Spectrum Synergy for Investigating Cloud Microphysics

Domenico Cimini⬤, Carmine Serio, Guido Masiello, Pietro Mastro,
Elisabetta Ricciardelli, Francesco Di Paola, Salvatore Larosa, Donatello Gallucci,
Tim Hultberg, Thomas August, and Filomena Romano

ABSTRACT: Observations from spaceborne microwave (MW) and infrared (IR) passive sensors are the backbone of current satellite meteorology, essential for data assimilation into modern numerical weather prediction and for climate benchmarking. While MW and IR observations from space offer complementary features with respect to cloud properties, their synergy for cloud investigation is currently underexplored, despite the presence of both MW and IR sensors on operational meteorological satellites such as the EUMETSAT Polar System (EPS) MetOp series. As such, several key cloud microphysical properties are not part of the operational products available from EPS MetOp sensors. In addition, the EPS Second Generation (EPS-SG) series, scheduled for launch starting from 2024 onward, will carry sensors such as the Microwave Sounder (MWS) and IASI Next Generation (IASI-NG), enhancing spatial and spectral resolutions and thus capacity to retrieve cloud properties. This article presents the Combined MWS and IASI-NG Soundings for Cloud Properties (ComboCloud) project, funded by EUMETSAT with the overall objective to specify, prototype, and validate algorithms for the retrieval of cloud microphysical properties (e.g., water content and drop effective radius) from the synergy of passive MW and IR observations. The article presents the synergy rationale, the algorithm design, and the results obtained exploiting simulated observations from EPS and EPS-SG sensors, quantifying the benefits to be expected from the MW–IR synergy and the new generation sensors.

KEYWORDS: Atmosphere; Clouds; Atmospheric profilers; Satellite observations; Cloud microphysics

AFFILIATIONS: Cimini, Ricciardelli, Di Paola, Larosa, Gallucci, and Romano—Institute of Methodologies for Environmental Analysis, National Research Council of Italy, Tito Scalo, Italy; **Serio, Masiello, and Mastro**—School of Engineering, University of Basilicata, Potenza, Italy; **Hultberg and August**—European Organisation for the Exploitation of Meteorological Satellites, Darmstadt, Germany

C louds play a key role at all different scales, from local weather to global climate (Fu et al. 2011; Stevens and Bony 2013; Stephens et al. 2015; Bony et al. 2015). Albeit measurements of cloud microphysics are possible with in situ cloud probes, these become impractical at a global scale. Conversely, observations from space offer the capability to monitor clouds over the entire globe, investigating cloud–radiation interactions and cloud microphysical properties (Stephens and Kummerow 2007; Stephens et al. 2018). Therefore, the bulk of cloud information comes from satellite observations, in particular from radiometric (passive) measurements of the upwelling electromagnetic radiation from the Earth–atmosphere system (Geer et al. 2019). Several spaceborne sensors, operating in the visible, infrared (IR), and microwave (MW) sections of the electromagnetic spectrum, are nowadays providing cloud products. In addition to the products extracted from a single sensor, more can be inferred from sensor synergy, leveraging on complementary features of atmosphere–radiation interaction at different wavelengths (August et al. 2012; Nalli et al. 2018; Smith and Barnet 2019). However, despite satellite MW and IR observations offer complementary information on cloud microphysics (Romano et al. 2007), their synergy is currently underexplored, with no operational cloud product exploiting that. In this context, ongoing research aims at addressing this gap by developing combined MW–IR algorithms for the retrieval of cloud microphysics products from next-generation polar-orbiting satellites to be operative for the next two decades (Mastro et al. 2022).

Rationale

Radiometric passive sensors measure the natural radiation coming from the observed scene, which consists of the integrated contributions from Earth's surface and the atmospheric constituents (gas, cloud particles, aerosols) along the path (Stephens and Kummerow 2007). Radiation interacts with cloud particles in such a way that the greatest sensitivity to the microphysical details of cloud in the atmosphere occurs when the wavelength of the radiation is close to the size of the particles (Geer et al. 2019): while shorter wavelengths (e.g., visible and infrared) are more sensitive to smaller particles, longer wavelengths (e.g., microwave) are more sensitive to larger particles. In addition, shorter wavelengths are more sensitive to the amount of cloud particles and water content, such that they see thin clouds that are transparent to longer wavelengths. On the other hand, shorter wavelengths are blocked by the top of thicker clouds, while longer wavelengths may be able to see throughout the cloud deck. This situation is pictured in Fig. 1. More quantitatively, Fig. 2 shows the sensitivity of satellite observations at typical IR and MW wavelengths to the integrated cloud liquid water path (CLWP). While IR is very sensitive to small water amounts, decreasing by up to 1.6 K $(g\ m^{-2})^{-1}$ of CLWP when CLWP < 25 g m^{-2}, it quickly saturates as the CLWP reaches 50 g m^{-2}, as typical even for relatively thin clouds (Kniffka et al. 2014). Conversely, MW shows a rather linear sensitivity with respect to CLWP [~0.05 K $(g\ m^{-2})^{-1}$], much smaller than for IR at low CLWP, but constant with increasing CLWP.

Thus, as the cloud becomes thicker, it blocks IR radiation coming from Earth's surface, and the IR radiation reaching a spaceborne radiometer tends to be generated close to the cloud top. On the other hand, MW radiation emitted from Earth's surface is only slightly affected when passing through the cloud and it carries information on the total vertical water content. A similar reasoning may be applied to ice clouds: IR radiation is greatly sensitive to low cloud ice water path (CIWP) while MW radiation is more sensitive to thicker ice clouds with higher CIWP (Sun and Weng 2012). Therefore, a proper combination of IR and MW measurements would be able to provide useful information over the whole range of cloud water paths and also help in determining the cloud vertical structure, particularly in case of overlapping cloud layers, as suggested by several investigators (e.g., Chevallier et al. 2002; Huang et al. 2006; Romano et al. 2007; Islam et al. 2014).

Current status

Nowadays, thanks to the presence of both MW and IR passive sensors on operational meteorological satellites, such as the MetOp satellites of the EUMETSAT Polar System (EPS) mission, the synergy of MW and IR observations is widely used for operational atmospheric temperature and humidity profiling in clear and cloudy conditions (August et al. 2012; Nalli et al. 2018; Smith and Barnet 2019). However, such a synergy is still underexplored for the retrieval and investigation of cloud microphysics. Despite that some previous investigations have demonstrated the feasibility and benefits of MW–IR synergy (Chevallier et al. 2002; Huang et al. 2006; Romano et al. 2007; Islam et al. 2014; Holl et al. 2014; Maddy and Boukabara 2021), several key parameters for cloud physical processes are not part of the operational level 2 products available from EPS sensors, such as the Advanced Microwave Sounding (AMSU), the Microwave Humidity Sounder (MHS), and the Infrared Atmospheric Sounding Interferometer (IASI). Among the key parameters are the cloud particle effective radius (Re), the profiles of cloud liquid water content (CLWC) and cloud ice water content (CIWC), and their path-integrated CLWP and CIWP. For these products no operational heritage from EPS is available. MW and
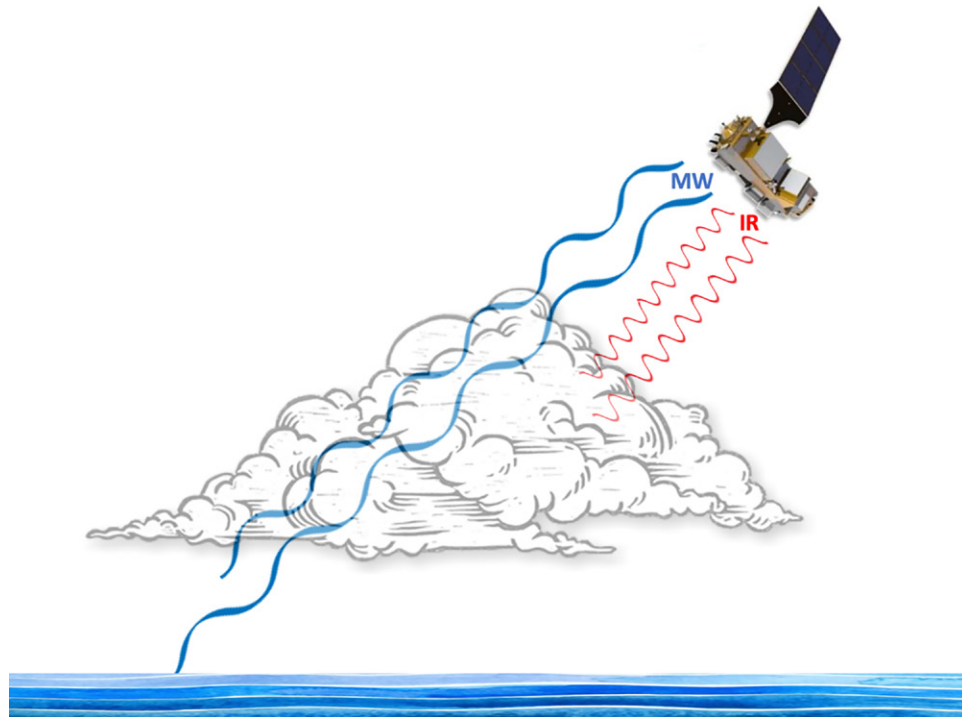

Fig. 1. Artistic view of complementary features of infrared (IR) and microwave (MW) observations of cloud properties as seen from the EUMETSAT Polar System Second Generation (EPS-SG) *MetOp-A* satellite (not in scale).

IR observations will continue to be available on future operational missions, such as EPS Second Generation (EPS-SG), with the Microwave Sounder (MWS) and IASI Next Generation (IASI-NG). The main features of these two instruments are given in the sidebar. In short, MWS and IASI-NG provide enhanced spatial and spectral resolutions, as well as lower instrumental noise, with respect to their predecessors (IASI and AMSU/MHS), allowing more sophisticated retrieval procedures to estimate new cloud products with enhanced accuracy.

In this context, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) funded a project to design, prototype, and validate algorithms for the retrieval of cloud products from combined MW–IR observations. This article summarizes the outcome of the project, named ComboCloud[1] (Combined MWS and IASI-NG Soundings for Cloud Properties), describing the retrieval approach, the retrieval performances, and the way forward.

Fig. 2. Simulated cloudy minus clear sky radiance difference at 11 μm and 31.4 GHz for a range of cloud liquid water path (CLWP). Both calculations use the same atmospheric settings: temperature and water vapor profiles are from typical midlatitude Atlantic Ocean conditions (from the 60-level diverse profile dataset available at https://nwp-saf.eumetsat.int/site/software/atmospheric-profile-data/), with 288.4-K sea surface temperature, 44.4 kg m^{-2} integrated water vapor, 280-m-thick stratiform cloud with top at 496.6 hPa, and 5-μm effective radius. The dataset considered in this work contains clouds at around this altitude for ~20% of the cases. Figure adapted to satellite observations from Turner (2007).

Algorithm development

The primary objective of the ComboCloud study is to specify, prototype, and validate retrieval algorithms for cloud microphysical properties (CLWC, CIWC, CLWP, CIWP, and Re) from the synergy of MW and IR measurements, exploiting combinations of current (AMSU/MHS and IASI) and future (MWS and IASI-NG) passive sensors. Secondary objectives are the quantitative evaluation of the benefits of (i) the MW–IR synergy with respect to MW-only and IR-only and (ii) the future sensors (MWS and IASI-NG) with respect to current ones (AMSU/MHS and IASI). The project requirements also included the development of a thin cirrus detection flag (Ricciardelli et al. 2022, manuscript submitted to *IEEE Trans. Geosci. Remote Sens.*). This activity is not reported here as it involves IR measurements only, with thin cirrus being nearly transparent to MW radiation.

Previous studies have demonstrated MW–IR synergy for cloud properties retrievals using a variety of methods, including variational (Chevallier et al. 2002), physical (Huang et al. 2006; Romano et al. 2007), and machine learning (Islam et al. 2014; Holl et al. 2014; Marke et al. 2016; Maddy and Boukabara 2021) approaches. Machine learning methods are increasingly applied to data inversion because of their ability to find nonlinear statistical relationships between target variables and input variables, such those arising in satellite remote sensing of geophysical parameters through spectral radiances (Holloway and Mengersen 2018; Belgiu and Drăguţ 2016; Izquierdo-Verdiguiera and Zurita-Millab 2020). Here, we also adopt
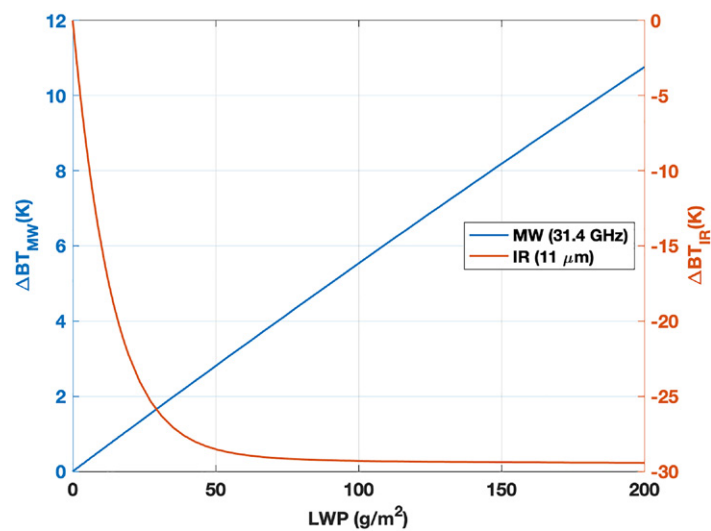
MW and IR sounders on EPS-SG (MWS and IASI-NG)

The EPS-SG Microwave Sounder (MWS; www.eumetsat.int/eps-sg-microwave-sounder) is a total power radiometer that measures the upwelling radiance. The primary objective of the MWS is to support numerical weather prediction (NWP) at regional and global scales by providing information on atmospheric temperature and water vapor profiles in clear and cloudy conditions. It is also sensitive to cloud water content. The MWS has a direct heritage from the microwave instruments AMSU-A (Advanced Microwave Sounding Unit A) and MHS (Microwave Humidity Sounder) on board EPS and NOAA satellites (Table SB1). The MWS spectral characteristic is enhanced in comparison to its predecessors by the addition of two temperature and three humidity sounding channels, including one at 229 GHz particularly sensitive to ice clouds.

Table SB1. Channel comparison between MWS and AMSU/MHS. In the AMSU/MHS combination, channels 1–15 belong to AMSU-A, channels 16–20 belong to MHS. Channels 5, 7, 20, 22, 24 of MWS are new with respect to AMSU/MHS. Channels 18, 19, 21, 23 have either central frequency, polarization, or passband differences with respect to the closest AMSU/MHS channels.

AMSU/MHS			MWS		
Ch. No.	Channel frequency (GHz)	NEΔT (K)	Ch. No.	Channel frequency (GHz)	NEΔT (K)
1	23.8	0.30	1	23.8	0.25
2	31.4	0.30	2	31.4	0.35
3	50.3	0.40	3	50.3	0.50
4	52.8	0.25	4	52.8	0.35
			5	53.246 ± 0.08	0.40
5	53.595 ± 0.115	0.25	6	53.596 ± 0.115	0.40
			7	53.948 ± 0.081	0.40
6	54.4	0.25	8	54.4	0.35
7	54.94	0.25	9	54.94	0.35
8	55.50	0.25	10	55.50	0.40
9	57.290344	0.25	11	57.290344	0.40
10	57.290344 ± 0.217	0.40	12	57.290344 ± 0.217	0.55
11	57.290344 ± 0.3222 ± 0.048	0.40	13	57.290344 ± 0.3222 ± 0.048	0.60
12	57.290344 ± 0.3222 ± 0.022	0.40	14	57.290344 ± 0.3222 ± 0.022	0.90
13	57.290344 ± 0.3222 ± 0.010	0.80	15	57.290344 ± 0.3222 ± 0.010	1.2
14	57.290344 ± 0.3222 ± 0.0045	1.20	16	57.290344 ± 0.3222 ± 0.0045	2.0
15	89.0	0.50			
16	89.0	0.22	17	89.0	0.25
17	157.0	0.38	18	165.5 ± 0.725	0.50
18	183.311 ± 1.0	0.42	23	183.311 ± 1.0	0.75
			22	183.311 ± 1.8	0.60
19	183.311 ± 3.0	0.57	21	183.311 ± 3.0	0.60
			20	183.311 ± 4.5	0.40
20	191.31	0.45	19	183.311 ± 7.0	0.40
			24	229	0.70

Table SB2. Channel comparison between IASI-NG and IASI. The corresponding noise levels are illustrated in Fig. SB1.

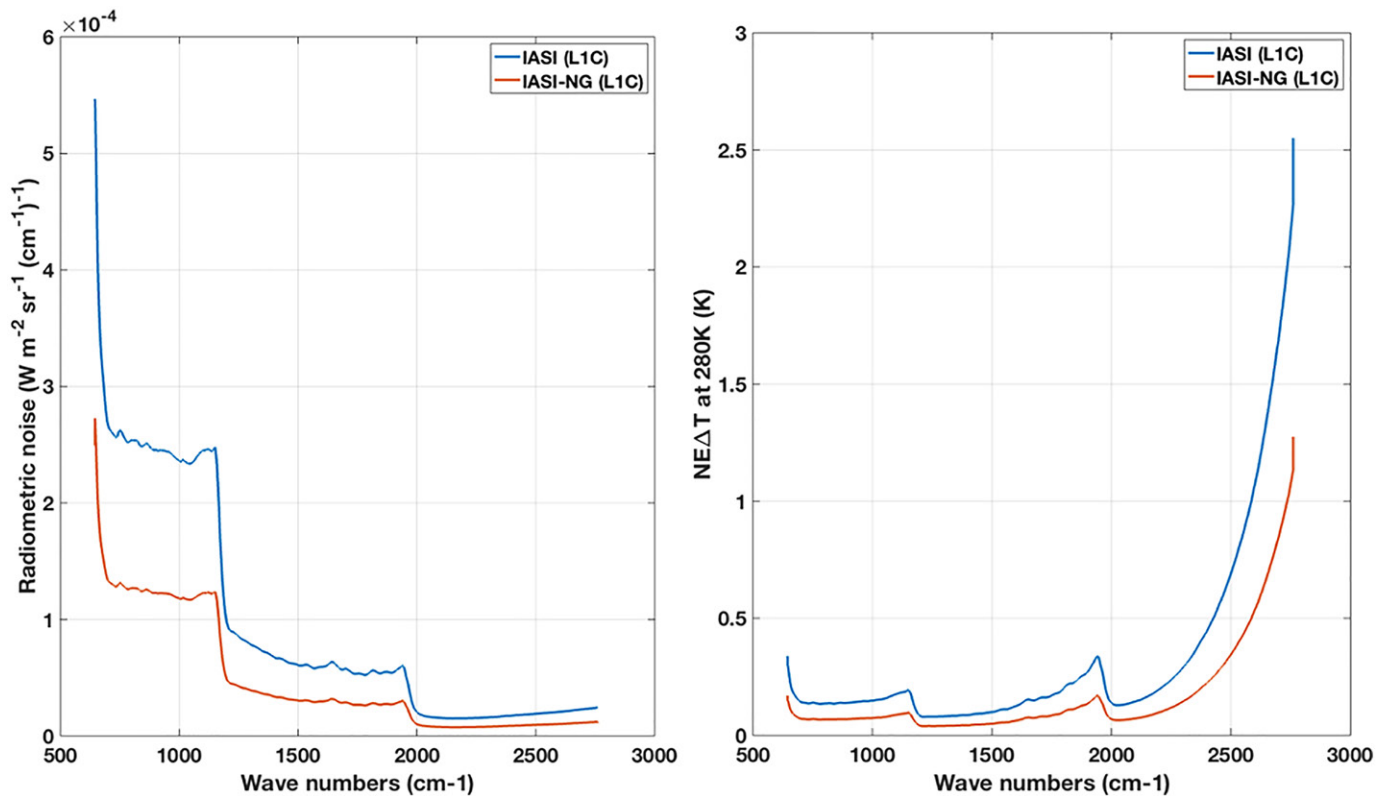| Characteristic	IASI	IASI-NG
Spectral coverage	645–2,760 cm^{-1}	645–2,760 cm^{-1}
Spectral sampling	0.250 cm^{-1}	0.125 cm^{-1}
No. of channels	8,461	16,921
Noise level	Blue curves in Fig. SB1	Red curves in Fig. SB1

Fig. SB1. Noise level for IASI (blue curves) and IASI-NG (red curves). (left) Radiometric noise in radiance units after the Gaussian apodization (L1C data). (right) Equivalent brightness temperature NEΔT at 280 K.

The MWS set of channels is thus similar to the ones of the Advanced Technology Microwave Sounder (ATMS) on board the NOAA–NASA *Suomi National Polar-Orbiting Partnership* (*SNPP*) satellite but with the additional high-frequency window channel at 229 GHz.

The Infrared Atmospheric Sounding Interferometer–New Generation (IASI-NG; www.eumetsat.int/eps-sg-iasi-ng) is an infrared interferometer which has the capability to measure the temperature and water vapor profiles of Earth's atmosphere. The primary objective of IASI-NG is to support NWP at regional and global scales. IASI-NG has a direct heritage from its predecessor IASI instrument on board EPS satellites. IASI-NG will provide observations at 16,921 spectral samples between 645 and 2,760 cm^{-1} with a spectral resolution of 0.25 cm^{-1}. With respect to its predecessor IASI, IASI-NG has double spectral resolution and signal-to-noise improved by a factor of 2 (Fig. SB1, Table SB2). These improvements are fundamental for increasing the accuracy in estimating the vertical profile of key atmospheric parameters for weather and climate studies (Crevoisier et al. 2014; Boynard et al. 2017).

machine learning approaches, two types in particular: (i) feedforward artificial neural networks (NN), also referred to as multilayer perceptron (MLP) systems and demonstrated to be "universal approximators" (Hornik et al. 1989), and (ii) random forests (RF), also referred to as ensemble learning methods, because they operate by constructing a structure composed of several decision trees (Breiman 2001). Within the ComboCloud project, NN algorithms were developed for the retrieval of CLWC, CIWC, CLWP, and CIWP, while an RF algorithm was developed for the retrieval of Re.

Both NN and RF require training of the coefficients. This is usually obtained through a dedicated dataset of collocated and simultaneous observations and target variables. Since ComboCloud mainly addresses future sensors (MWS and IASI-NG will fly on board the MetOp-SG spacecraft from 2024 onward), an observational dataset is not available yet and thus a synthetic dataset of simulated observations had to be generated. Indeed, a simulated approach provides the flexibility needed for evaluating the relative improvements brought by the MW–IR synergy and by future sensors with respect to what is currently available. Figure 3 pictures the design, implementation, and validation of ComboCloud algorithms. A simulated dataset of realistic observations was produced processing global data from ERA5, the fifth generation of European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis, with

state-of-the-art radiative transfer models (RTM) codes, namely, RTTOV-SCAT (Saunders et al. 2018) and sigma-IASI-as (Liuzzi et al. 2017). Input data from ERA5 are mainly atmospheric thermodynamic profiles (pressure, temperature, humidity, cloud, and precipitation water contents) and surface meteorological data (pressure, temperature, humidity, wind). Global data at four days (1 January, 1 April, 1 July, 1 October) and four synoptic hours (0000, 0600, 1200, 1800 UTC) have been processed to catch typical seasonal and diurnal cycles. The considered dataset contains high (33%), medium (45%), and low (30%) clouds, inclusive of overlap cases. Thus, the considered dataset is assumed to include all types of clouds with realistic relative frequencies. Histograms for other cloud characteristics (i.e., pressure level, CLWP, CIWP) are illustrated in Fig. 4.
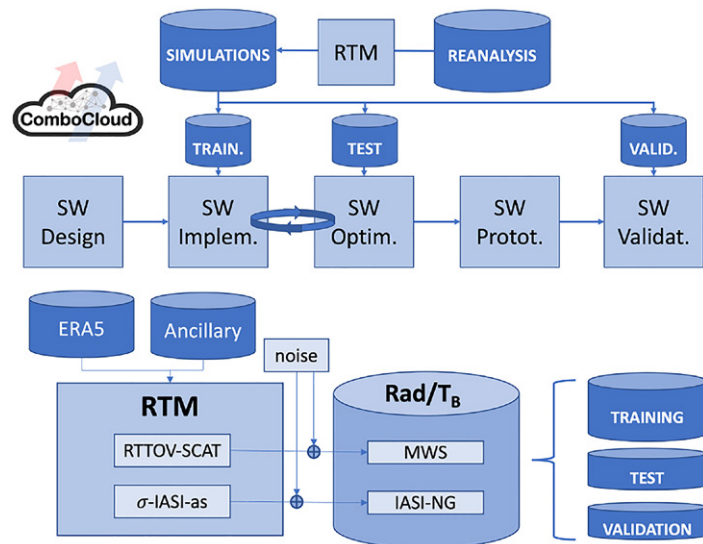

Fig. 3. (top) Flow diagram of ComboCloud algorithm design, implementation, and validation. Radiative transfer model (RTM) codes are used to generate simulated realistic observations from global reanalysis fields. (bottom) Zooming in the RTM input/output. Ancillary data include surface emissivity and cloud particle size distributions.

Radiance simulations at IR wavelengths assume cloud optical properties from the LBLDIS library for either liquid or ice particles (Turner et al. 2003), considering crystal aggregates for the latter (Martinazzo et al. 2021). Realistic values for cloud effective radius and size distributions are computed from ERA5 data using parameterized methods, developed for either liquid (Martin et al. 1994) or ice particles (Wyser 1998). At MW wavelengths, recommended RTTOV-SCATT optical properties have been used (Geer and Baordo 2014), including Mie spheres for rain and nonprecipitating hydrometeors and sector snowflakes for solid precipitation. The processed dataset thus contains a fairly large number (66,144) of simulated radiance and brightness temperature spectra at MW and IR channels corresponding to current EPS (AMSU/MHS and IASI) and future EPS-SG (MWS and IASI-NG) sensors. Typical emissivity values obtained from physical model and atlas are considered in the simulation of MW and IR radiances. MW emissivity is computed with the TESSEM2 model and TELSEM2 atlas over ocean and land, respectively, available as part of the RTTOV package. For simulation over sea background, the emissivity is modified according to wind speed to account for ocean roughness. IR surface emissivity is computed based on the Masuda model for sea surface (available for the 15 IASI Field of Regard angles), while the IR emissivity for land surface is derived from the University of Wisconsin Global Infrared land Surface Emissivity[2] (Seemann et al. 2008). However, only data simulated over ocean have been considered in this study. Three nadir angles have been considered (2°, 21°, and 44°), spanning the range of useful observing angles of most satellite sounders. Synthetic noise is added to simulations accord-

ing to the instrument requirements (Tables SB1 and SB2 in the sidebar). The obtained dataset has been divided into three subsets for algorithm training (70%), test (15%), and validation (15%). The designed NN and RF algorithms are implemented and optimized based on the training and test datasets, and finally evaluated with the validation dataset. The target reference values correspond to the cloud properties values from the global reanalysis used to generate
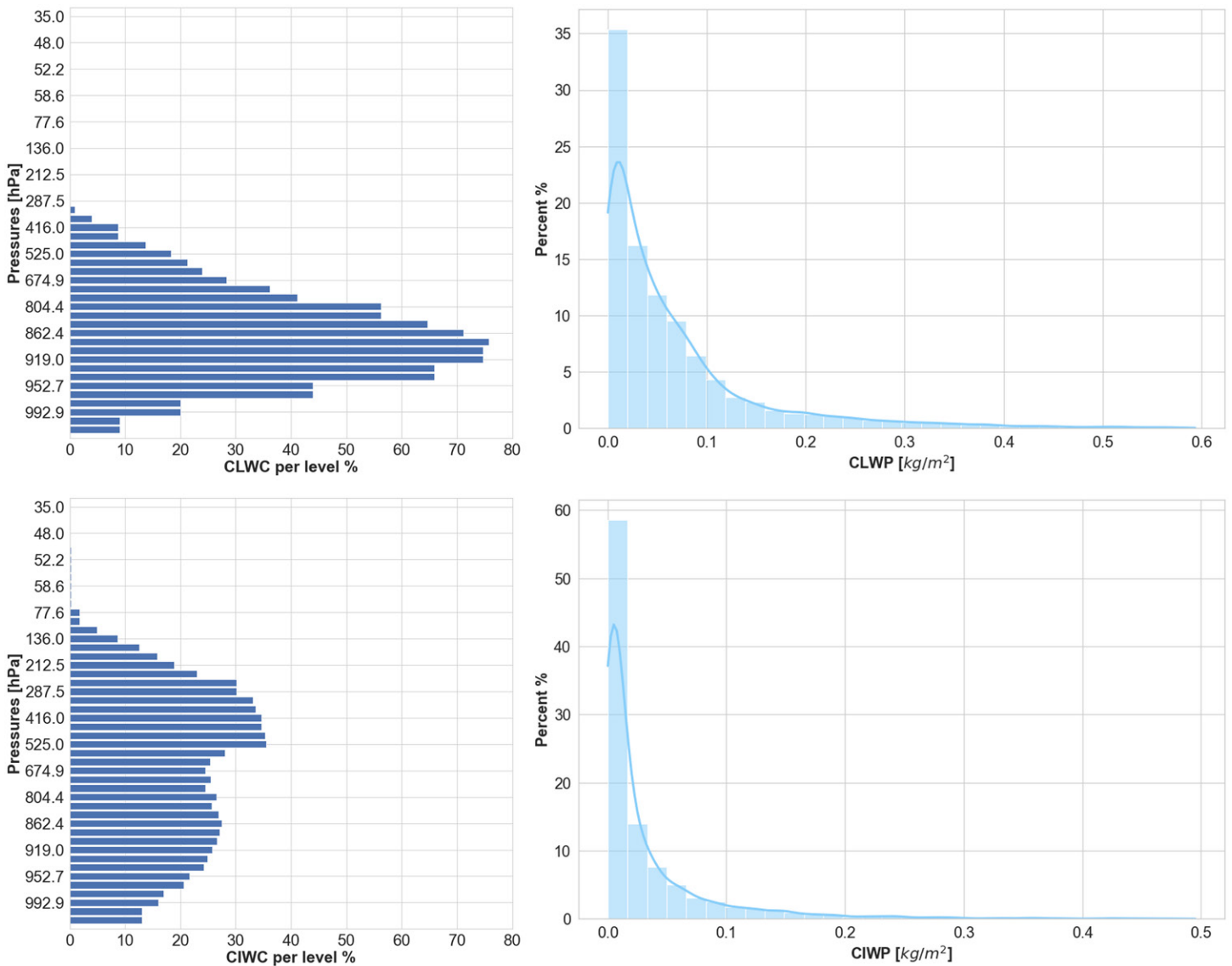
Fig. 4. (top) Histograms of (left) pressure and (right) water content for liquid clouds within the considered dataset. (bottom) As in the top panels, but for ice clouds.

the simulated observations. The algorithms were trained and optimized on the global dataset over ocean (day and night, including liquid and solid precipitation). The dataset was filtered to avoid extremely large CLWP/CIWP values (i.e., CLWP and CIWP larger than 0.6 and 0.5 kg m^{-2}, respectively) that may harm the NN training, screening out less than 0.3% of the initial dataset. Algorithm optimization was achieved by objective methods, like hyperband tuners (Li et al. 2018; Akiba et al. 2019), random search (Bergstra and Bengio 2012; Kandasamy et al. 2018), and early stopping (Caruana et al. 2000; Yao et al. 2007). Four independent modules have been developed, one for the retrieval of CLWC and CIWC profiles (SW1), one for the retrieval of CLWP and CIWP (SW2), one for the retrieval of the cloud effective radius (SW3), and one for the detection of thin cirrus (SW4). All four modules exploit machine learning techniques (NN for SW1-2-4, RF for SW3) and are designed to ingest all channels from either current (AMSU/MHS and IASI) or future (MWS and IASI-NG) sensors. SW1-2-3 are designed to take in input either MW-only, IR-only, or combined MW–IR, while SW4 is designed to take in input IR-only. Each module thus consists of multiple independent algorithms, with the same output, but different input and architecture. Note that the original dimensionality of the hyperspectral data (8,461 channels for IASI and 16,921 for IASI-NG) is reduced using principal component analysis (PCA). This step allows a compact representation of the spectra,

much more efficient for the development and operational use of retrieval algorithms. All modules are applicable at one-pixel resolution and include an estimate of the per-pixel uncertainty. Further information on the algorithm development is given in Mastro et al. (2021, 2022).

Validation

The validation data subset was used to quantify the algorithm retrieval performances, comparing retrieved cloud products obtained from simulated satellite observations (CLWC, CIWC, CLWP, CIWP, Re) with the corresponding ERA5 data from which the simulated observations were generated. The validation dataset does not include cases used during the algorithm training phase, though it is generated using the same radiative transfer code and settings. As the validation set was selected randomly from the original dataset, we assume it retains the main characteristics, i.e., it includes all kind of clouds with realistic relative frequencies. Performances of configurations based on each single instrument or their combination can be compared to quantify the benefits of future sensors and MW–IR synergy with respect to current state. Figure 5 shows the statistics of the retrieval error and also the error decrease in CLWC and CIWC retrievals using six configurations. The retrieval error decrease is defined as the percentage decrease in root-mean-square error (rmse) with respect to the mean state; the latter being defined as simply the mean CLWC (or CIWC) profile derived from the training set:

$$f_{\mathrm{rmse}} = 100 \times \frac{\mathrm{rmse}_{\mathrm{clim}} - \mathrm{rmse}_{\mathrm{retr}}}{\mathrm{rmse}_{\mathrm{clim}}}. \tag{1}$$

Among the six configurations, three are based on current sensors (MW-only: AMSU/MHS; IR-only: IASI; MW–IR synergy: AMSU/MHS and IASI) and three on future sensors (MW-only: MWS; IR-only: IASI-NG; MW–IR synergy: MWS and IASI-NG). Figure 5 confirms quantitatively the qualitative expectations: the MW–IR synergy outperforms the configurations based in either MW- or IR-only for both liquid and ice clouds throughout the vertical domain, from few percent up to ~20%. The benefits of future sensors with respect to current ones is not always evident, but for MWS, which usually provides higher retrieval error decrease than the current AMSU/MHS for both liquid and ice clouds. However, it is worth noting that the combination of new sensors leads to a substantial error decrease (~10%) for CLWC at 500–600 hPa, which may be relevant for NWP applications. Note also that NWP centers do not necessarily assimilate all channels, but just an opportune down-selection (e.g., Collard 2007), which may hide the benefits of higher resolution. However, PCA-based retrievals and data assimilation (Masiello et al. 2011; Matricardi and McNally 2013) are likely to grow with the advent of future higher-resolution sensors like IASI-NG.

Similarly, Fig. 6 shows the retrieval performances for path-integrated CLWP and CIWP. Here, we use the Taylor diagram (Taylor 2001) for a graphical, intuitive, and concise representation to compare the retrieval performances. The performances are quantified in terms of three statistics: (i) the Pearson correlation coefficient, (ii) the rmse, and (iii) the variability defined as one standard deviation (std) of the validation subset. Practically, the closer a marker is to the reference point, the better are the retrieval performances, so the relative merit of the different configurations is evident. MW–IR synergy shows the best retrieval performances for both CLWP and CIWP, followed by MWS-only and IR-only. Quantitatively, the MW–IR combination provides CLWP with 2% (14%) higher correlation and a 1.4 (2.4) factor lower rmse with respect to MW-only (IR-only). For CIWP, the MW–IR combination provides 4% (8%) higher correlation and a 1.7 (2.1) factor lower rmse with respect to MW-only (IR-only). With respect to current MetOp sensors (IASI and AMSU/MHS), the combination of future
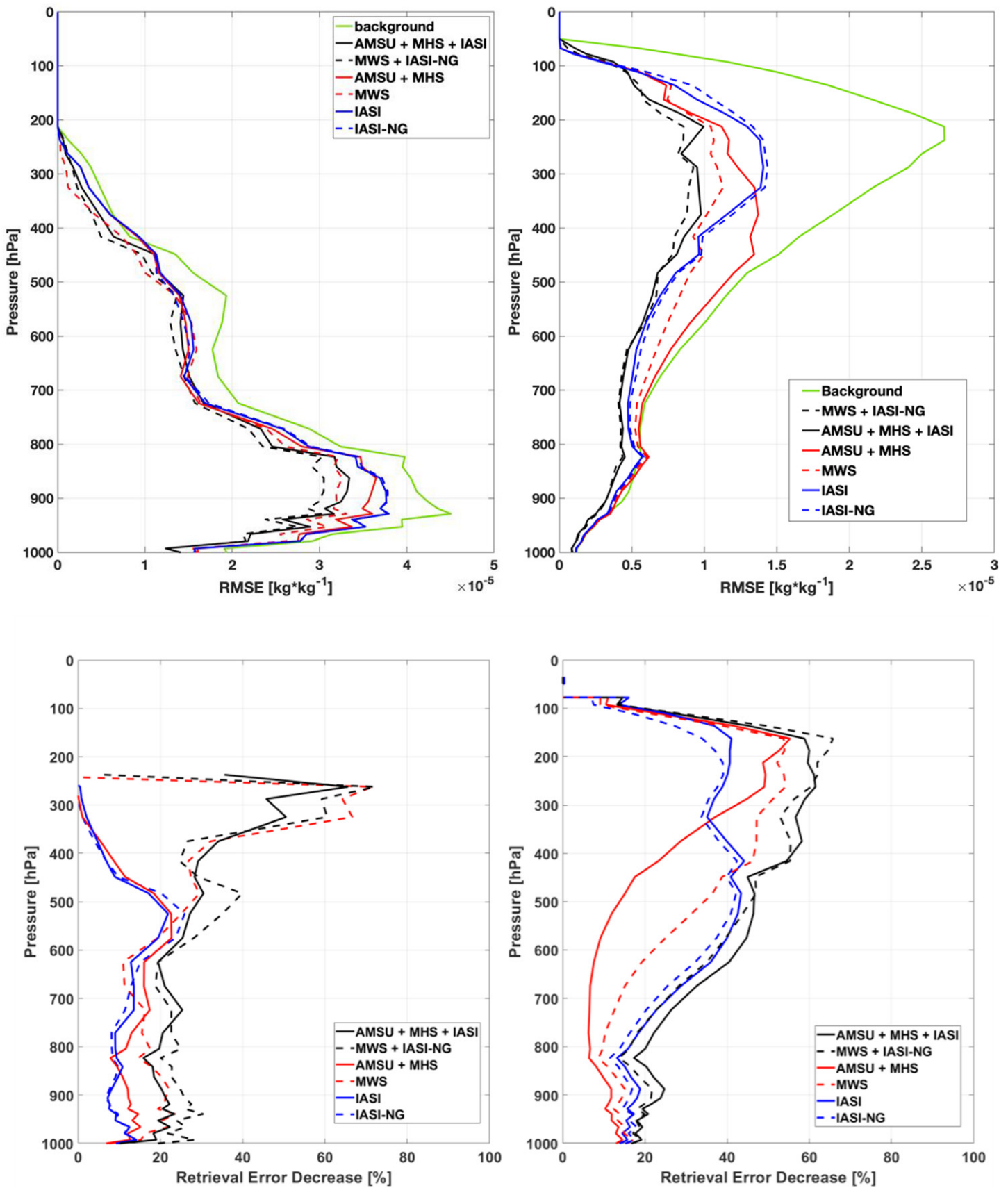
Fig. 5. (top) Retrieval error (kg kg^{-1}) for (left) cloud liquid water content (CLWC) and (right) cloud ice water content (CIWC). IR-only is shown in blue (solid: IASI; dashed: IASI-NG), MW-only in red (solid: AMSU+MHS, dashed: MWS), MW–IR combination in black (solid: AMSU+MHS+IASI; dashed: MWS+IASI-NG). (bottom) As in the top panels, but for retrieval error decrease (%) with respect to background.
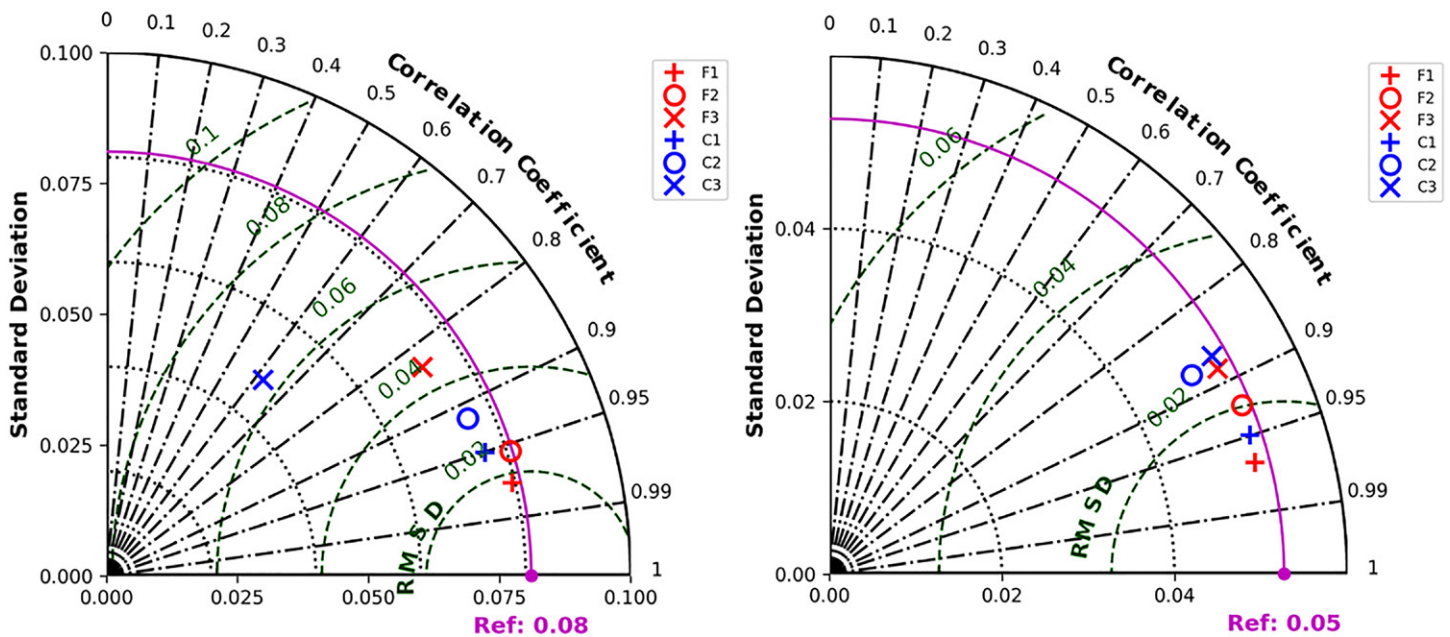
Fig. 6. Taylor diagrams for (left) CLWP and (right) CIWP retrievals. "Ref" marks the best adherence to the refence truth. Six NN architectures are compared, three corresponding to future sensors (F, red markers) and three to current sensors (C, blue markers). Plus markers (+) indicate combined MW and IR sensors (F1: MWS and IASI-NG; C1: IASI and AMSU/MHS). Circles indicate MW-only (F2: MWS; C2: AMSU/MHS-only). Cross markers (X) indicate IR-only (F3: IASI-NG-only; C3: IASI-only).

MetOp-SG sensors (IASI-NG and MWS) provides CLWP with 0.4% higher correlation and a 1.2 factor lower rmse and CIWP with 0.2% higher correlation and a 1.4 factor lower rmse.

For the retrieval of cloud drop effective radius (Re), Fig. 7 shows the performances obtained with the combined IASI-NG and MWS configuration. Re for liquid and ice clouds are retrieved separately, and so the validation is performed separately. The reference Re have been determined from ECMWF data with the same parameterized methods used in the training [Martin et al. (1994) and Wyser (1998) for liquid and ice clouds, respectively]. The scatterplots in Fig. 7 show that the MW–IR combined approach offers some skills in retrieving Re throughout the 4–100-μm range. The validation shows 0.68-μm rmse and 0.79 correlation for liquid clouds and 11.67-μm rmse and 0.86 correlation for ice clouds, with a tendency for the retrievals to overestimate small values and underestimate large values. Using IASI-NG alone, the scores reduce to 0.87-μm rmse and 0.62 correlation for liquid clouds and 12.1-μm rmse and 0.85 correlation for ice clouds. With respect to current MetOp sensors (IASI and AMSU/MHS), the combination of future MetOp-SG sensors (IASI-NG and MWS) provides Re with 1.6% higher correlation and a 2.4% lower rmse for liquid clouds and with 1.9% higher correlation and a 4.7% lower rmse for ice clouds. These results provide overall global performances, although it is likely that the accuracy of Re estimates depends on cloud characteristics (e.g., optical thickness and top temperature).

Uncertainty estimation

A single-pixel uncertainty estimate for each cloud product has been implemented within the four retrieval modules. The single-pixel uncertainty is estimated for each module through the validation process, computing a binning analysis of the residuals between the values of the output cloud product and the corresponding target. Although this uncertainty is not based on error propagation, as for example in the optimal estimation method (Maahn et al. 2020), it is based on the statistics of realistic retrieval residuals. The residuals account for the uncertainty due to the inversion technique as well as the expected instrumental random uncertainty, as instrumental noise was added to the simulated radiances. Note that the
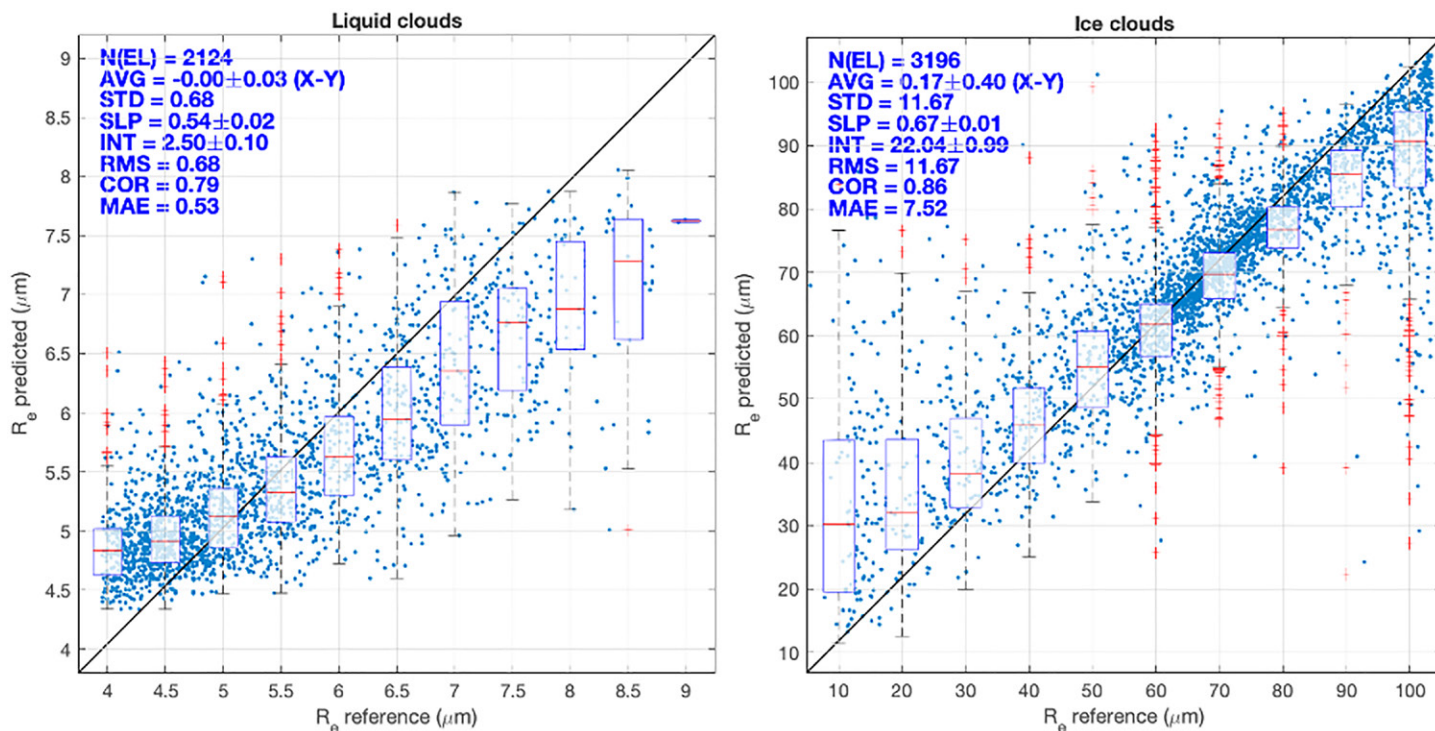
Fig. 7. Superimposed scatterplot and box-and-whisker diagrams of Re retrievals for (left) liquid and (right) ice clouds. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles (Q1 and Q3, respectively), the whiskers extend to the most extreme datapoints that are not considered outliers, while the outliers are plotted individually with red cross markers. Points are considered as outliers if they are outside 99.3% coverage of a normal distribution.

residuals are not spatially correlated, as they are computed on single observations that are independent from each from other. Figure 7 shows the results from the binned analysis for selected cloud products (CLWC, CIWC, CLWP, CIWP, Re for liquid and ice clouds). Both the systematic and random error components are estimated, although they are combined into the total error to estimate the retrieval uncertainty (u_t). A polynomial curve is fitted through the total error resulting from the binned analysis, and the single-pixel uncertainty estimate is generated evaluating the fit of the error model as a function of the retrieved product value:

$$u_t\left(P_i\right)=\text{fit}\left(\varepsilon_M,P_i\right). \tag{2}$$

For example, the ith-pixel uncertainty associated to CLWC is computed as

$$u_t\left(\text{CLWC}_i\right)=\text{fit}\left(\text{rms},\text{CLWC}_i\right). \tag{3}$$

Figure 8 shows that retrieval uncertainty is ~120% for CLWC, 70%–100% for CIWC, ~20% for CLWP, 15%–30% for CIWP, 10%–15% for liquid-cloud Re, and finally 10%–75% for ice-cloud Re. The achieved retrieval uncertainty may be compared with the user requirements for each of the considered cloud products. Table 1 reports uncertainty values for global NWP applications, as retrieved from World Meteorological Organization (WMO) Observing Systems Capability Analysis and Review Tool (OSCAR).[3] Three levels of performances are specified: threshold (T), breakthrough (B), and objective (O). Table 1 also reports the values set by the EPS-SG End User Requirements Document (EURD 2019), limited to the cloud products considered therein. Finally, Table 1 reports the performance levels obtained in ComboCloud: the goal (G), as set initially based on a literature review

[3] www.wmo-sat.info/oscar/requirements

Table 1. Cloud product accuracy requirements for global NWP applications. Threshold (T), breakthrough (B), and objective (O) levels are from WMO OSCAR. EURD levels are from the EPS-SG End User Requirements Document (EPS-SG EURD gives two accuracy values for CLWP: 5% using MWS and IASI-NG and 50% using MWI and ICI). The goal (G) level was set initially based on the ComboCloud literature survey. The achieved (A) values report results from the ComboCloud validation from the simulated dataset.

Product	Units	WMO OSCAR Global NWP			EUMETSAT	ComboCloud	
		T	B	O	EURD	G (initial)	A (final)
CLWC	g kg^{-1}	100%	50%	20%		100%	120%
CIWC	g kg^{-1}	100%	50%	20%		100%	70%
CLWP	g m^{-2}	50 g m^{-2}	20 g m^{-2}	10 g m^{-2}	5%–50%	50 g m^{-2}	21 g m^{-2}/20%
CIWP	g m^{-2}	20 g m^{-2}	10 g m^{-2}	5 g m^{-2}	50%	20 g m^{-2}	17 g m^{-2}/20%
Re	μm	5 μm	2 μm	1 μm	5 μm	4 μm	1 μm (liq)
							12 μm (ice)

(e.g., Lin et al. 1998; Huang et al. 2006; Romano et al. 2007; Holl et al. 2014; Nalli et al. 2018), and the achieved (A) performances resulting from the validation using the simulated dataset. Note that the estimated per-pixel uncertainty does not account for the additional uncertainty related to mapping the products into model space, i.e., from satellite observations (level 2 data) to regular grids (level 3 data) (Sun et al. 2018).

Summary and conclusions

The ComboCloud study demonstrated quantitatively the value of combining MW and IR observations to retrieve cloud properties with respect to using MW-only or IR-only. The synergistic retrievals were demonstrated to combine the sensitivity of MW and IR to cloud properties, thus outperforming the performances achievable by either one. In addition, the study demonstrated quantitatively the added value of next generation EPS-SG sensors (IASI-NG and MWS) with respect to the EPS sensors currently available (IASI and AMSU/MHS). The impact of sensor synergy and next generation sensors depends on the considered cloud product and can lead to considerable improvements (up to 20%).

The expected performances for the retrieval of cloud microphysical properties have been compared against current WMO and user requirements. Although the MW–IR synergy outperforms MW- and IR-only methods, the achieved performances for liquid cloud profiling (CLWC, 120%) are ~20% worse than the minimum threshold (100%) level set by WMO for global NWP applications (Table 1). Conversely, the achieved performances for ice cloud profiling (CIWC, 70%) are below the threshold (100%) and closer to the breakthrough (50%) level. Similarly, the performances for the integrated liquid and ice paths (CLWP and CIWP, respectively 21 and 17 g m^{-2}) are between the threshold (respectively 50 and 20 g m^{-2}) and breakthrough (respectively 20 and 10 g m^{-2}) levels. The performances for CLWP and CIWP correspond to ~20% of the corresponding total range. For CIWP, this is well within the EPS-SG End User Requirements (EURD 2019). For CLWP, the achieved performances are better than what was expected from the combined two EPS-SG imagers (MWI and ICI, 50%), but worse than what was expected from the combined MWS and IASI-NG (5%). Our analysis seems to indicate that the EURD requirement for CLWP from the combined MWS and IASI-NG is quite optimistic. Finally, for the retrieval of effective radius Re, the achieved performances meet both the EURD and the WMO objective level when limited to liquid clouds. For ice clouds, the achieved performances are a factor of ~2 larger than the requirements. But given the much larger range of ice particle size, the achieved performances result in less than 12%. Note that no distinction is made in EURD and WMO OSCAR between Re of liquid and ice clouds.
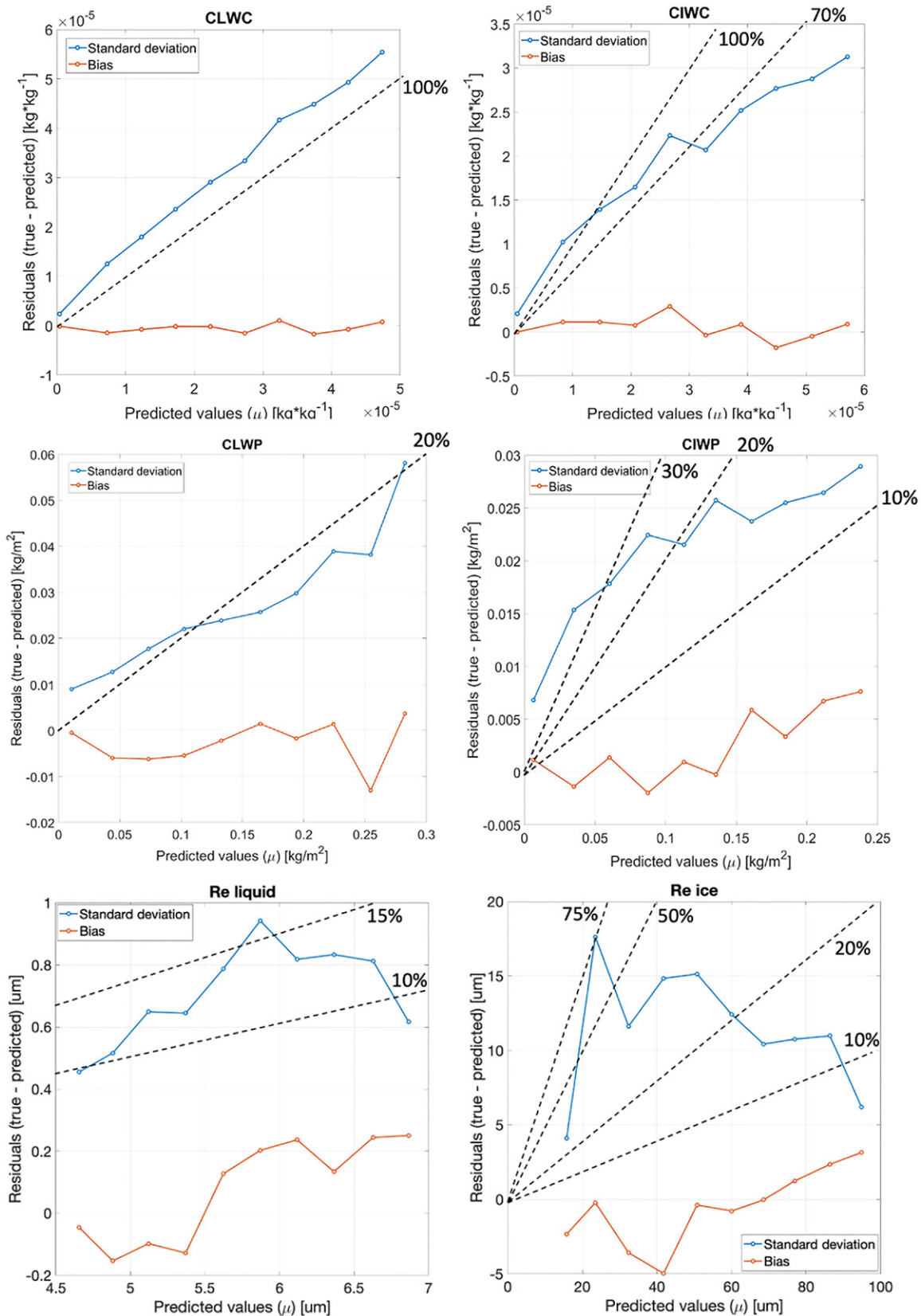
Fig. 8. Retrieval uncertainty estimates by binned analysis for the configurations exploiting combined MW and IR observations from future sensors (MWR and IASI-NG). Systematic (bias) and random (std) error components are estimated: (top left) CLWC, (top right) CIWC, (middle left) CLWP, and (middle right) CIWP, (bottom left) liquid-cloud Re, and (bottom right) ice-cloud Re.

This analysis shows that the improvements expected by future sensors depend on the considered cloud product. In general, the combination of future sensors provides performance closer to the WMO requirements for the retrieval of cloud microphysical properties. The scientific impacts of these improvements include better understanding of cloud microphysics, refined cloud climatology, and radiative forcing. All these contribute to reducing the large uncertainty associated with cloud essential climate variables in climate predictions and will also provide more accurate reference for the assessment of climate models. Few directions are identified for further refining the current analysis:

(i) Consider additional nadir angles in the training, test, and validation datasets, possibly covering the whole set of operational IASI/IASI-NG nadir angles.

(ii) Consider refining the sampling procedure of the training dataset in order to extend the range and thus the ability to estimate products in extreme events.

(iii) After the launch, consider algorithm output validation and calibration against reference estimate of cloud parameters; identified references for validation are observations from active sensors deployed both on spaceborne platforms (e.g., *CloudSat*) and ground-based networks (e.g., Cloudnet). In addition, cross-comparison is foreseen with cloud products from other sensors exploiting also visible channels, such as the low-orbiting VIIRS and the geostationary Flexible Combined Imager (FCI).

(iv) After the launch, investigate the feasibility to perform training, test, and validation with a dataset of real observations, in order to mitigate effects related to inherent inconsistencies between radiance observations and radiative transfer simulations.

In summary, the ComboCloud project set the stage for the exploitation of artificial intelligence methods using MW–IR synergy for cloud microphysical product retrievals, paving the road for further developments and the operational exploitation with EPS-SG MWS and IASI-NG from 2024 onward.

Data availability statement. Data analyzed in this study were a postprocessing of existing data (ERA5), which are openly available at locations cited in the reference section. The radiative transfer simulations were obtained using RTTOV v12.3 (https://nwp-saf.eumetsat.int/site/download/documentation/rtm/docs_rttov12/NWPSAF_ReleaseNote_RTTOV12.3.pdf) and sigma-IASI-AS (version as in reference Liuzzi et al. 2017). All the information needed to replicate the simulations is provided. The output radiative transfer simulations are available from Domenico Cimini (domenico.cimini@imaa.cnr.it).

References

Akiba, T., S. Sano, T. Yanase, T. Ohta, and M. Koyama, 2019: Optuna: A next-generation hyperparameter optimization framework. *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Association for Computing Machinery, 2623–2631, https://doi.org/10.1145/3292500.3330701.

August, T., and Coauthors, 2012: IASI on Metop-A: Operational level 2 retrievals after five years in orbit. *J. Quant. Spectrosc. Radiat. Transfer*, **113**, 1340–1371, https://doi.org/10.1016/j.jqsrt.2012.02.028.

Belgiu, M., and L. Drăguţ, 2016: Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.*, **114**, 24–31, https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Bergstra, J., and Y. Bengio, 2012: Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, **13**, 281–305.

Bony, S., and Coauthors, 2015: Clouds, circulation and climate sensitivity. *Nat. Geosci.*, **8**, 261–268, https://doi.org/10.1038/ngeo2398.

Boynard A., C. Clerbaux, S. Bauduin, P. Prunet, B. Tournier, J. Hadji-Lazaro, D. Hurtmans, and P.-F. Coheur, 2017: Air quality monitoring with current (IASI) and future (IASI-NG/MetOp-SG, IRS/MTG) space-borne thermal infrared sounders. *Geophysical Research Abstracts*, Vol. 19, Abstract 16047, https://meetingorganizer.copernicus.org/EGU2017/EGU2017-16047.pdf.

Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5–32, https://doi.org/10.1023/A:1010933404324.

Caruana, R., S. Lawrence, and C. Giles, 2000: Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. *Proc. Advances in Neural Information Processing Systems 13 (NIPS 2000)*, Denver, CO, Conference on Neural Information Processing Systems, 402–408.

Chevallier, F., P. Bauer, J.-F. Mahfouf, and E. Morcrette, 2002: Variational retrieval of cloud profile from ATOVS observations. *Quart. J. Roy. Meteor. Soc.*, **129**, 2511–2525, https://doi.org/10.1256/qj.01.153.

Collard, A. D., 2007: Selection of IASI channels for use in numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **133**, 1977–1991, https://doi.org/10.1002/qj.178.

Crevoisier, C., and Coauthors, 2014: Towards IASI-New Generation (IASI-NG): Impact of improved spectral resolution and radiometric noise on the retrieval of thermodynamic, chemistry and climate variables. *Atmos. Meas. Tech.*, **7**, 4367–4385, https://doi.org/10.5194/amt-7-4367-2014.

EURD, 2019: EPS-SG End User Requirements Document (EURD). EUM/PEPS/REQ/09/0151 v5, 168 pp., accessed 18 February 2020.

Fu, D., X. Guo, and C. Liu, 2011: Effects of cloud microphysics on monsoon convective system and its formation environments over the South China Sea: A two-dimensional cloud-resolving modeling study. *J. Geophys. Res.*, **116**, D07108, https://doi.org/10.1029/2010JD014662.

Geer, A. J., and F. Baordo, 2014: Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech*., **7**, 1839–1860, https://doi.org/10.5194/amt-7-1839-2014.

——, N. Bormann, K. Lonitz, P. Weston, R. Forbes, and S. English, 2019: Recent progress in all-sky radiance assimilation. *ECMWF Newsletter*, No. 161, ECMWF, Reading, United Kingdom, 8 pp., www.ecmwf.int/en/newsletter/161/meteorology/recent-progress-all-sky-radiance-assimilation.

Holl, G., S. Eliasson, J. Mendrok, and S. A. Buehler, 2014: SPARE-ICE: Synergistic ice water path from passive operational sensors. *J. Geophys. Res. Atmos.*, **119**, 1504–1523, https://doi.org/10.1002/2013JD020759.

Holloway, J., and K. Mengersen, 2018: Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sens.*, **10**, 1365, https://doi.org/10.3390/rs10091365.

Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feedforward networks are universal approximators. *Neural Network*, **2**, 359–366, https://doi.org/10.1016/0893-6080(89)90020-8.

Huang, J., P. Minnis, B. Lin, Y. Yi, T.-F. Fan, S. Sun-Mack, and S. K. Ayers, 2006: Determination of ice water path in ice-over-water cloud systems using combined MODIS and AMSR-E measurements. *Geophys. Res. Lett.*, **33**, L21801, https://doi.org/10.1029/2006GL027038.

Islam, T., P. K. Srivastava, Q. Dai, and M. Gupta, 2014: Ice cloud detection from AMSU-A, MHS, and HIRS satellite instruments inferred by cloud profiling radar. *Remote Sens. Lett.*, **5**, 1012–1021, https://doi.org/10.1080/2150704X.2014.990643.

Izquierdo-Verdiguiera, E., and R. Zurita-Millab, 2020: An evaluation of guided regularization random forest for classification and regression tasks in remote sensing. *Int. J. Appl. Earth Obs. Geoinf.*, **88**, 102051, https://doi.org/10.1016/j.jag.2020.102051.

Kandasamy, K., W. Neiswanger, J. Schneider, B. Póczos, and E. P. Xing, 2018: Neural architecture search with Bayesian optimisation and optimal transport. *Proc. 32nd Int. Conf. on Neural Information Processing Systems (NIPS'18)*, Montreal, QC, Canada, Conference on Neural Information Processing Systems, 2020–2029, https://dl.acm.org/doi/10.5555/3326943.3327130.

Kniffka, A., M. Stengel, M. Lockhoff, R. Bennartz, and R. Hollmann, 2014: Characteristics of cloud liquid water path from SEVIRI onboard the Meteosat Second Generation 2 satellite for several cloud types. *Atmos. Meas. Tech.*, **7**, 887–905, https://doi.org/10.5194/amt-7-887-2014.

Li, L., and Coauthors, 2018: Hyperband: A novel bandit-based approach to hyperparameter optimization. *J. Mach. Learn. Res.*, **18**, 1–52.

Lin, B., P. Minnis, B. Wielicki, D. R. Doelling, R. Palikonda, D. F. Young, and T. Uttal, 1998: Estimation of water cloud properties from satellite microwave, infrared and visible measurements in oceanic environments: 2. Results. *J. Geophys. Res.*, **103**, 3887–3905, https://doi.org/10.1029/97JD02817.

Liuzzi, G., G. Masiello, C. Serio, D. Meloni, C. Di Biagio, and P. Formenti, 2017: Consistency of dimensional distributions and refractive indices of desert dust measured over Lampedusa with IASI radiances. *Atmos. Meas. Tech.*, **10**, 599–615, https://doi.org/10.5194/amt-10-599-2017.

Maahn, M., D. D. Turner, U. Löhnert, D. J. Posselt, K. Ebell, G. G. Mace, and J. M. Comstock, 2020: Optimal estimation retrievals and their uncertainties: What every atmospheric scientist should know. *Bull. Amer. Meteor. Soc.*, **101**, E1512–E1523, https://doi.org/10.1175/BAMS-D-19-0027.1.

Maddy, E. S., and S. A. Boukabara, 2021: MIIDAPS-AI: An explainable machine-learning algorithm for infrared and microwave remote sensing and data assimilation preprocessing – Application to LEO and GEO sensors. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **14**, 8566–8576, https://doi.org/10.1109/JSTARS.2021.3104389.

Marke, T., K. Ebell, U. Löhnert, and D. D. Turner, 2016: Statistical retrieval of thin liquid cloud microphysical properties using ground-based infrared and microwave observations. *J. Geophys. Res. Atmos.*, **121**, 14558–14573, https://doi.org/10.1002/2016JD025667.

Martin, G. M., D. W. Johnson, and A. Spice, 1994: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *J. Atmos. Sci.*, **51**, 1823–1842, https://doi.org/10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2.

Martinazzo, M., D. Magurno, W. Cossich, C. Serio, G. Masiello, and T. Maestri, 2021: Assessment of the accuracy of scaling methods for radiance simulations at far and mid infrared wavelengths. *J. Quant. Spectrosc. Radiat. Transfer*, **271**, 107739, https://doi.org/10.1016/j.jqsrt.2021.107739.

Masiello, G., C. Serio, and P. Antonelli, 2011: Inversion for atmospheric thermodynamical parameters of IASI data in the principal components space. *Quart. J. Roy. Meteor. Soc.*, **138**, 103–117, https://doi.org/10.1002/qj.909.

Mastro, P., G. Masiello, D. Cimini, F. Romano, E. Ricciardelli, F. Di Paola, T. Hultberg, and T. August, 2021: Cloud liquid and ice water content estimation from satellite: A regression approach based on neural networks. *Proc. SPIE*, **11859**, 118590H, https://doi.org/10.1117/12.2600124.

——, and Coauthors, 2022: Combined IASI-NG and MWS observations for the retrieval of cloud liquid and ice water path: A deep learning artificial intelligence approach. *IEEE J. Select. Top. Appl. Earth Obs. Remote Sens.*, **15**, 3313–3322, https://doi.org/10.1109/JSTARS.2022.3166992.

Matricardi, M., and A. P. McNally, 2013: The direct assimilation of principal components of IASI spectra in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **140**, 573–582, https://doi.org/10.1002/qj.2156.

Nalli, N. R., and Coauthors, 2018: Validation of atmospheric profile retrievals from the SNPP NOAA-unique combined atmospheric processing system. Part 1: Temperature and moisture. *IEEE Trans. Geosci. Remote Sens.*, **56**, 180–190, https://doi.org/10.1109/TGRS.2017.2744558.

Romano, F., D. Cimini, R. Rizzi, and V. Cuomo, 2007: Multilayered cloud parameters retrievals from combined infrared and microwave satellite observations. *J. Geophys. Res.*, **112**, D08210, https://doi.org/10.1029/2006JD007745.

Saunders, R., and Coauthors, 2018: An update on the RTTOV fast radiative transfer model (currently at version 12). *Geosci. Model Dev.*, **11**, 2717–2737, https://doi.org/10.5194/gmd-11-2717-2018.

Seemann, S. W., E. E. Borbas, R. O. Knuteson, G. R. Stephenson, and H.-L. Huang, 2008: Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements. *J. Appl. Meteor. Climatol.*, **47**, 108–123, https://doi.org/10.1175/2007JAMC1590.1.

Smith, N., and C. D. Barnet, 2019: Uncertainty characterization and propagation in the Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS). *Remote Sens.*, **11**, 1227, https://doi.org/10.3390/rs11101227.

Stephens, G. L., and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation from space: A review. *J. Atmos. Sci.*, **64**, 3742–3765, https://doi.org/10.1175/2006JAS2375.1.

——, D. O'Brien, P. J. Webster, P. Pilewski, S. Kato, and J. Li, 2015: The albedo of Earth. *Rev. Geophys.*, **53**, 141–163, https://doi.org/10.1002/2014RG000449.

——, D. Winker, J. Pelon, C. Trepte, D. Vane, C. Yuhas, T. L'Ecuyer, and M. Lebsock, 2018: CloudSat and CALIPSO within the A-Train: Ten years of actively observing the Earth system. *Bull. Amer. Meteor. Soc.*, **99**, 569–581, https://doi.org/10.1175/BAMS-D-16-0324.1.

Stevens, B., and S. Bony, 2013: Water in the atmosphere. *Phys. Today*, **66**, 29, https://doi.org/10.1063/PT.3.2009.

Sun, K., and Coauthors, 2018: A physics-based approach to oversample multi-satellite, multispecies observations to a common grid. *Atmos. Meas. Tech.*, **11**, 6679–6701, https://doi.org/10.5194/amt-11-6679-2018.

Sun, N., and F. Weng, 2012: Retrieval of ice water path from Special Sensor Microwave Imager/Sounder (SSMIS). *J. Appl. Meteor. Climatol.*, **51**, 366–379, https://doi.org/10.1175/JAMC-D-11-021.1.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192, https://doi.org/10.1029/2000JD900719.

Turner, D. D., 2007: Improved ground-based liquid water path retrievals using a combined infrared and microwave approach. *J. Geophys. Res.*, **112**, D15204, https://doi.org/10.1029/2007JD008530.

——, S. A. Ackerman, B. A. Baum, H. E. Revercomb, and P. Yang, 2003: Cloud phase determination using ground-based AERI observations at SHEBA. *J. Appl. Meteor.*, **42**, 701–715, https://doi.org/10.1175/1520-0450(2003)042<0701:CPDUGA>2.0.CO;2.

Wyser, K., 1998: The effective radius in ice clouds. *J. Climate*, **11**, 1793–1802, https://doi.org/10.1175/1520-0442(1998)011<1793:TERIIC>2.0.CO;2.

Yao, Y., L. Rosasco, and A. Caponnetto, 2007: On early stopping in gradient descent learning. *Constr. Approx.*, **26**, 289–315, https://doi.org/10.1007/s00365-006-0663-2.