (Kilic et al. 2023), are distributed with RTTOV and thus are considered here. While SURFEM provides parameterized sea surface emissivity, TELSEM2 provides parameterized surface emissivity for land, snow and sea ice. These models are commonly used for MWI and ICI simulations, albeit with some limitations (e.g., no frequency dependence of sea-ice emissivity above 183 GHz due to the lack of available information). TELSEM2 emissivity has been validated up to 325 GHz against airborne observations from the International Submillimeter Airborne Radiometer (ISMAR) and the Microwave Airborne Radiometer Scanning System (MARSS), reporting consistent estimates in spatially homogeneous regions, especially at 89 and 157 GHz (Wang et al. 2017). The reported biases and standard deviations for the modeled vs. retrieved emissivity are on the order of 0.01 and 0.04, respectively. Thus, a conservative value of 0.05 uncertainty was assumed for land surface emissivity. Similarly, a 0.018 uncertainty for sea surface emissivity is assumed here, as derived from the data reported by Kilic et al. (2023). Thus, the uncertainty in the surface emissivity has been propagated to simulated BT considering the six climatological conditions introduced above to quantify the BT uncertainty due to surface emissivity. The results are reported in Fig. 3 for both MWI and ICI, obtained as the difference between two spectra (control and control perturbed by one std uncertainty). As one may expect, the surface emissivity uncertainty leads to a large BT uncertainty at most transparent lower-frequency channels. Note that the impact of surface emissivity does depend on the atmospheric conditions affecting atmospheric opacity. The most evident case is at 243 GHz (ICI channel 4), for which the contribution of surface emissivity uncertainty is negligible (<0.1 K) under warm and humid conditions (tropical and midlatitude summer) but becomes substantial (0.5–1.5 K) under cold and dry conditions (e.g., midlatitude and subarctic winters).

3.10 Total uncertainty

The uncertainty contributions evaluated above are then combined following Eq. (1), considering no correlation between them, to provide an estimate of the total uncertainty of the vicarious calibration of MWI and ICI using radiosondes. The expected total uncertainty is reported in Fig. 4 for both MWI and ICI channels; results show that the total uncertainty for the vicarious calibration is below 1 K for most of the MWI and ICI channels, but for MWI most transparent channels (e.g., 1–4), heavily impacted by surface emissivity, and ICI most opaque channels (e.g., 9–11). Note that these values are preliminary and will be re-evaluated after further investigations. For example, the uncertainty at most opaque ICI channels (e.g., 9 and 10) is likely overestimated due to the assumption of linear relationship with atmospheric opacity, as discussed in the vertical discretization section. In addition, the review of uncertainty sources indicated remaining knowledge gaps; suggestions for their future evaluation are given in the next section. However, the estimated values in Fig. 4 are already useful for an initial planning of the MWI and ICI cal/val campaign that will follow their launch. In fact, one important aspect for planning the cal/val activities is how many comparison pairs

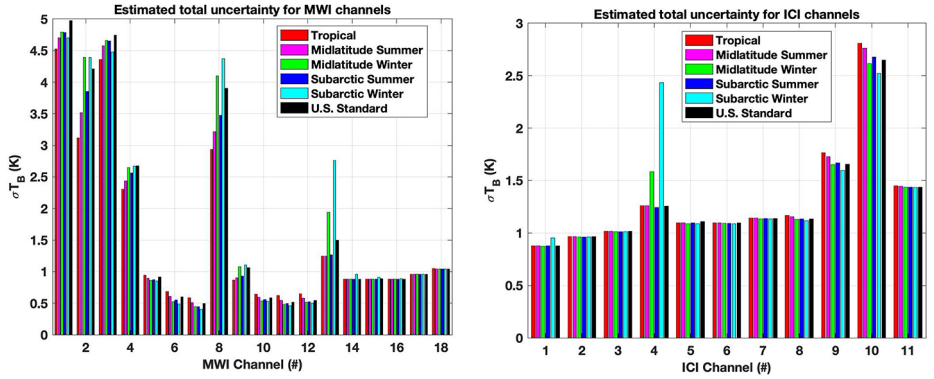

Fig. 4 Typical values of total uncertainty for the MWI (left) and ICI (right) channels estimated as described within the main text. The colored bars indicate the six typical climatological conditions considered (tropical, midlatitude summer, midlatitude winter, subarctic summer, subarctic winter, U.S. standard). Note that the results are preliminary and will be re-evaluated after further investigations, as detailed in the main text. The uncertainty values at ICI channels 9 and 10 are likely overestimated (see the section on vertical discretization)

are needed for a robust estimation of the bias. This can be evaluated rigorously through Eq. (4) by considering the desired uncertainty for bias determination (u_b) and replacing with the estimated total uncertainty. Since the estimated total uncertainty depends on the instrument, channel, and climatology, the resulting N would as well. Considering an objective uncertainty u_b =0.2 K for the bias and typical uncertainties as in Fig. 1, the resulting N is ≤100 for channels with total uncertainty ≤1 K, while N > 2000 for uncertainty ≥4.5 K. The resulting N would increase (decrease) as more (less) stringent requirements are set for u_b.

4 Summary and outlook

The VICIRS study has identified various sources of uncertainty that impact the vicarious calibration of satellite radiometric MW observations using radiosondes. It also provided an initial quantification of the overall uncertainty for MWI and ICI vicarious calibration, which is useful for planning their postlaunch cal/val activities. While some of these sources have been evaluated within the VICIRS study, others remain uncharacterized, indicating knowledge gaps. For these remaining sources, suggestions for future analysis are given hereafter. One such uncharacterized source of uncertainty arises from the spatial and temporal representativeness of radiosonde data within the adopted collocation criteria, that is, how well the radiosonde profiles represent the atmospheric spatial and temporal variability within the selected target area and time window. The radiosonde temporal representativeness could be characterized by analyzing the typical temporal variability within different time windows mapped into clear-sky simulated BTs, i.e., computing simulated clear-sky BTs from a dataset of realistic atmospheric profiles (e.g., high resolution reanalysis) for the same site but within time windows of different amplitudes (e.g., +/-1 h to +/-3 h), and then evaluating how the difference changes with temporal distance. This likely depends upon meteorological conditions, and could be characterized through proxies, such as total column water vapor, instability indices, convective available potential energy, or wind speed/direc-

tion. The radiosonde spatial representativeness could be evaluated following the analysis outlined in Calbet et al. (2018, 2022), who derived the spatial structure functions (closely related to autocorrelation) of atmospheric water vapor and temperature from sequential radiosonde launches. Considering that water vapor and temperature are the main drivers of MWI/ICI simulated observations, their structure functions could be used as proxies for computing the BT autocovariance function. Other approaches may also be considered, e.g., exploiting (*i*) available NWP data to evaluate the variability of BT for the set of NWP profiles falling within the considered target area, or (*ii*) available airborne observations (e.g., Fox et al. 2024) to compute the BT autocorrelation function for each channel. Another unaccounted uncertainty contribution is the contamination of undetected clouds within the field of view, i.e., relatively thin clouds with small water amounts that are not detected by the applied cloud tests. These clouds cause a residual signal that is not modeled by clear-sky RT calculations. Brogniez et al. (2016) evaluated the uncertainty related to undetected clouds for 183 GHz channels by comparing observations detected as clear-sky with simulations from ECMWF profiles in either clear-sky or all-sky computations, showing that the all-sky calculations lead to smaller biases in the lower peaking channels (e.g. by 0.4 K in the 183 ± 7 GHz channel). A similar approach may be extended to other MWI and ICI channels. Another contribution that has not been considered explicitly is the uncertainty related to the surface parameters not directly measured by the radiosonde, e.g., the skin temperature (English 2008). If NWP surface data are used to cover this measurement gap, their uncertainty shall be propagated into the radiative transfer to evaluate the contribution to BT simulations. Finally, the vertical correlation of radiosonde uncertainty between levels has not yet been considered, although it is recognized to potentially have a substantial impact on the estimated BT uncertainty from radiosondes, as indicated by Calbet et al. (2017) for infrared hyperspectral observations. To this end, a working group has been established within GRUAN to exploit the available information about radiosonde uncertainty vertical correlation in radiative transfer calculations. The results are expected to be reported at the next GRUAN coordination and implementation meeting (ICM-16, scheduled for fall 2025). Finally, although full metrological closure is unlikely to be achieved, future work will be dedicated to advancing the awareness and quantification of the missing uncertainty contributions and in their implementation within the EUMETSAT postlaunch cal/val procedures for MWI and ICI.

Acknowledgements The authors acknowledge the support of EUMETSAT studies ISMAR (contract EUM/CO/20/4600002477/VM) and VICIRS (contract EUM/CO/22/4600002714/FDA). The use of GRUAN data with associated uncertainties is acknowledged. Use of the RHARM data as stated in the Copernicus license agreement is also acknowledged.

Author contributions D.C., F.D.A., V.M., F.R., and E.R. designed the research. E.R. led the VICIRS proposal. A.F., F.C., F.M., T.H., A.V.E. helped in shaping the theoretical approach. F.D.P., D.G., S.G., S.L., M.M., F.R., M.R., M.V. contributed to data processing and analysis. V.B. contributed to testing and validation data analysis. D.C. wrote the main manuscript text and prepared Figs. 1, 2, 3 and 4. All authors reviewed the manuscript.

Funding This work was supported by EUMETSAT through the ISMAR (contract EUM/CO/20/4600002477/VM) and VICIRS (contract EUM/CO/22/4600002714/FDA) studies.
Open access funding provided by Consiglio Nazionale Delle Ricerche (CNR) within the CRUI-CARE Agreement.

Data availability All data available upon request.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Brogniez H, English S, Mahfouf J-F, Behrendt A, Berg W, Boukabara S, Buehler SA, Chambon P, Gambacorta A, Geer A, Ingram W, Kursinski ER, Matricardi M, Odintsova TA, Payne VH, Thorne PW, Tretyakov MYu, Wang J (2016) A review of sources of systematic errors and uncertainties in observations and simulations at 183 GHz. Atmos Meas Tech 9(5):2207–2221. https://doi.org/10.5194/amt-9-2207-2016

Buehler SA, Kuvatov M, John VO, Leiterer U, Dier H (2004) Comparison of microwave satellite humidity data and radiosonde profiles: a case study. J Geophys Res 109(D13):2004JD004605. https://doi.org/10.1029/2004JD004605

Calbet X, Carbajal Henken C, DeSouza-Machado S, Sun B, Reale T (2022) Horizontal small-scale variability of water vapor in the atmosphere: implications for intercomparison of data from different measuring systems. Atmos Meas Tech 15(23):7105–7118. https://doi.org/10.5194/amt-15-7105-2022

Calbet X, Peinado-Galan N, DeSouza-Machado S, Kursinski ER, Oria P, Ward D, Otarola A, Rípodas P, Kivi R (2018) Can turbulence within the field of view cause significant biases in radiative transfer modeling at the 183 GHz band? Atmos Meas Tech 11(12):6409–6417. https://doi.org/10.5194/amt-11-6409-2018

Calbet X, Peinado-Galan N, Rípodas P, Trent T, Dirksen R, Sommer M (2017) Consistency between GRUAN sondes, LBLRTM and IASI. Atmos Meas Tech 10(6):2323–2335. https://doi.org/10.5194/amt-10-2323-2017

Carminati F, Migliorini S, Ingleby B, Bell W, Lawrence H, Newman S, Hocking J, Smith A (2019) Using reference radiosondes to characterise NWP model uncertainty for improved satellite calibration and validation. Atmos Meas Tech 12(1):83–106. https://doi.org/10.5194/amt-12-83-2019

Cimini D, Hocking J, De Angelis F, Cersosimo A, Di Paola F, Gallucci D, Gentile S, Geraldi E, Larosa S, Nilo S, Romano F, Ricciardelli E, Ripepi E, Viggiano M, Luini L, Riva C, Marzano FS, Martinet P, Song YY, Ahn MH, Rosenkranz PW (2019) RTTOV-gb v1.0– updates on sensors, absorption models, uncertainty, and availability. Geosci Model Dev 12(5):1833–1845. https://doi.org/10.5194/gmd-12-1833-2019

Cimini D, Rosenkranz PW, Tretyakov MY, Koshelev MA, Romano F (2018) Uncertainty of atmospheric microwave absorption model: impact on ground-based radiometer simulations and retrievals. Atmos Chem Phys 18(20):15231–15259. https://doi.org/10.5194/acp-18-15231-2018

Dirksen RJ, Bodeker GE, Thorne PW, Merlone A, Reale T, Wang J, Hurst DF, Demoz BB, Gardiner TD, Ingleby B, Sommer M, Von Rohden C, Leblanc T (2020) Managing the transition from Vaisala RS92 to RS41 radiosondes within the global climate observing System Reference Upper-Air Network (GRUAN): a progress report. Geosci Instrum Method Data Syst 9(2):337–355. https://doi.org/10.5194/gi-9-337-2020

Dirksen RJ, Sommer M, Immler FJ, Hurst DF, Kivi R, Vömel H (2014) Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde. Atmos Meas Tech 7(12):4463–4490. https://doi.org/10.5194/amt-7-4463-2014

Duncan DI, Geer AJ, Bormann N, Dahoui M (2024) Vicarious calibration monitoring for MWI and ICI using NWP fields. ECWMF Contract Report Serieshttps://doi.org/10.21957/7c2d18d2e1. Report no: EUM/CO/22/4600002673/SDM

Durre I, Yin X, Vose RS, Applequist S, Arnfield J (2018) Enhancing the Data Coverage in the Integrated Global Radiosonde Archive. Journal of Atmospheric and Oceanic Technology 35(9):1753–1770. https://doi.org/10.1175/JTECH-D-17-0223.1

English SJ (2008) The importance of Accurate skin temperature in assimilating Radiances from Satellite Sounding instruments. IEEE Trans Geosci Remote Sens 46(2):403–408. https://doi.org/10.1109/TGRS.2007.902413

Fassò A, Ignaccolo R, Madonna F, Demoz B, Franco-Villoria M (2014) Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles. Atmos Meas Tech 7:1803–1816. https://doi.org/10.5194/amt-7-1803-2014

Fox S, Lee C, Moyna B, Philipp M, Rule I, Rogers S, King R, Oldfield M, Rea S, Henry M, Wang H, Harlow RC (2017) ISMAR: an airborne submillimetre radiometer. Atmos Meas Tech 10(2):477–490. https://doi.org/10.5194/amt-10-477-2017

Fox S, Mattioli V, Turner E, Vance A, Cimini D, Gallucci D (2024) An evaluation of atmospheric absorption models at millimetre and sub-millimetre wavelengths using airborne observations EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-229

GAIA-CLIM (2017) Traceability model diagrams. http://www.gaia-clim.eu/page/traceability-model-diagrams. Accessed 20 December 2017

Gallucci D, Cimini D, Turner E, Fox S, Rosenkranz PW, Tretyakov MY, Mattioli V, Larosa S, Romano F (2024) Uncertainty in simulated brightness temperature due to sensitivity to atmospheric gas spectroscopic parameters from the centimeter- to submillimeter-wave range. Atmos Chem Phys 24:7283–7308. https://doi.org/10.5194/acp-24-7283-2024

Hocking J (2014) NWPSAF report Technical Report No: 590, Interpolation methods in the RTTOV radiative transfer model. https://digital.nmla.metoffice.gov.uk/download/file/digitalFile_911bd873-f30f-4617-9810-ad73b5457ea1, last access: 06/05/2024

Ignaccolo R, Franco-Villoria M, Fassò A (2015) Modelling collocation uncertainty of 3D atmospheric profiles. Stoch Env Res Risk Assess 29(2):417–429. https://doi.org/10.1007/s00477-014-0890-7

Immler FJ, Dykema J, Gardiner T, Whiteman DN, Thorne PW, Vömel H (2010) Reference Quality Upper-Air measurements: guidance for developing GRUAN data products. Atmos Meas Tech 3(5):1217–1231. https://doi.org/10.5194/amt-3-1217-2010

JCGM (2008) Evaluation of measurement data– Guide to the expression of uncertainty in measurement (GUM), Tech. Rep. JCGM 100: 2008, International Bureau of Weights and Measures (BIPM), https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf

JCGM (2012) International Vocabulary of Metrology– Basic and General Concepts and Associated Terms (VIM3), Tech. Rep. JCGM 200: 2012, International Bureau of Weights and Measures (BIPM), https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf

Kilic L, Prigent C, Jimenez C, Turner E, Hocking J, English S, Meissner T, Dinnat E (2023) Development of the SURface fast Emissivity Model for Ocean (SURFEM-Ocean) based on the PARMIO Radiative transfer model. Earth Space Sci 10(11):e2022EA002785. https://doi.org/10.1029/2022EA002785

Madonna F, Tramutola E, Sy S, Serva F, Proto M, Rosoldi M, Gagliardi S, Amato F, Marra F, Fassò A, Gardiner T, Thorne PW (2022) The New Radiosounding HARMonization (RHARM) Data Set of Homogenized Radiosounding temperature, humidity, and wind profiles with uncertainties. JGR Atmos 127(2):e2021JD035220. https://doi.org/10.1029/2021JD035220

Moradi I, Ferraro RR, Eriksson P, Weng F (2015) Intercalibration and Validation of observations from ATMS and SAPHIR Microwave sounders. IEEE Trans Geosci Remote Sens 53(11):5915–5925. https://doi.org/10.1109/TGRS.2015.2427165

Moradi I, Soden B, Ferraro R, Arkin P, Vömel H (2013) Assessing the quality of humidity measurements from global operational radiosonde sensors. JGR Atmos 118(14):8040–8053. https://doi.org/10.1002/jgrd.50589

Newman S, Carminati F, Lawrence H, Bormann N, Salonen K, Bell W (2020) Assessment of New Satellite missions within the Framework of Numerical Weather Prediction. Remote Sens 12(10):1580. https://doi.org/10.3390/rs12101580

NWPSAF (2023) MW optical depth coefs and RTTOV-SCATT optical properties. https://nwp-saf.eumetsat.int/site/software/rttov/download/coefficients/coefficient-download/#MW_optical_depth_coefs_and_RTTOV-SCATT_optical_properties Accessed 30 March 2023

Papa P, Mattioli M, Avbelj V, FS Marzano J (2021) Assessing the Spaceborne 183.31-GHz Radiometric Channel Geolocation using high-Altitude Lakes, Ice shelves, and SAR Imagery. IEEE Trans Geosci Remote Sens 59(5):4044–4061. https://doi.org/10.1109/TGRS.2020.3024677

Saunders R, Hocking J, Turner E, Rayer P, Rundle D, Brunel P, Vidot J, Roquet P, Matricardi M, Geer A, Bormann N, Lupu C (2018) An update on the RTTOV fast radiative transfer model (currently at version 12). Geosci Model Dev 11(7):2717–2737. https://doi.org/10.5194/gmd-11-2717-2018

Saunders RW, Blackmore TA, Candy B, Francis PN, Hewison TJ (2013) Monitoring Satellite Radiance biases using NWP models. IEEE Trans Geosci Remote Sens 51(3):1124–1138. https://doi.org/10.1109/TGRS.2012.2229283

Seidel DJ, Sun B, Pettey M, Reale A (2011) Global radiosonde balloon drift statistics. J Geophys Res 116(D7):D07102. https://doi.org/10.1029/2010JD014891

Sommer M, von Rohden C, Simeonov T, Oelsner P, Naebert T, Romanens G, Jauhiainen H, Survo P, Dirksen R (2023) GRUAN characterisation and data processing of the Vaisala RS41 radiosonde, GRUAN-TD-8, v1.0.0 https://www.gruan.org/documentation/gruan/td/gruan-td-8

Verhoelst T, Granville J, Hendrick F, Köhler U, Lerot C, Pommereau J-P, Redondas A, Van Roozendael M, Lambert J-C (2015) Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons. Atmos Meas Tech 8(12):5039–5062. https://doi.org/10.5194/amt-8-5039-2015

Wang D, Prigent C, Kilic L, Fox S, Harlow C, Jimenez C, Aires F, Grassotti C, Karbou F (2017) Surface emissivity at microwaves to millimeter waves over Polar regions: parameterization and evaluation with aircraft experiments. J Atmos Ocean Technol 34(5):1039–1059. https://doi.org/10.1175/JTECH-D-16-0188.1

Wilheit TT (2013) Comparing calibrations of similar Conically scanning window-Channel Microwave Radiometers. IEEE Trans Geosci Remote Sens 51(3):1453–1464. https://doi.org/10.1109/TGRS.2012.2207122

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Domenico Cimini[1,2] · Vasileios Barlakas[3,4] · Fabien Carminati[5] · Francesco De Angelis[4] · Francesco Di Paola[1] · Alessandro Fassò[6] · Donatello Gallucci[1] · Sabrina Gentile[1,2] · Tim Hewison[4] · Salvatore Larosa[1] · Fabio Madonna[1,7] · Vinia Mattioli[4] · Mario Montopoli[2,8] · Filomena Romano[1] · Marco Rosoldi[1] · Mariassunta Viggiano[1] · Axel Von Engeln[4] · Elisabetta Ricciardelli[1]

✉ Domenico Cimini
domenico.cimini@cnr.it

[1] National Research Council of Italy– Institute of Methodologies for Environmental Analysis (CNR-IMAA), Potenza, Italy

[2] Center of Excellence CETEMPS, University of L'Aquila, L'Aquila, Italy

[3] CS Group– Germany GmbH, ECOS Office Center, Darmstadt, Germany

[4] EUMETSAT, Darmstadt, Germany

[5] Met Office, Exeter, UK

[6] Department of Economics, University of Bergamo, Bergamo, Italy

[7] Department of Physics, University of Salerno, Salerno, Italy

[8] National Research Council of Italy, Institute of Atmospheric and Climate Sciences (CNR-ISAC), Rome, Italy