A Feedforward Neural Network Approach for the Detection of Optically Thin Cirrus From IASI-NG

Elisabetta Ricciardelli, Francesco Di Paola, Domenico Cimini, Salvatore Larosa, Pietro Mastro, *Graduate Student Member, IEEE*, Guido Masiello, Carmine Serio, Tim Hultberg, Thomas August, and Filomena Romano

***Abstract*— The identification of optically thin cirrus is crucial for their accurate parameterization in climate and Earth's system models. This study exploits the characteristics of the infrared atmospheric sounding interferometer—new generation (IASI-NG) to develop an algorithm for the detection of optically thin cirrus. IASI-NG has been designed for the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) polar system second-generation program to continue the service of its predecessor IASI from 2024 onward. A thin-cirrus detection algorithm (TCDA) is presented here, as developed for IASI-NG, but also in parallel for IASI to evaluate its performance on currently available real observations. TCDA uses a feedforward neural network (NN) approach to detect thin cirrus eventually misidentified as clear sky by a previously applied cloud detection algorithm. TCDA also estimates the uncertainty of "clear-sky" or "thin-cirrus" detection. NN is trained and tested on a dataset of IASI-NG (or IASI) simulations obtained by processing ECMWF 5-generation reanalysis (ERA5) data with the σ-IASI radiative transfer model. TCDA validation against an independent simulated dataset provides a quantitative statistical assessment of the improvements brought by IASI-NG with respect to IASI. In fact, IASI-NG TCDA outperforms IASI TCDA by 3% in probability of detection (POD), 1% in bias, and 2% in accuracy, and the false alarm ratio (FAR) passes from 0.02 to 0.01. Moreover, IASI TCDA validation against state-of-the-art cloud products from Cloudsat/CPR and CALIPSO/Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) real observations reveals a tendency for IASI TCDA to underestimate the presence of thin cirrus (POD = 0.47) but with a low FAR (0.07), which drops to 0.0 for very thin cirrus.**

Manuscript received 27 November 2022; revised 20 April 2023 and 3 July 2023; accepted 26 July 2023. Date of publication 7 August 2023; date of current version 21 August 2023. This work was supported by EUMETSAT (www.eumetsat.int/combocloud), under the framework of the ComboCloud (Combined MWS and IASI-NG Soundings for Cloud Properties) Project, under Contract EUM/CO/19/4600002352/THH. *(Elisabetta Ricciardelli and Francesco Di Paola contributed equally to this work.) (Corresponding author: Francesco Di Paola.)*

Elisabetta Ricciardelli, Francesco Di Paola, Domenico Cimini, Salvatore Larosa, and Filomena Romano are with the Institute of Methodologies for Environmental Analysis, Italian National Research Council (CNR), 85050 Tito Scalo, Italy (e-mail elisabetta.ricciardelli@cnr.it; francesco.dipaola@cnr.it; domenico.cimini@cnr.it; salvatore.larosa@cnr.it; filomena.romano@cnr.it).

Pietro Mastro is with the Institute for the Electromagnetic Sensing of the Environment, Italian National Research Council (CNR), 80124, Naples, Italy (e-mail: p.mastro@irea.cnr.it).

Guido Masiello and Carmine Serio are with the School of Engineering, University of Basilicata, 85100 Potenza, Italy (e-mail: guido.masiello@unibas.it; carmine.serio@unibas.it).

Tim Hultberg and Thomas August are with the European Organization for the Exploitation of Meteorological Satellites, 64297 Darmstadt, Germany (e-mail: tim.hultberg@eumetsat.int; thomas.august@eumetsat.int).

Digital Object Identifier 10.1109/TGRS.2023.3303268

***Index Terms*— Feedforward neural network (NN), next-generation hyperspectral infrared (IR) data, optically thin-cirrus detection, thin-cirrus-detection error.**

NOMENCLATURE

2C-ICE	Cloudsat and CALIPSO Ice Cloud Property Product.
AVHRR	Advanced Very-High-Resolution Radiometer.
C3S	Copernicus Climate Change Service.
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization.
CBH	Cloud base height.
CCF	Cloud cover fraction.
CCL	Cloud cover layer.
CER	Cloud effective radius.
CF	Cloud fraction.
CI	Cloud information.
CIWC	Cloud ice water content.
CL	Cloud layer.
CNES	Centre National d'Études Spatiales.
ComboCloud	Combined MWS and IASI-NG soundings for cloud properties.
COPs	Cloud optical properties.
COT	Cloud optical thickness.
CPR	Cloud profiling radar.
CPRinIASI	CPR FOV included in IASI IFOV.
CT	Cloud type.
CTP	Cloud top pressure.
ECMWF	European Center for Medium-Range Weather Forecasts.
EDR	Environmental data record.
EFOV	Elementary fields of view.
EOS	NASA Earth Observing System.
ERA5	ECMWF 5 generation reanalysis.
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites.
FAR	False alarm ratio.
FOR	Field of regard.
FOV	Field of view.
GMT	Greenwich Mean Time.
GOES	Geostationary Operational Environmental Satellite.
HCC	High cloud cover.

IASI	Infrared atmospheric sounder interferometer.
IASI-NG	Infrared atmospheric sounder interferometer—new generation.
IFOV	Instantaneous field of view.
IFS	Integrated forecast system.
IR	Infrared.
LB	Layer base.
LCC	Low cloud cover.
M∗D35	Both MOD35 and MYD35 products.
MCC	Medium cloud cover.
MetOp	Meteorological Operational Satellite.
MetOp-SG	Meteorological Operational Satellite—Second Generation.
MOD35	Cloud Mask Level 2 MODIS on Terra satellite.
MODIS	Moderate Resolution Imaging Spectroradiometer.
MODinIASI	MODIS FOV included in IASI IFOV.
MSG	Meteosat Second Generation.
MWS	Microwave sounder.
MYD35	Cloud Mask Level 2 MODIS on Aqua satellite.
NN	Neural network.
NOAA	National Oceanic and Atmospheric Administration.
OCA	Optimal cloud analysis.
PC	Principal component.
PCA	Principal component analysis.
PL	Pressure levels.
POD	Probability of detection.
PPC	Perform parallax correction.
PWV	Precipitable water vapor.
RGB	Red green blue.
RMSE	Root-mean-square error.
SCIW	Specific cloud ice water content.
SCLW	Specific cloud liquid water content.
SEVIRI	Spinning Enhanced Visible and Infrared Imager.
SEVIRIinIASI	SEVIRI FOV included in IASI IFOV.
SSP	Sub satellite point.
NPP	National Polar-Orbiting Partnership.
TCC	Total cloud cover.
TCDA	Thin-cirrus detection algorithm
TCIW	Total column ice water content.
TCLW	Total column liquid water content.
TP	Total precipitation.
VIIRS	Visible Infrared Imaging Radiometer Suite.
VIIRSinIASI	VIIRS FOV included in IASI IFOV.

I. Introduction

THE accurate identification of thin cirrus is crucial for climatological studies and for the study of Earth radiation budget [1]. In the upper troposphere, they cause both atmospheric cooling, by reflecting back the incoming shortwave solar radiation, and atmospheric heating, by partially trapping outgoing longwave terrestrial radiation [2]. The dominant process depends on the cloud properties. Despite their importance for the radiation budget, their detection is challenging due to relatively low contrast with the underlying surface, and the misidentification of thin cirrus as clear sky introduces errors in the retrieval of atmospheric and surface parameters [3]. Several studies focus on the detection of thin cirrus alone or as a distinct class in a cloud classification scheme using physical methods based on threshold tests applied to IR and/or Visible (VIS) observations [4], [5]. A classical physical method, very effective for the characterization of optically thin cirrus, is the CO_2 slicing method [6]. It is based on IR observations from 13 to 15 μm to estimate CTP at high altitudes and to distinguish semitransparent from opaque clouds. McHardy et al. [7] also exploited PWV information in detecting thin cirrus by using physical-based algorithms.

Among the studies based on statistical methods, Bankert [8] applied a probabilistic NN to the AVHRR data to assign a defined sample area to one of the ten considered cloud classes. The method identified about 75% of the examined samples correctly, with a high rate of misclassification occurring between high thin cirrus and cirrostratus. In a later study, Bankert et al. [9] applied a one-nearest-neighbor classification to the GOES observations using a training dataset of expertly labeled image samples. In this way, they improved the choice of samples to be included in the various classes training dataset with particular attention to thin cirrus. Strandgren et al. [10] used an NN approach trained with data from SEVIRI and CALIOP. Maestri et al. [11] proposed a machine learning algorithm for cloud detection and classification, using simulated high spectral resolution radiances. They demonstrated that very thin cirrus are better detected exploiting the full IR spectrum rather than the mid-IR part alone [12], as it could be expected from previous works [13], [14], showing numerical simulations in the thermal IR to be sensitive to variations in cirrus optical depth and ice crystal size as well as in ice crystal shape. It is evident that, compared to other passive remote sensing instruments, high spectral resolution IR sounders carry more information on cirrus properties. However, detecting thin-cirrus clouds remains an unsolved challenge due to their spectral signatures being very similar to those of clear sky. Consequently, cloud classification algorithms often misclassify thin-cirrus clouds as clear sky [15], [16], [17], with nonnegligible consequences on the estimation of atmospheric parameter whose correct knowledge is fundamental for the Earth's radiation budget. To overcome the frequent misidentification of thin cirrus in clear sky, it is possible to exploit the improvements in the spatial and spectral resolutions of future satellite-borne sensors that are promising for detecting clouds and estimating their properties with greater precision. In this study, based on previous considerations, the whole IR spectrum [645, 2760] cm^{-1} of the IASI-NG was considered for thin and very thin-cirrus detection using a feedforward NN: the TCDA. In detail, TCDA focuses only on the area classified as clear sky by a previous-applied cloud detection algorithm and it aims to detect the thin cirrus previously misidentified as clear sky providing an estimate of the detection error. It was developed primarily for IASI-NG, but also in parallel for its

predecessor, IASI, so to quantify the improvement brought by IASI-NG with respect to IASI and to evaluate the TCDA performance on real observations, currently available for IASI only. The core of the TCDA is the NN, preferred over simpler approaches, like threshold algorithms or linear regressions, due to its higher versatility, power, and widespread use in the scientific field to solve complex problems. The TCDA NN was trained and tested with IASI-NG (or IASI) simulated observations, calculated processing the global numerical weather reanalysis of the ECMWF—5 generation reanalysis (ERA5) archive [18] with the σ-IASI radiative transfer model [19]. TCDA was developed in the framework of the ComboCloud project, funded by EUMETSAT [20].

This article is structured as follows. Section II describes the dataset used for the TCDA development and the criteria adopted for selecting thin-cirrus profiles to be used for IASI-NG/IASI simulation. It also describes the cloud products from active and passive sensors used for IASI-TCDA validation and comparison. Section III describes the TCDA methodology, with an overview of NN, the method used to estimate the total error to associated with TCDA output and, finally, the validation results of both IASI and IASI-NG TCDA on an independent simulated dataset. Section IV discusses the comparison of IASI-TCDA with MODIS, VIIRS, and SEVIRI cloud products and its validation against CPR/CALIOP cloud products. To facilitate the reading of this article, the Nomenclature provides a comprehensive list of acronyms and abbreviations used throughout the text.

II. Instrument and Data Descriptio

This section describes the sensors involved in this study. Section II-A describes IASI and IASI-NG and the criteria adopted to select the dataset used for training and validation. Section II-F describes the cloud products from active and passive sensors collected to build the observational datasets for IASI-TCDA validation.

A. IASI/ IASI-NG: Instrument Description

IASI is a high spectral resolution sounding instrument based on a Fourier transform spectrometer. It measures 8461 spectral samples in the range [645, 2760] cm^{-1}, with a spectral resolution of 0.5 cm^{-1} and a spectral sampling of 0.25 cm^{-1}. It has been flying on EUMETSAT polar system (EPS) Metop A-B-C satellites series since 2006. IASI scans across-track 30 EFOV. Each EFOV contains four IFOVs disks with a ground diameter of 12 km at the SSP. Moving away from the SSP, the IFOV assumes an elliptic shape with axes increasing to about 39 km (across-track direction) and 20 km (along-track direction) at the swath edge. At the SSP, the four IFOVs are 24 km apart within the EFOV, whose dimension is 48 $\times$ 48 km.

IASI-NG is the evolution of IASI. It differs from IASI for the improved radiometric and spectral characteristics as well as for the scan geometry. IASI-NG will measure 16 921 samples in the spectral range [645, 2760] cm^{-1} with a spectral resolution (0.25 cm^{-1}), spectral sampling (0.125 cm^{-1}), and signal-to-noise ratio improved by a factor of 2 with respect to IASI. The higher spectral resolution and signal-to-noise ratio will directly result in higher vertical resolution and product accuracy [21]. Regarding scan geometry, IASI-NG will scan across-track 14 FORs; each FOR will contain 4 $\times$ 4 IFOVs whose dimension at the SSP and the swath edge will be the same as IASI IFOV. In detail, the distance between two consecutive IFOV centers will be about 23.83 km and the distance between FORs will be about 32 km, resulting in a 100 $\times$ 100 km FOR at the SSP [22].

B. IASI/ IASI-NG: Simulated Dataset and Criteria Adopted for Thin-Cirrus Dataset Implementation

The TCDA implements an NN approach, which is trained, validated, and tested using a dataset of IASI-NG and IASI simulated observations. Two NNs were trained separately for IASI and IASI-NG using a dedicated global dataset of 300 000 samples (over both land and sea surfaces), 50% characterized by clear sky and 50% by thin-cirrus presence. IASI and IASI-NG simulated observations are obtained by applying the σ-IASI radiative transfer code to a set of surface and atmospheric data selected from the ERA5 climate dataset [19]. The σ-IASI is a fast line-by-line radiative transfer scheme that simulates both clear and cloudy spectral radiances for a given set of geophysical parameters. The radiometric noise provided by CNES is used for IASI [23], while radiometric noise for IASI-NG was assumed to be half of that of IASI [24]. In σ-IASI, cloud ice particles were represented using spheres. Realistic assumptions on CER and size distributions have been made to produce IR radiative transfer calculations, ice cloud d_e (diameter effective) from Wyser [25] in which the shape distribution $n(L)$ is determined by the use of a mixed distribution, Γ distribution [26] for small particles ($L < 20$ μm), and power-law distribution [27] for the larger one ($L > 20$ μm). Both these distributions have been parameterized with respect to the B parameter

$$B = -2 + 10^{-3}(273 - T)^{1.5}\log_{10}\left(\frac{\text{CIWC}}{\text{CIWC}_0}\right) \quad (1)$$

where T is the atmospheric layer temperature (K), CIWC is the corresponding ice water content ($g{\cdot}m^{-3}$), and $\text{CIWC}_0 = 50\,g{\cdot}m^{-3}$. Thus, d_e is obtained with B using a 3rd-order polynomial approximation

$$d_e = 377.4 + 203.3B + 37.91B^2 + 2.3696B^3. \quad (2)$$

According to Wyser, the d_e range of values is limited in $[10, 100]\mu$m. The set of ERA5 atmospheric state vectors, needed for the radiative transfer calculations, includes the thin-cirrus dataset used for IASI and IASI-NG TCDA implementation. Both ERA5 hourly data on PLs and single level have been used for building the dataset of atmospheric profiles on 37 PLs and surface parameters, respectively. Global data for four representative days (1 January, April, July, and October 2019), each at four synoptic hours (00, 06, 12, and 18), equally spaced on a regular latitude–longitude grid at 0.125° $\times$ 0.125° resolution, have been selected to capture both seasonal and diurnal cycles. The initial ERA5 dataset consists in 2880 $\times$ 1441 $\times$16 (longitude $\times$ latitude $\times$ time dimensions) samples for each variable.

Here, only the subset of profiles satisfying thin-cirrus selection criteria is used for simulating IASI and IASI-NG radiances. The ERA5-profiles thin-cirrus subset is hereafter denominated TCDA-db. The ERA5 geophysical parameters considered for the selection of isolated thin-cirrus samples are given as follows:

1) HCC, MCC, and LCC, providing CF at three levels in atmosphere;
2) TCC indicating CF throughout the atmosphere;
3) CCF giving the portion of grid box covered by cloud for each of the 37 PLs;
4) TP;
5) TCIW;
6) SCIW;
7) TCLW;
8) SCLW;
9) CBH.

A sample is considered thin cirrus contaminated if it satisfies the following tests.

1) Test 1 requires that only ERA5 profiles with TP $=$ 0 mm, TCC $>$ 0, LCC $=$ MCC $=$ 0 and HCC $>$ 0, and CCF $=$ 0 for PLs higher than 450 mb can be considered thin cirrus, so to exclude rainy samples and thin cirrus overlapping low or/and medium clouds.
2) Test 2 considers the presence of cloud ice content, i.e., the TCIW $>$ 0 and SCIW $>$ 0 as indicators of cirrus presence in conjunction with the absence of cloud liquid water (TCLW $=$ 0 and SCLW $=$ 0). The TCIW and SCIW upper limits for thin cirrus given by Mace et al. [28] are adopted in this study because they were determined for a thin-cirrus subset representative of different atmospheric conditions related to cold and warm seasons in large-scale ascent and large-scale subsidence (TCIW $<$ 18 g·m^{-2} and SCIW $<$ 0.012 g·m^{-3}).
3) Test 3 considers CBH. Taking as a reference the latitudinal distribution of cirrus height determined by Sassen et al. [29] on one year of CloudSat and CALIPSO combined observations, only the ERA5 samples with CBH $\geq$ 5 km were included in TCDA-db. The CBH is determined for the highest PLs where 0 $<$ SCIW $\leq$ 0.012 g·m^{-3}. Test 3 is used as a further control with respect to Test 1, in order to exclude ERA5 profiles with HCC $>$ 0 and MCC $=$ LLC $=$ 0 but with SCIW $>$ 0 at CBH $<$ 5 km.

The TCDA-db obtained from the initial ERA5 dataset after verifying the thin-cirrus tests includes ~260 000 samples. This was further reduced to 50 000 samples because of the high computational cost required for simulating hyperspectral radiances and, successively, for the NN training. TCDA-db subset is denominated TCDA-db-50k. Fig. 1 shows the distribution of thin-cirrus occurrences for the variables used in the three tests, to assess their agreement with other studies reported in the literature and the representativeness of the database. Fig. 1 (top left) shows the latitudinal distribution of CBH for the TCDA-db-50k profiles. The CBH range agrees with the values reported in [28] and [30]. The spatial distribution, with maximum values for thin-cirrus CBH occurring in the tropical belt, is similar to the latitudinal distribution of cirrus CBH derived by Sassen et al. [29] analyzing one-year data detected by CloudSat and Calypso. The combined detection characteristics of both Cloudsat radar and Calypso lidar were also used by Haladay and Stephens [2] to build a two-year tropical thin-cirrus dataset between 20° N and 20° S latitude. They found that thin-cirrus CBH ranges from 11 to 16 km in the tropical belt, while Fig. 1 (top left) shows CBH values lower (~ 9 km $\leq$ CBH $\leq$~ 13 km) in the same area. Moreover, the ranges determined by Haladay and Stephens [2] for SCIW (0.002 g·m^{-3} $\leq$ SCIW $\leq$ 5 $\times$ 10^{-3} g·m^{-3}) and TCIW (1g·m^{-2} $\leq$ TCIW $\leq$ 6g·m^{-2})are similar to the SCIW and TCIW ranges spanned by TCDA-db-50k: 0.002 $\times$ 10^{-3}g·m^{-3}$\leq$ SCIW $\leq$ 3.6 $\times$ 10^{-3}g·m^{-3} [see Fig. 1 (bottom left)] and 0.03g·m^{-2} $\leq$ TCIW $\leq$ 7.3g·m^{-2} [see Fig. 1 (bottom right)].

Overall, Fig. 1 shows that the TCDA-db-50k covers all the ranges of thin-cirrus base height, SCIW, and TCIW for all the latitudes in agreement with the abovementioned studies. So long as the TCDA aims to detect isolated thin cirrus not included in multilayer clouds, limiting the need of a very large data set including different clouds combinations, TCDA-db-50k can be deemed representative for the different types of thin cirrus and therefore suitable for training NNs. To this end, TCDA-db-50k was processed with σ-IASI to simulate 50 000 $\times$ 16 921 IASI-NG and 50 000 $\times$ 8421 IASI radiances at three vertical zenith angle (VZA) (0°, 20°, and 44°) both for clear-sky and cirrus conditions.

C. Validation Observational Dataset: CPR and CALIOP

The CPR and CALIOP are the instrument onboard CloudSat and CALIPSO, respectively. The joined CloudSat/CPR and CALIPSO/CALIOP 2B-Geoprof-Lidar [31], [32] and 2C-ICE [33] cloud products were considered for TCDA validation purposes. 2B-Geoprof-Lidar combines the CPR and CALIOP observations to determine cloud properties such as the CF, the CL, and the LB for each CL. 2C-ICE is the CloudSat and CALIPSO Ice Cloudy Property Product that takes as input a combination of the CPR reflectivity and CALIOP attenuated backscattering coefficients at 532 nm to retrieve cloud properties information more accurately than the radar-only product. In particular, the 2C-ICE cloud properties used in this study to define thin-cirrus-contaminated IASI IFOV are the COT and the ice water path (IWP). The 2B-Geoprof-Lidar and 2C-ICE cloud properties were used simultaneously to define the collocated IASI clear or thin-cirrus-contaminated IFOVs. Since an IASI IFOV is never completely covered by the CPR pixels, a homogeneous criterion involving AVHRR measurements collocated within IASI on the METOP platform has been used. In detail, the 10.8-μm radiance and the CI included in AVHRR Level 1B product have been used for the selection of IASI homogeneous IFOVs [34].

D. Comparison Observational Dataset: MODIS

The MODIS [35] Cloud Mask Level 2 product, available at 1-km and 250-m spatial resolutions at the SSP, has been exploited for comparison with TCDA results. This product is

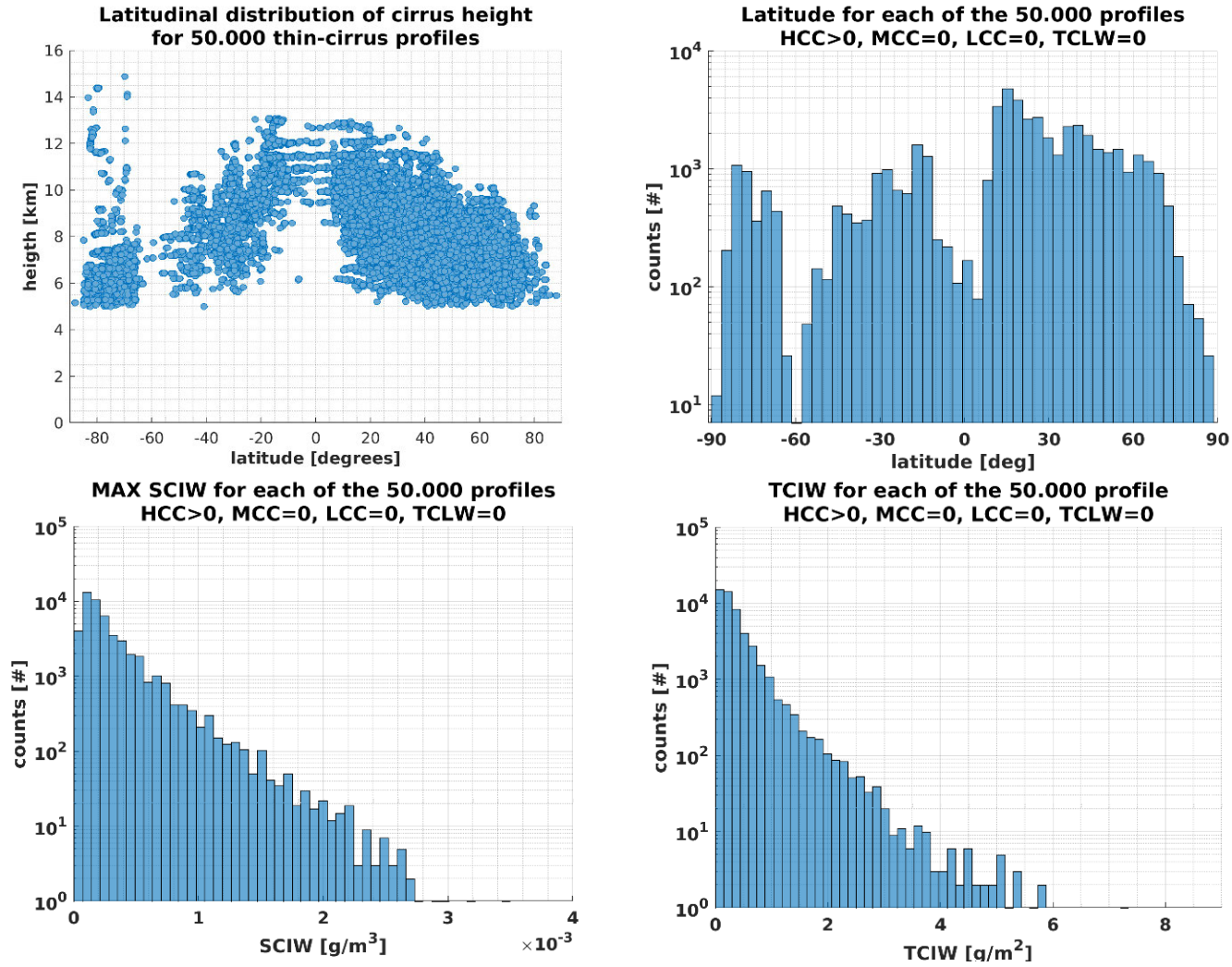

Fig. 1. (Top left) Latitudinal distribution of CBH for the TCDA-db-50k profiles. (Top right) Histogram of latitude for the TCDA-db-50k profiles. (Bottom left) Histogram of max SCIW for TCDA-db-50k profiles. (Bottom right) Histogram of TCIW for TCDA-db-50k profiles.

named MOD35 for MODIS on the Terra satellite and MYD35 for MODIS on the Aqua satellite (hereinafter M∗D35 for both MOD35 and MYD35). The M∗D35 algorithm [35], [36] employs a series of visible and IR threshold and consistency tests to define confidence that a MODIS FOV is cloudy. In detail, the M∗D35 algorithm includes two thin-cirrus tests: the IR-thin-cirrus and the 1.38-μm-thin-cirrus tests. The IR-thin-cirrus test is based on more ice absorption at larger wavelengths; it applies the split window technique [37], using the BT difference between the MODIS bands centered at 11 and 12 μm to indicate the presence of thin cirrus during daytime and nighttime. The 1.38-μm-thin-cirrus test compares the reflectance in 1.38 μm with a threshold [4] to reveal the presence of transmissive cirrus clouds in the upper troposphere under daytime viewing conditions.

E. Comparison Observational Dataset: VIIRS

A further comparison dataset for the TCDA was implemented based on the VIIRS cloud products. VIIRS is aboard two satellites: Suomi NPP (Suomi NPP) since 2011 and NOAA-20 since 2017 [38]. The VIIRS cloud products used for comparison with TCDA are the VIIRS EDR CCL and COT [39]. The EDR-CCL and EDR-COT are cloud products gridded at 6-km resolution and derived from several other intermediate products, e.g., COPs, CTP, PPC, and CBH at the original 750-m spatial resolution. In particular, EDR-CCL gives information about the CT (Stratus, Altocumulus, Cumulus, Cirrus, and Cirrocumulus) and CF on four layers from the top of the atmosphere (layer 1) to layer near surface (layer 4). EDR-COT values are retrieved by using observations at 0.672, 1.24, and 1.61 μm during daytime (solar approach based on two-channel correlation technique of Nakajima and King [40]) and observations at 3.70, 8.55, 10.76, and 12.01 μm during nighttime (IR approach that follows the two-channel cirrus technique of Ou et al. [41]).

F. Comparison Observational Dataset: SEVIRI

The MSG/ SEVIRI [42] OCA product, developed and distributed by EUMETSAT [43], has been considered for comparison with IASI-TCDA results. The OCA algorithm provides key cloud parameters through an optimal estimation method ingesting all the SEVIRI spectral measurements

simultaneously. For each cloudy SEVIRI FOV, the OCA algorithm classifies as multilayered, single-layer water, or single-layer ice clouds. Moreover, it gives information about COT, CTP, and CER for up to two atmospheric layers.

III. TCDA Methodology

This section describes the TCDA. Section III-A details the NNs configurations and training. Section III-B details the error estimation procedure. Section III-C validates the algorithm using a simulated dataset.

A. Definition of NNs

The core of the TCDA consists of two feedforwards fully connected NNs with two hidden layers, whose general principles are summarized in "Appendix A." One NN was developed for IASI-NG and the other was developed for IASI, by using the configuration described in "Appendix B." Both the NNs use the same procedure for training, architecture definition, and input selection, and similar training datasets with 300 000 samples (see Section II-B). Each dataset is randomly split into three sub-datasets, namely, the training dataset (60%, 180 000 samples), validation dataset (20%, 60 000 samples), and test dataset (20%, 60 000 samples), used to calculate weights and biases, to tune some hyperparameters and to assess the performance, respectively. Each sample of the datasets consists of the following 104 variables:

1) the first 100 PCA outputs of the IASI-NG/IASI radiances [44], [45];
2) the cosine of the scan angle;
3) the cosine of the latitude;
4) the land fraction;
5) the flag 0/1 for clear sky/thin-cirrus presence.

The preliminary selected 103 variables were examined as possible inputs for the NNs, through a systematic procedure of analysis and removal of the unnecessary inputs as detailed in "Appendix C." The output values 0/1 have been chosen so that their application to real data returns values ranging continuously in [0,1]. Although 100 PCs may seem too many, the adopted procedure for input analysis and removal of the unnecessary ones requires a preliminary overestimation of the number of inputs. Table I summarizes the selected inputs and the main NN configurations.

The output values can be considered as the probability of thin-cirrus presence; the final output of the TCDA is then "clear-sky condition" if the NN output is in the range [0, 0.5] or "thin-cirrus presence" if the NN output is in the range [0.5,1]. It was chosen as a continuous output in the range [0,1], instead of a discrete output 0/1, in order to evaluate it in a metric for the estimation of the associated error, as explained in Section III-B. In this way, however, values slightly lower or higher than 0.5, for instance, 0.49 or 0.51, would be classified as "clear-sky condition" and "thin-cirrus presence," respectively, despite being very close to each other. To overcome this issue, an estimation error will be provided together with the TCDA output, as explained in Section III-B, developed so that its value is maximum when the NN outputs are around the value 0.5, while it decreases toward zero as the outputs reach 0 or 1.

TABLE I
Main Features of the Optimized NN Architectures for TCDA Retrievals

Sensor	Input	- Hidden-layers nodes number - Transfer functions
IASI-NG	67 IASI-NG PCs (1÷41, 43÷66, 70, 72) 3 ancillary (scan-angle cosine; latitude cosine; land fraction)	10 - 4 tangent sigmoid – logarithmic sigmoid – logarithmic sigmoid
IASI	39 IASI PC (1÷36, 38÷40) 2 ancillary (scan-angle cosine; latitude cosine)	23 - 8 logarithmic sigmoid – logarithmic sigmoid – logarithmic sigmoid

B. Total Error Estimation

A method was developed to associate the NNs output with an estimate of the thin-cirrus detection error, exploiting the dichotomous statistic on the validation dataset. To this aim, hits are the thin-cirrus-in-reference observations correctly detected by TCDA, misses are the thin-cirrus-in-reference observations not-correctly detected by TCDA, false alarms are the clear-sky-in-reference observations not-correctly detected by TCDA, and correct negatives are the clear-sky-in-reference observations correctly detected by TCDA. Using these quantities, we define the total error, which is the ratio of the sum of misses and false alarms to the total observations

$$\text{total error} = 1 - \text{accuracy} = \frac{\text{misses} + \text{false alarms}}{\text{total}} \tag{3}$$

where accuracy indicates the fraction of thin-cirrus and clear-sky observation correctly detected. To obtain an estimated total error, the output of the validation dataset has been divided into 20 equally spaced intervals with a minimum amplitude of 0.05. The amplitude is gradually increased with steps of 10^{-4} to ensure at least 100 samples in each interval, as shown in Fig. 2 by means of the horizontal bars, for both IASI-NG and IASI algorithms. For each interval, a contingency table and the related total error are determined, and finally, two weighted quadratic fits are calculated. By using the coefficients of these fits, it is possible to estimate, for each NN output, the corresponding estimated total error that represents the estimated uncertainty of clear-sky or thin-cirrus detection. For example, considering the IASI-NG fit [see Fig. 2 (left)], the value 0.05 of the NN output leads to a clear-sky scenario because it is less than 0.5, with an estimated total error of 5.7%

$$y = -0.36 \cdot 0.05^2 + 1.11 \cdot 0.05 + 0.0023 = 0.057(5.7\%). \tag{4}$$

Otherwise, the value 0.75 of the NN output leads to a thin-cirrus presence because it is greater than 0.5, with an estimated total error of 29%

$$y = -0.63 \cdot 0.75^2 - 0.06 \cdot 0.75 + 0.69 = 0.29(29\%). \tag{5}$$

C. TCDA Validation With Simulated Dataset

The statistical assessment of TCDA is evaluated on the test dataset, which consists of 20% of the initial database

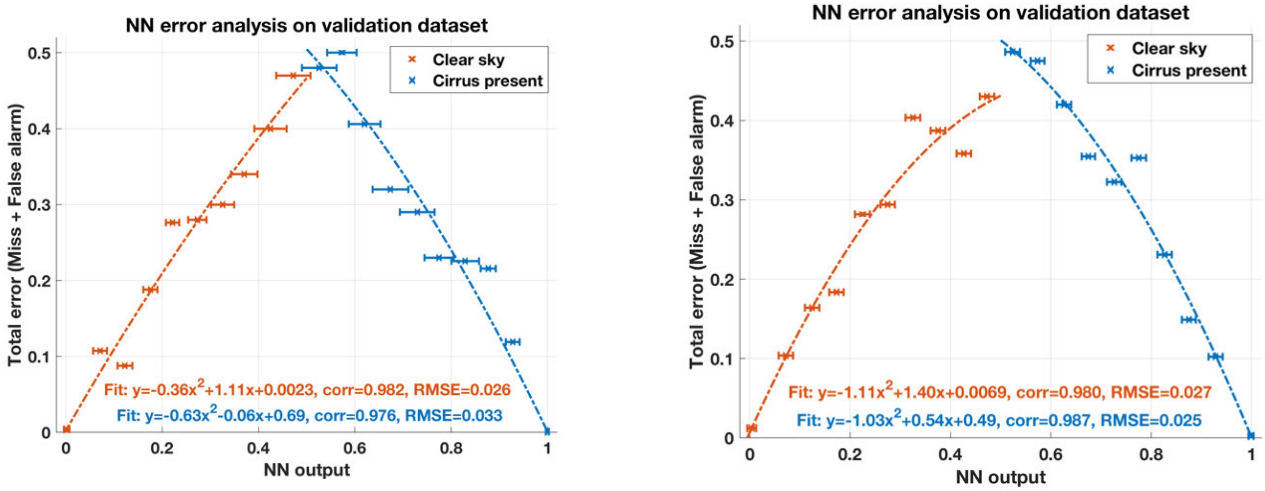

Fig. 2. Result of the weighted quadratic fit for (left) IASI-NG and (right) IASI.

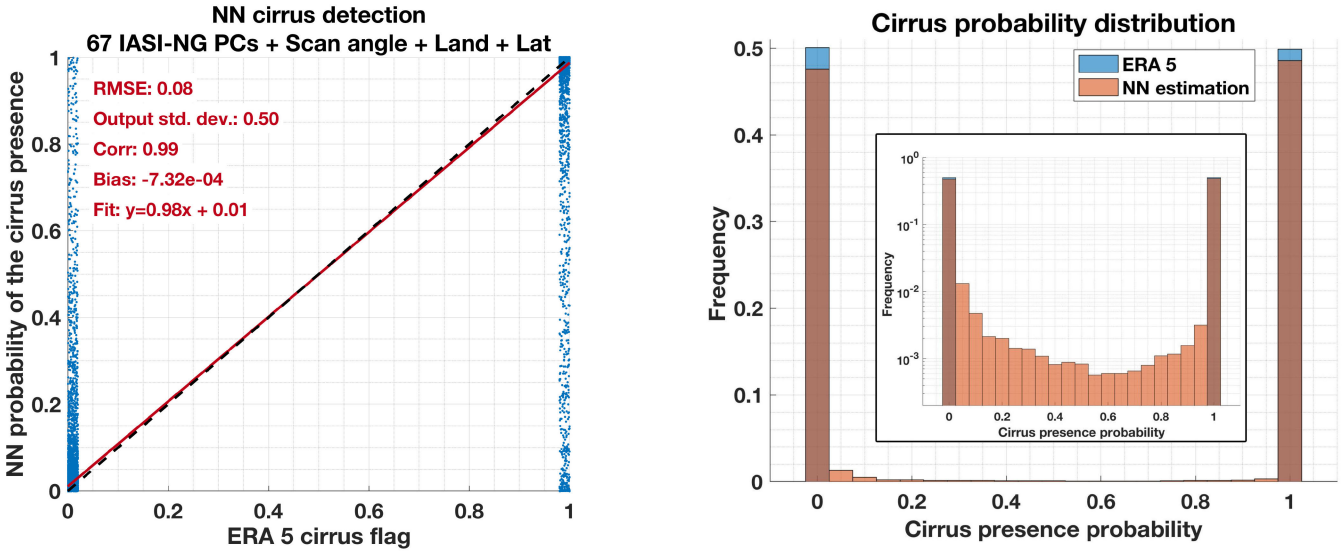

Fig. 3. Validation of TCDA for IASI-NG. (Left) Scatter plot of NN versus reference truth (ERA5). (Right) Histograms of NN output and reference truth (the onset is on a logarithmic scale).

not used in the previously described processes (training and validation). Figs. 3 and 4 show the scatterplots between the NN output and the reference target, both for IASI-NG and IASI, together with the histograms, showing that the distributions of NN outputs and reference targets are rather similar. The correlation between the NN output and the reference target is about 0.99 for IASI-NG and 0.96 for IASI, and the bias error is about 10^{-3} for IASI-NG and 10^{-2} for IASI, while the RMSE is about 16% and 30% of the target standard deviation for IASI-NG and IASI, respectively. The differences between retrievals over land and sea surfaces are negligible. As expected, the statistical results obtained for IASI-NG are better than those obtained for its predecessor, which will be also confirmed by the dichotomous statistical assessment in the next section.

The POD, the FAR, the bias, and the accuracy are defined as follows.

1) POD = hits/(hits + misses) (range: 0–1; perfect score: 1).
2) Bias = (hits + false alarms)/(hits+ misses) (range: 0–∞; perfect score: 1).
3) FAR = false alarms /(hits + false alarms) (range: 0–1; perfect score: 0).
4) Accuracy = (hits + correct negatives)/total (range: 0–1; perfect score: 1).

These statistical parameters are fully described in [46].

Finally, from Table II, the dichotomous scores are summarized and compared in Table III.

Overall, the POD, FAR, bias, and accuracy are very close to their perfect values, thus indicating a good ability of the TCDA to correctly detect the presence of thin cirrus on the independent test dataset. The statistical scores separately for land and sea surface are not shown because they are identical to those of Table III. This result implicitly confirms the

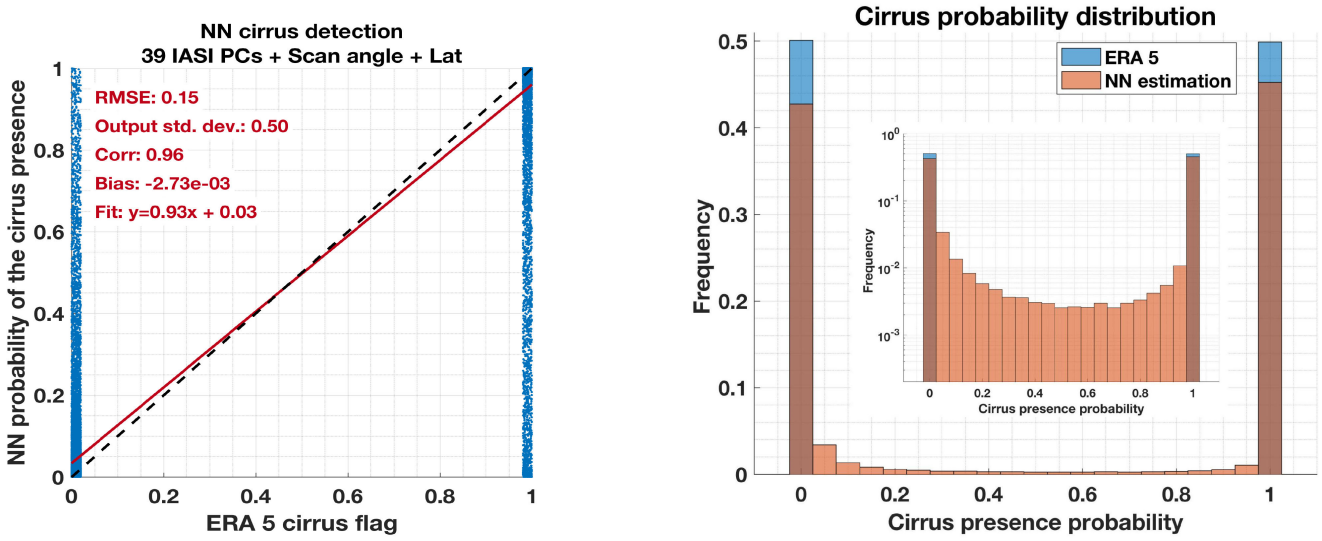

Fig. 4. Statistical results from validation of TCDA for IASI. (Left) Scatter plot of NN versus reference truth (ERA5). (Right) Histograms of NN output and reference truth (the onset is on a logarithmic scale).

TABLE II

CONTINGENCY TABLE FOR TCDA FROM IASI/IASI-NG. (aCORRECT POSITIVES—bFALSE ALARMS—cMISS—dCORRECT NEGATIVES)

		ERA5 cirrus presence (reference)		
		Yes	No	Total NN
NN cirrus presence	Yes	28867/29610[a]	619/160[b]	29486/29770
	No	1073/335[c]	29441/29895[d]	30514/30230
	Total ERA5	29940/29945	30060/30055	60000/60000

TABLE III

SUMMARY OF DICHOTOMOUS SCORES FOR TCDA FROM IASI AND IASI-NG

SCORE	IASI	IASI-NG	PERFECT VALUE
SAMPLE	60000	60000	
POD	0.96	0.99	1.0
FAR	0.02	0.01	0.0
BIAS	0.98	0.99	1.0
ACCURACY	0.97	0.99	1.0

goodness of the choice to use a unique NN for land and sea surface, instead of two separate NN, as frequently happens in the retrieval of atmospheric parameters by means of satellite observations. Statistical scores for TCDA applied to IASI-NG are better than those for TCDA applied to its predecessor IASI. This is likely a result of the IASI-NG improvements in spectral resolution and noise level. In detail, IASI-NG TCDA outperforms IASI TCDA by 3% in POD, 50% in FAR, 1% in bias, and 2% in accuracy.

IASI and IASI-NG test datasets were also used to validate the error estimation, as derived in Section III-D. For both the sensors, the NN output obtained from the test dataset was divided into 12 equally spaced intervals whose width was initially 0.08 and it was gradually increased, without exceeding the maximum width of 0.5, to ensure at least 100 samples in each interval. For each interval (vertical bars in Fig. 5), the corresponding contingency table and the true total error were computed. Fig. 5 shows the scatterplot of the true total error determined for NN versus NN estimated error determined by using the equations in Fig. 2. The excellent correlation (∼0.97 for IASI-NG and 0.99 for IASI) and low RMSE (∼24% and 13% of target standard deviation, respectively) suggest good capability for this method to estimate the total error.

IV. IASI TCDA VALIDATION AND COMPARISON WITH REAL DATA

In addition to the validation against simulated data reported in Section III-C, IASI TCDA was validated against real observations obtained from CPR/CALIOP cloud products. Active sensors, such as lidar, are able to recognize thin cirrus with matchless accuracy [2]. In particular, the joint use of CPR-radar and CALIOP-Lidar gives better results in identifying thin cirrus than using them separately. In fact, while CPR is able to penetrate thick layers that attenuate the CALIOP signal, CALIOP is able to detect the scattering from very tenuous clouds as well as to sense the top of optically thin ice clouds that are transparent to CPR [30]. Moreover, IASI TCDA was also compared with three different datasets of cloud products from the passive sensors MODIS, VIIRS, and SEVIRI described in Sections II-D, II-E, and II-F, respectively.

For spatial collocation, IASI was designated as the principal instrument; each IASI IFOV is defined as clear or thin-cirrus contaminated according to the cloud properties of the reference data falling within it. Only the IASI IFOVs completely covered by reference FOVs have been considered for validation and comparisons.

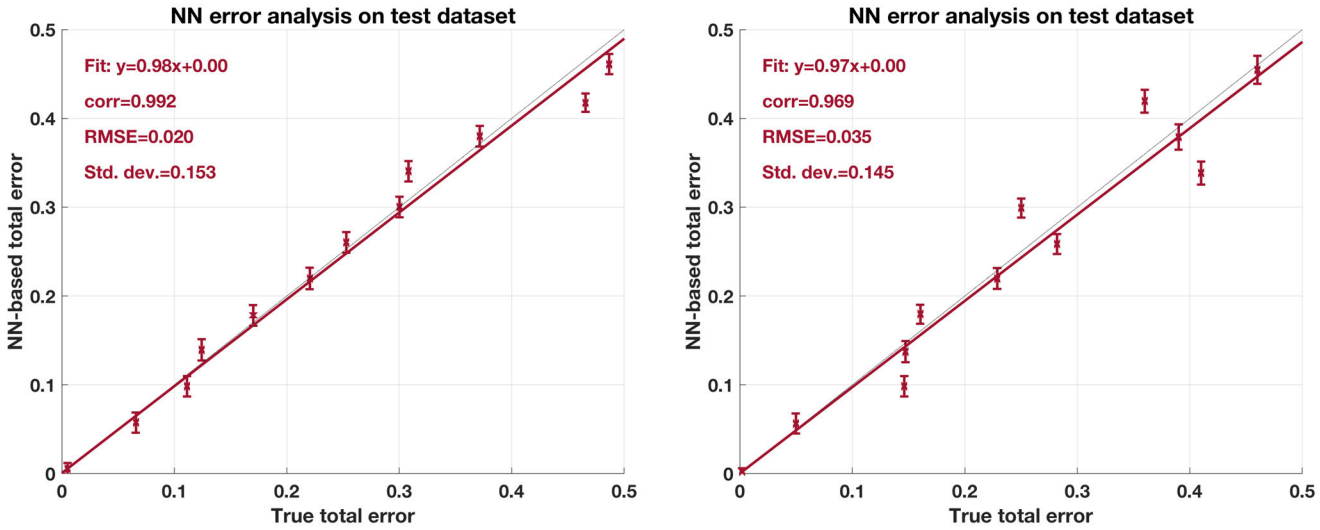

Fig. 5. True versus NN-estimated total error for (left) IASI and (right) IASI-NG.

The following sections show the results of validation carried out against CPR-CALIOP/2B-Geoprof and 2C-ICE cloud products and the results of comparison with the three passive sensors/products MODIS/M∗D35, VIIRS/EDR, and SEVIRI/OCA. A dichotomous statistical assessment for TCDA was performed, including the accuracy, POD, FAR, and bias.

A. Results of TCDA Validation Against CPR/CALIOP-2B-Geoprof-Lidar and 2C-ICE Products

2B-Geoprof-Lidar and 2C-ICE files collocated with MetOp-A/B IASI overpass in 2016–2017 were jointly used to define a dataset of clear and thin-cirrus-contaminated IASI IFOVs. The 2C-ICE and 2B-Geoprof-Lidar are sampled at the CPR spatial resolution (https://www.cloudsat.cira.colostate.edu/cloudsat-static/info/dl/2b-geoprof-lidar/2B-GEOPROF-LIDAR_PDICD.P2_R04.20070604. pdf). The CPR FOV is considered thin-cirrus contaminated when only the highest levels are cloudy ($LC > 0$ and LB $\leq$ 450 hPa) and the related COT and IWP satisfy the conditions $0 < \text{COT} \leq 1.72$ and $0 < \text{IWP} \leq 18$ g·m^{-2} [28]. The IWP value is not always available, so the thin-cirrus selection was principally based on the COT test. In total, 314 IASI-CPR matchups have been collected for TCDA validation, all located at high latitudes (65°N–74°N). For simplicity, the CPR FOV included in IASI IFOV is denominated CPRinIASI. The number of CPRinIASI along the CloudSat depends on the satellite view angle and ranges from 2 to 15. The IASI IFOV collocated with CPR is considered for validation only when all the CPRinIASIs are clear (in this case, the IASI IFOV is considered "clear") or a fraction of CPRinIASIs is thin cirrus and the remaining CPRinIASIs are clear (in this case, the IASI IFOV is considered "thin-cirrus contaminated"). However, this information cannot be simply extended to the entire IASI FOVs since IASI IFOVs (12 × 12 km at the SSP) are always larger than the area covered by the CPR swath (2.5 × 1.4 km). CPR overlaps IASI IFOV along a circle arc, and thus, the IASI IFOV is never completely covered by the CPR pixels. In order to get information on the entire IASI IFOVs previously classified as thin-cirrus contaminated or clear based on the 2B-Geoprof-Lidar/2C-ICE information only, a further test for investigating the homogeneity of the IASI IFOVs has been done by considering the AVHRR-CI (https://www.eumetsat.int/media/38675) and the 11.5-μm radiances collocated within the IASI IFOV. The homogeneity test consisted in applying the criteria proposed in [47] that established the homogeneity of an IASI IFOV based on the overall AVHRR cluster statistics. Farouk et al. assumed that an IASI IFOV, including several classes (i.e., clear class and cloudy classes), characterized by small standard deviations and similar mean radiances, can be more homogeneous than an IASI IFOV with a single cloudy or clear class. In detail, only the IASI IFOVs containing a cirrus class (ensemble of AVHRR FOVs satisfying only the CI cirrus test) and/or a clear class (ensemble of AVHRR FOVs not satisfying any CI cloud test) have been considered for homogeneous test. Finally, the IASI IFOV was declared thin-cirrus-contaminated only after satisfying the homogeneity criteria [48]. Among the 314 prescreened thin-cirrus-contaminated IASI IFOVs, 248 passed the homogeneity test. The same procedure was applied to the IASI IFOVs previously classified as clear according to 2B-Geoprof-Lidar/2C-ICE cloud products, so to select 248 clear homogeneous IASI IFOVs. In applying the homogeneity criterion, we assumed to consider the CPR/CALIOP information as "truth" when it disagrees with the AVHRR-CI in detecting thin-cirrus cloud. In fact, some IASI IFOVs defined as thin-cirrus contaminated by the 2B-Geoprof-Lidar/2C-ICE cloud products were classified as "clear" in terms of the AVHRR-CI, but this is due to the better ability of active sensor to detect optically thin layer if compared with passive sensors that often misidentify thin cirrus as clear sky [16], [17].

TCDA detects 117 IASI-IFOVs as thin cirrus among the 248 thin-cirrus-contaminated IASI IFOVs and it correctly detects as clear 237 IASI IFOVs among the 248 clear IASI

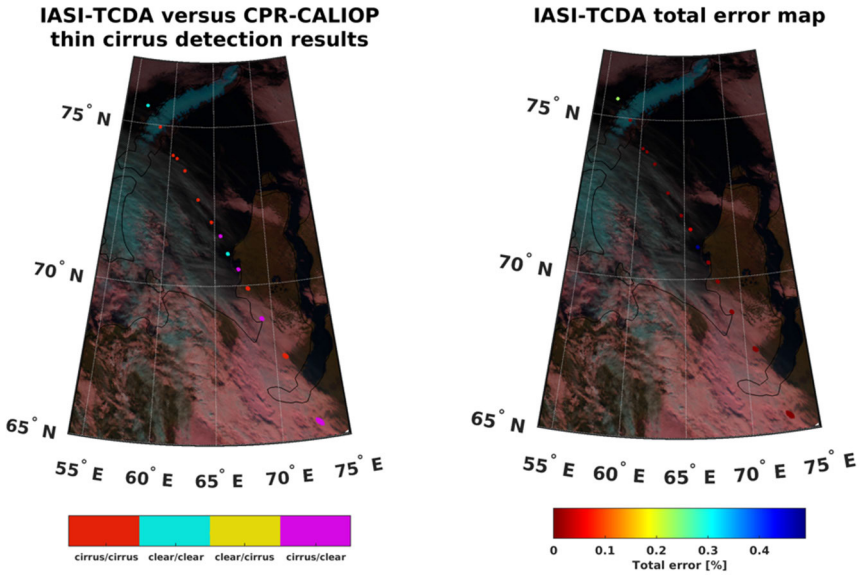

Fig. 6. 25 August 2017 from 17:38 to 17:41 GMT. (Left) IASI-TCDA versus CPR-CALIOP thin-cirrus detection results. (Right) IASI-TCDA total error map. The AVHRR natural-color RGB image consists in the composition of the 1.6-μm reflectance (visualized in red), the 0.87-μm reflectance (visualized in green), and the 0.63-μm reflectance (visualized in blue).

TABLE IV

CONTINGENCY TABLE FOR DICHOTOMOUS STATISTICAL ASSESSMENT OF TCDA APPLIED TO REAL IASI OBSERVATIONS AGAINST COLOCATED 2B-GEOPROF AND 2C-ICE (1), M∗D35(2), VIIRS-EDR(3), AND SEVIRI-OCA(4) CLOUD PRODUCTS

		Reference cloud product cirrus presence		
		Yes	No	Total NN
NN thin-cirrus presence	Yes	$117^{(1)}$ $8049^{(2)}$ $4470^{(3)}$ $1893^{(4)}$	$11^{(1)}$ $4111^{(2)}$ $5722^{(3)}$ $1868^{(4)}$	$128^{(1)}$ $12160^{(2)}$ $10192^{(3)}$ $3761^{(4)}$
	No	$131^{(1)}$ $1830^{(2)}$ $4761^{(3)}$ $119^{(4)}$	$237^{(1)}$ $3139^{(2)}$ $3509^{(3)}$ $144^{(4)}$	$368^{(1)}$ $4969^{(2)}$ $8270^{(3)}$ $263^{(4)}$
	Total	$248^{(1)}$ $9879^{(2)}$ $9231^{(3)}$ $2012^{(4)}$	$248^{(1)}$ $7250^{(2)}$ $9231^{(3)}$ $2012^{(4)}$	$496^{(1)}$ $17129^{(2)}$ $18421^{(3)}$ $4024^{(4)}$

IFOVs. Table IV shows the contingency table for dichotomous statistical assessment of TCDA applied to real IASI observations against colocated 2B-Geoprof and 2C-Ice cloud products [the related samples are indexed with $^{(1)}$] as well as against the other cloud products considered for the comparison. This results in POD = 0.47, bias = 0.52, FAR = 0.09, and accuracy = 0.71. When considering only the 64 IASI IFOVs characterized by COT $\leq$ 0.3 and comparing them with the same number of clear ones, the statistical scores are quite similar: POD = 0.47, bias = 0.47, FAR = 0.0, and accuracy = 0.73. The low FAR indicates the tendency of TCDA to not misidentify clear sky as thin cirrus. Only one IASI IFOV with COT $<$ 0.03 (subvisual cirrus) is present in the dataset and it is correctly classified by TCDA. These results should be taken with care due to the uncertainty caused by the partial coverage of the IASI IFOV by the CPRinIASIs so that the information on the remaining IASI IFOV is derived by the homogeneity tests involving AVHRR radiances that may not reveal inhomogeneity due to presence of very thin cirrus easily misidentified as clear sky at the AVHRR wavelengths. An example of IASI-TCDA cloud detection map and the related TCDA-total-error map are shown on the left and the right of Fig. 6, respectively, for the Metop B-IASI granule acquired on 25 August 2017 from 07:38 to 07:41 GMT. TCDA was applied to the IASI IFOVs declared clear and thin-cirrus contaminated based on the CPR/CALIOP-2B-Geoprof-Lidar and 2C-ICE products. The TCDA-IASI results have been collocated on the corresponding AVHRR RGB composition. The color of the thin-cirrus clouds in the natural-color RGB image depends on the underlying surface. In Fig. 6, the thin-cirrus clouds are recognizable by their elongated shape and cyan color shades. In detail, the IASI IFOVs in red are correctly detected as thin cirrus by TCDA accordingly to the CPR/CALIOP-2B-Geoprof-Lidar and 2C-ICE cloud products and the RGB image. The IASI TCDA results in magenta are related to IASI IFOVs detected as thin cirrus by TCDA at odds with CPR/CALIOP-2B-Geoprof-Lidar and 2C-ICE cloud products classifying them as clear sky. However, the area corresponding to the IASI IFOVs magenta between about 68°N and 70°N of latitude seems to be contaminated by a very thin cirrus, whose presence can be deduced from the hazy effect on the underlying surface. The corresponding total error values, shown on the right of Fig. 6, are lower than 1% testifying the TCDA output precision for this example.

B. Results of IASI TCDA Comparison With MODIS/M∗D35 Product

MODIS orbits acquired between the years 2016 and 2018 with a good temporal and spatial overlap with IASI orbits have been considered for building a dataset of thin-cirrus and

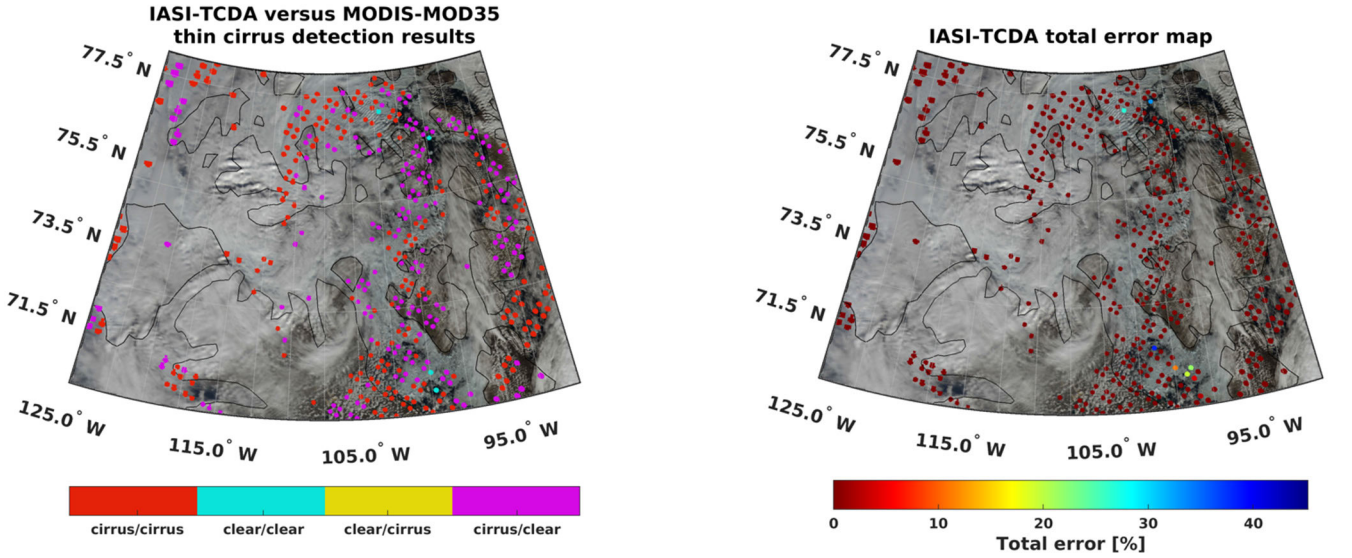

Fig. 7. 18 August 2018 from 18:17 to 18:20 GMT. (Left) IASI-TCDA versus MODIS-MOD35 thin-cirrus detection results. (Right) IASI-TCDA total error map. The MODIS true-color RGB image consists in the composition of the 0.645-μm reflectance (visualized in red), the 0.555-μm reflectance (visualized in green), and the 0.469-μm reflectance (visualized in blue).

clear IASI IFOVs based on the M∗D35 product; 68 orbits were considered, all located at high latitudes. MODIS IFOVs collocated within each IASI IFOV are denominated MODinIASI. Each IASI IFOV is classified as clear when all the MODinIASIs are classified as clear by M∗D35. An IASI IFOV is considered thin-cirrus contaminated when a fraction of the MODinIASIs satisfies at least one of the two M∗D35 thin-cirrus tests described above (see Section II-D), and the remaining are clear. The samples related to the TCDA dichotomous statistics are indexed with $^{(2)}$ in Table IV. In detail, the TCDA dichotomous statistics was applied to 17 129 IASI IFOVs (9879 clear IFOVs and 7250 thin-cirrus IFOVs). The corresponding statistical scores are: POD = 0.81, FAR = 0.34, and bias = 1.23, indicating good detecting skills with some tendency to overestimate thin-cirrus presence.

An example of TCDA application to IASI IFOVs declared clear/thin-cirrus contaminated based on the M∗D35 product is shown in Fig. 7, where the IASI-TCDA cloud detection map and the related TCDA-total-error map are shown on the left panel and the right panel, respectively, for the Metop A-IASI granule acquired on 18 August 2016 from 18:17 to 18:20 GMT. The IASI IFOVs classified as thin-cirrus (in red/magenta) and clear (in cyan/yellow) have been collocated on MODIS true-color RGB image, where the high thick clouds are white, while the high thin ice clouds are characterized by a transparent white color and are recognizable by their filament shape. In particular, the red IASI IFOVs are classified as thin cirrus by the TCDA in agreement with the MOD35 cloud product. Most of the red samples correspond to high thick cloud or multilayered cloud overlaid by thin cirrus. The IFOVs in magenta are clear for MOD35 and thin cirrus for TCDA that, based on the MODIS true-color RGB, mostly detects them correctly (e.g., magenta IFOVs on Bathurst Island, approximately around 75°N latitude and 98°W longitude, as well as those near Stefansson Island, approximately around 72°N latitude and 104°W longitude). The associated total error is mostly around 10%. The statistical scores related to this example are POD = 0.95, FAR = 0.40, bias = 1.59, and accuracy = 0.65.

C. Comparisons With VIIRS/EDR Product

The TCDA was further compared with cloud products available from VIIRS. Unlike the MODIS/M∗D35 cloud mask product, there is no explicit reference to the identification of thin-cirrus clouds in VIIRS cloud products. Because of this, thin-cirrus-contaminated IASI IFOVs were defined based on CI provided by the VIIRS EDR-CCL and VIIRS EDR-COT (described in Section II-E). VIIRS-IASI spatial/temporal coincidences for the years 2016 and 2017 were investigated. For simplicity, VIIRS FOVs collocated within an IASI IFOV are denominated VIIRSinIASI. Only the IASI IFOVs completely covered by the VIIRS FOVs classified as clear or thin cirrus have been considered to implement the clear and thin-cirrus datasets to be used for validation. A VIIRS FOV is considered thin-cirrus contaminated if the following conditions hold.

1) CF greater than 0 only for the highest layer (i.e., EDR-CCL-CF > 0 only for layer 1 and EDR-CCL-CF = 0 for the remaining three layers).
2) CT identified as cirrus (i.e., EDR-CCL-CT equal to "cirrus").
3) COT smaller than 1.72 (i.e., 0 < EDR-COT ≤ 1.72 [49].

An IASI IFOV is considered clear when all the VIIRSinIASIs are clear (EDR-CCL-CF = 0 for all the four layers), while it is considered thin-cirrus contaminated when a fraction of VIIRSinIASIs is classified as thin cirrus and all the remainder VIIRSinIASIs are clear. A total of 9231 thin-cirrus-contaminated and 9231 clear IASI IFOVs were so identified. All the VIIRS FOVs used to validate TCDA are located at high latitudes, as for M∗D35 product.

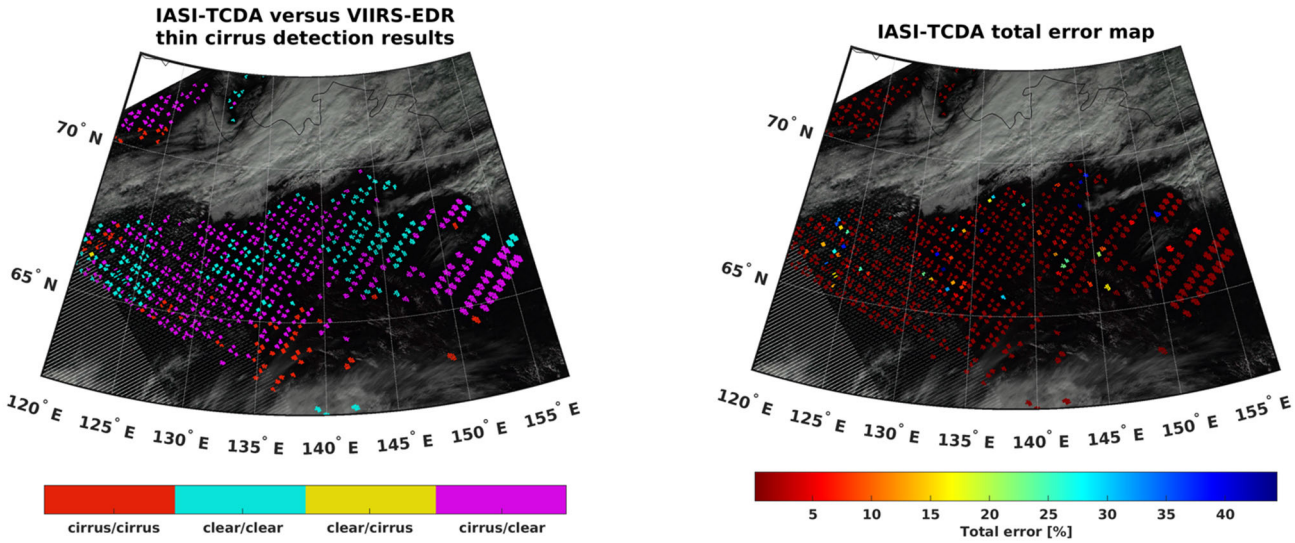

Fig. 8. 16 August 2016 from 02:20 to 02:23 GMT. (Left) IASI-TCDA versus VIIRS-EDR thin-cirrus detection results. (Right) IASI-TCDA total error map.

The samples of the contingency table for the IASI TCDA algorithm comparison with VIIRS EDR product are indexed with (3) in Table IV. The dichotomous statistical scores are characterized by a high FAR (0.56) and a low POD (0.48) with a bias (1.10) that shows a tendency of IASI TCDA algorithm to overestimate thin-cirrus presence when compared with VIIRS EDR product. When considering only the 779 IASI IFOVs characterized by COT $\leq$ 0.3 and comparing them with the same number of clear ones, the statistical results improve slightly, in particular, VIIRS and IASI agree in detecting about 90% of clear IASI IFOVS (FAR = 0.17) and about the 50% of the thin-cirrus-contaminated IASI IFOVs (POD = 0.51). Statistical scores obtained against VIIRS-EDR are worse than those obtained against MODIS-M$*$D35. This may be a consequence of the indirect detection of thin-cirrus-contaminated VIIRS IFOVs. In fact, unlike thin-cirrus dataset built based on M$*$D35 thin-cirrus tests, the IASI IFOV is considered thin-cirrus-contaminated or clear by considering the combination of three different VIIRS-EDR products, which in turn derive from several other intermediate products. Moreover, the high FAR is also due to the low accuracy of VIIRS in detecting cloud area over snow-covered land [49], as evidenced also in [15] where the CRIS-cloud detection algorithm trained on VIIRS cloud mask sometimes misidentifies thin clouds as clear sky. Fig. 8 shows an example of TCDA applied to IASI IFOVs classified as clear or thin cirrus based on VIIRS-EDR cloud product information. As for comparison examples in Figs. 6 and 7, the red/cyan IASI IFOVs are detected as thin cirrus/clear both by TCDA and reference cloud product, while the magenta IASI IFOVs correspond to false alarms. The example is related to the IASI granule acquired on 6 June 2017 from 18:50 to 18:53 GMT collocated on VIIRS true-color RGB image. Comparing the TCDA results with VIIRS true-color RGB image, it can be deduced that TCDA correctly detects clear (in cyan) and thin-cirrus (in red) IASI IFOVs in agreement with VIIRS-EDR cloud product. The FAR (0.87) for this example is very high as it can be seen from the high number of magenta IASI IFOVs in Fig. 8. However, some magenta IFOVs are related to areas where the thin-cirrus presence is correctly detected by TCDA (e.g., below 65°N latitude). Moreover, based on VIIRs-EDR information, IASI IFOVs characterized by the same RGB colors are classified differently (e.g., IASI IFOVs in the area between 64°N–66°N latitude and 135°E–140°E longitude as well as in the area between 64°N–70°N latitude and 150°E–155°E). Due to the high number of false alarms, for this example, the statistical scores are not good (POD = 0.51, bias = 4.16, and accuracy = 0.68). The total error values are about 10% for most IASI IFOVs. Total error values higher than 10% are related to IASI IFOVs detected as clear both by TCDA and VIIRS-EDR product.

D. Results of TCDA Validation Against SEVIRI/OCA Product

All the validation/comparison results obtained by comparing IASI TCDA output with MODIS, VIIRS, and CPR products are based on high-latitude data, because of the collocation between different polar orbiting platforms. To extend the TCDA comparisons to middle latitudes, we also considered the MSG-SEVIRI OCA product (see Section II-E). For simplicity, the SEVIRI FOV included in IASI IFOV is denominated SEVIRIinIASI. Based on the OCA product, the IASI IFOV colocated with MSG-SEVIRI is considered thin-cirrus contaminated when all the cloudy SEVIRIinIASIs are classified as "single layer ice cloud" with 0 < COT $\leq$ 1.72 and CTP < 450 mb. Conversely, the IASI IFOV is considered clear when all the SEVIRIinIASI FOVs are clear. A total of 4024 match-ups were collected for years 2021 and 2022 using IASI on MetOp-A/B. As for MODIS and VIIRS products, also for MSG-SEVIRI, the IASI IFOVs are considered for validation only when they are completely covered by MSG-SEVIRI FOVs. The number of SEVIRI FOVs included in IASI IFOV ranges from $\sim$4 to $\sim$65 depending on the dimensions of the SEVIRI FOVs and IASI IFOV. The samples related to the IASI-TCDA versus SEVIRI-OCA comparison are indexed with (4) in Table V. The corresponding statistical scores reveal

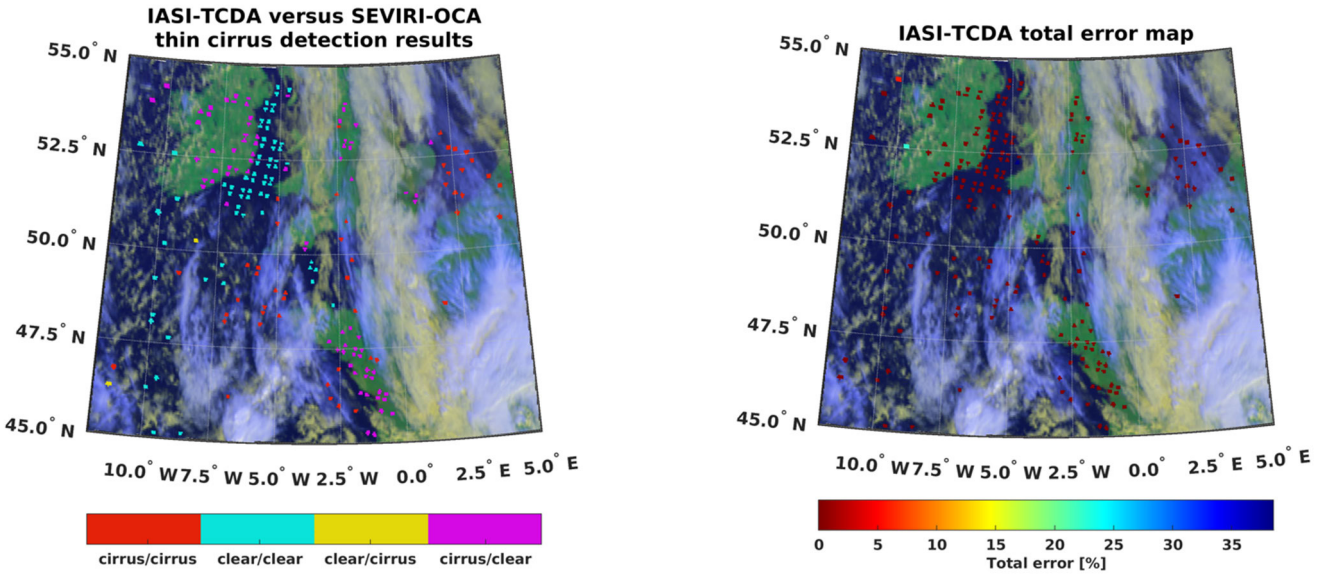

Fig. 9. 19 September 2021 from 10:23 to 10:26 GMT. (Left) IASI-TCDA versus SEVIRI-OCA thin-cirrus detection results. (Right) IASI-TCDA total error map. The SEVIRI natural-color RGB image consists in the composition of the 0.8-μm reflectance (visualized in red), the 0.6-μm reflectance (visualized in green), and the reflectance (visualized in blue).

TABLE V
STATISTICAL SCORES OF IASI TCDA AGAINST MODIS-M*D35, VIIRS-EDR, SEVIRI-OCA, AND CPR/CLOUDSAT-2B-TAU CLOUD PRODUCT

		Dichotomous scores			
		Accuracy	**Bias**	**POD**	**FAR**
Sensor/ product	MODIS/M*D35	0.65	1.23	0.81	0.34
	VIIRS/EDR	0.43	1.10	0.48	0.56
	2B-GEOPROF-Lidar & 2C-ICE/CPR&CALIOP	0.71	0.52	0.47	0.09
	SEVIRI/OCA	0.51	1.87	0.94	0.50

a good POD (0.94) but also a considerable FAR (0.50). Similar statistical results are obtained by analyzing the thin-cirrus-contaminated IASI-IFOVs with COT $\leq$ 0.3. As for the VIIRS and MODIS comparisons, the comparison with SEVIRI shows that TCDA tends to overestimate the thin-cirrus presence (bias = 1.87). This overestimation could also be a consequence of the OCA's high errors in estimating COT related to optically thin cirrus [43].

Fig. 9 shows the results of TCDA applied to IASI IFOVs classified as clear/thin cirrus based on the SEVIRI-OCA cloud product. The TCDA results have been collocated on SEVIRI Natural Color. The example is related to the IASI granule acquired on 19 September 2021 from 10:23 to 10:26 GMT. The thin-cirrus and clear-sky IASI IFOVs are correctly detected according to the SEVIRI-OCA cloud product and RGB image. Some false alarms (in magenta) are related to areas that seem contaminated by very thin cirrus (e.g., the area between 52.5°N–54°N latitude and 2.5°W–0°W longitude) or to the area contaminated by low clouds (e.g., the area between 52.5°N–54°N latitude and 10°W–7.5°W longitude). The statistical scores related to this example are POD = 0.95, FAR = 0.74, bias = 3.77, and accuracy = 0.34, and the total error is lower than 10% for most IASI IFOVs samples.

V. CONCLUSION

The main novelties characterizing the TCDA proposed in this study are given as follows.

1) TCDA is not a standalone algorithm, but it is applied to the IASI/IASI-NG IFOVs classified as clear sky by any cloud detection algorithm in order to refine cloud screening by detecting the thin cirrus previously misidentified as clear sky.
2) TCDA also provides an estimate of the classification error, which indicates the probability that the algorithm misclassified the scenario.
3) TCDA has been developed for both IASI-NG and IASI, thus allowing the following:
   a) continuous use of the algorithm in the transitional phase between the two sensors;
   b) the comparison of IASI/IASI-NG TCDA results to highlight the differences between the two instruments.

TCDA consists in an approach based on feedforward NN for the detection of optically thin cirrus from the new generation sensor IASI-NG and from its predecessor IASI. TCDA exploits the whole IR spectrum in the range [645, 2760] cm^{-1} with a spatial resolution of 0.125 cm^{-1} for IASI-NG and 0.25 cm^{-1} for IASI to give in output the probability that the IASI (IASI-NG) IFOV is thin-cirrus contaminated (TCDA NN output > 0.5) and the corresponding estimated total error. The estimated total error represents the estimated uncertainty of clear-sky or thin-cirrus detection. TCDA uses two distinct NNs for IASI and IASI-NG. Although the NNs were defined and optimized through the same procedure for IASI and IASI-NG, the resulting NN configurations differ in architecture and input selection. This was to be expected, as the two sensors have different characteristics, especially the improved spectral resolution and radiometric signal-to-noise ratio of IASI-NG.

These improvements result in the better statistical scores obtained for IASI-NG with respect to IASI, as demonstrated by validating TCDA with a simulated dataset. In fact, IASI-NG TCDA outperforms IASI TCDA by 3% in POD, 50% in FAR, 1% in bias, and 2% in accuracy. In order to evaluate the performance of TCDA on real observations, IASI TCDA was validated against the combined CPR/CALIOP 2B-Geoprof-Lidar and 2C-ICE cloud products and compared with three observation datasets from passive sensors, i.e., the MODIS-M∗D35, VIIRS-EDR, and SEVIRI-OCA cloud products.

The dichotomous statistical scores are summarized in Table V. The statistics obtained by comparing IASI TCDA with the cloud products from passive sensors seem to indicate a tendency for IASI-TCDA to overestimate thin cirrus, but this tendency is not confirmed by the IASI TCDA validation against active sensor CPR and CALIOP cloud product where the FAR is very low (0.09). The best results are given by comparing IASI-TCDA with MODIS-M∗D35 product, the only cloud product providing an explicit detection of thin-cirrus presence. This is not the case for VIIRS-EDR and SEVIRI-OCA products, which give information about COT and cloud levels, from which the thin-cirrus presence has been derived.

Note that, because of their low COT, thin cirrus is the hardest clouds to detect. This study only partially addresses cloud detection, as TCDA was intended to work on IASI/IASI-NG IFOVs classified as clear by a cloud mask previously applied.

A future study will be dedicated to the implementation of a cloud classification mask, including TCDA, with the aim to exploit IASI-NG hyperspectral data for the detection of multilayer clouds and eventually overlapping thin cirrus.

Appendix A-Neural Network Overview

Under specific conditions, the NNs are the universal approximators of any continuous function with arbitrary accuracy [50]. An NN is a computing system based on perceptrons (neurons o nodes) organized in layers connected in a chain, whose overall length gives the depth of the NN. In the feed-forward NN, the input information $\vec{x}$ moves always forward through the first layer called input layer, then follows into the intermediate levels or hidden layers, and finally into the output layer to obtain the output results $\vec{y}$. In the fully connected NN used in this study, each node of a layer is connected to each node of the next layer by a weight $w_{i,j}^k$, where i and j indicate the indexes of the two nodes of the two contiguous layers k and $k+1$. In the input layer, the output of each node corresponds to a single input, while in the other layers, the output of the node j of the layer k, y_j^k, is

$$y_j^k = \text{trfun}\left[\sum_{i=1}^{M}\left(w_{i,j}^k y_i^{k-1}\right) + b_j^k\right] \tag{A1}$$

where b_j^k is a bias, k is the considered layer with M nodes, and trfun is a transfer function properly chosen to modulate the summation result. The choices for the number of hidden layers, their nodes, and the transfer functions for each layer define the NN architecture. The procedure to calculate the weights and biases is the so-called training phase, which in this study was carried out by using the Levenberg–Marquardt backpropagation algorithm [51], [52] that iteratively adjusts weights and biases minimizing the mean square error (MSE).

To minimize the risk of overfitting and reducing the computational cost by limiting the number of training epochs (i.e., the training cycles used by the backpropagation algorithm through the full training dataset), two early stopping rules were used in this study.

1) The MSE calculated on the validation dataset increases for five consecutive epochs.
2) The MSE gradient with respect to the weights calculated on the training dataset drops below 0.1‰of the variance of the NN output.

When the training stops, the weights and biases corresponding to the epoch with the minimum of MSE calculated on the validation dataset are chosen.

Appendix B-Definition of the NN Architecture

In general, NNs with at least one hidden layer are universal approximators of any continuous function, as long as the transfer function of the nodes in the hidden layers is continuous, limited, and not-constant, with pure linear activation function in the output layer and with enough hidden units available [50]. However, although one hidden layer is enough to approximate any continuous function, NNs with two hidden layers generally produce better results [53]. Traditional NNs with two hidden layers are chosen to keep the complexity of the proposed architecture as low as feasible. To define the number of nodes in the hidden layers, an iterative trial-and-error approach was adopted. Starting with only one node for both hidden layers, for each iteration, two NNs were configured by adding one node separately both on the first and on the second hidden layer. The two NNs were subsequently trained and compared, finally choosing the one that produced the least MSE. The iterative procedure continues with the progressive addition of nodes and ends when the gradient of the three-point moving averages of MSE calculated on the validation dataset increases ten times consecutively. The final configuration is chosen in correspondence to the minimum of the MSE calculated on the validation dataset with a tolerance of 1%, in the minimum input number direction. To evaluate the contribution of the transfer functions to the NN architecture, the three most common transfer functions were considered.

1) Hyperbolic tangent sigmoid function $f(x) = (2)/(1 + e^{-2n}) - 1$.
2) Logarithmic sigmoid function $f(x) = (1)/(1 + e^{-n})$.
3) Linear function $f(x) = x$.

The above-described iterative procedure was performed for each of the 12 combinations of the transfer functions, obtained by using the first two functions for the two hidden layers and all the three functions for the output layer. The combination of the transfer functions that produce the minimum MSE calculated on the validation dataset was finally selected. For IASI, the best architecture consists of 23 nodes on the firsthidden layer and eight nodes in the second hidden layer, with the transfer functions [logarithmic sigmoid—logarithmic sigmoid—logarithmic sigmoid]. For IASI-NG, the best architecture consists of ten nodes on the first hidden layer and four

nodes in the second hidden layer, with the transfer functions [tangent sigmoid—logarithmic sigmoid—logarithmic sigmoid] for [first hidden layer—second hidden layer—output layer].

Appendix C-Selection of NN Inputs

To reduce the risk of overfitting of the NNs, an iterative input removal procedure was adopted, which starts from the less important ones and ends until the desired degree of tolerance is reached. In this procedure, at the first iteration, all the inputs are removed, one at a time, and the relative weight redistributed among the remaining inputs by minimizing the MSE by using the approach described in [54]. The NN corresponding to the input whose removal returns the minimum MSE calculated on the training dataset is selected and its weights and bias are updated, by using the Levenberg–Marquardt algorithm and the same two early stopping rules introduced above. The procedure is then reiterated starting from the number of inputs obtained in the previous iteration until only one input remains. The final subset of inputs was chosen by selecting those corresponding to the minimum MSE calculated on the validation dataset, with a 1% tolerance, in the direction of the minimum input number.

References

[1] K. Liou, "Influence of cirrus clouds on weather and climate processes: A global perspective," *Monthly Weather Rev.*, vol. 114, no. 6, pp. 1167–1199, Jun. 1986, doi: 10.1175/1520-0493(1986)114<1167:IOCCOW>2.0.CO;2.

[2] T. Haladay and G. Stephens, "Characteristics of tropical thin cirrus clouds deduced from joint CloudSat and CALIPSO observations," *J. Geophys. Res.*, vol. 114, Apr. 2009, doi: 10.1029/2008JD010675.

[3] J. E. Russell, J. D. Haigh, "Detecting thin cirrus clouds in high-spectral-resolution infrared data," *Proc. SPIE*, vol. 4168, Jan. 2001, Art. no. 413844, doi: 10.1117/12.413844.

[4] B.-C. Gao, A. F. H. Goetz, and W. J. Wiscombe, "Cirrus cloud detection from airborne imaging spectrometer data using the 1.38 μm water vapor band," *Geophys. Res. Lett.*, vol. 20, no. 4, pp. 301–304, Feb. 1993, doi: 10.1029/93GL00106.

[5] K. D. Hutchison, K. R. Hardy, and B.-C. Gao, "Improved detection of optically thin cirrus clouds in nighttime multispectral Meteorological Satellite imagery using total integrated water vapor information," *J. Appl. Meteorol.*, vol. 34, no. 5, pp. 1161–1168, May 1995.

[6] W. P. Menzel, D. P. Wylie, and K. I. Strabala, "Seasonal and diurnal changes in cirrus clouds as seen in four years of observations with the VAS," *J. Appl. Meteorol.*, vol. 31, no. 4, pp. 370–385, Apr. 1992.

[7] T. McHardy et al., "GOES ABI detection of thin cirrus over land," *J. Atmos. Ocean. Technol.*, vol. 39, no. 9, pp. 1415–1429, Mar. 2022. [Online]. Available: https://journals.ametsoc.org/view/journals/atot/39/9/JTECH-D-210160.1.xml

[8] R. L. Bankert, "Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network," *J. Appl. Meteorol.*, vol. 33, no. 8, pp. 909–918, Aug. 1994.

[9] R. L. Bankert, C. Mitrescu, S. D. Miller, and R. H. Wade, "Comparison of GOES cloud classification algorithms employing explicit and implicit physics," *J. Appl. Meteorol. Climatol.*, vol. 48, no. 7, pp. 1411–1421, Jul. 2009.

[10] J. Strandgren, L. Bugliaro, F. Sehnke, and L. Schröder, "Cirrus cloud retrieval with MSG/SEVIRI using artificial neural networks," *Atmos. Meas. Techn.*, vol. 10, no. 9, pp. 3547–3573, Sep. 2017, doi: 10.5194/amt-10-3547-2017.

[11] T. Maestri, W. Cossich, and I. Sbrolli, "Cloud identification and classification from high spectral resolution data in the far infrared and mid-infrared," *Atmos. Meas. Techn.*, vol. 12, no. 7, pp. 3521–3540, Jul. 2019, doi: 10.5194/AMT-12-3521-2019.

[12] L. Palchetti, G. Di Natale, and G. Bianchini, "Remote sensing of cirrus cloud microphysical properties using spectral measurements over the full range of their thermal emission," *J. Geophys. Res. Atmos.*, vol. 121, pp. 10804–10819, Sep. 2016, doi: 10.1002/2016JD025162.

[13] Q. Yue and K. N. Liou, "Cirrus cloud optical and microphysical properties determined from AIRS infrared spectra," *Geophys. Res. Lett.*, vol. 36, no. 5, 2009, Art. no. L05810, doi: 10.1029/2008GL036502.

[14] F. Romano, D. Cimini, R. Rizzi, and V. Cuomo, "Multilayered cloud parameters retrievals from combined infrared and microwave satellite observations," *J. Geophys. Res.*, vol. 112, no. D8, 2007, Art. no. D08210, doi: 10.1029/2006JD007745.

[15] Q. Liu et al., "Hyperspectral infrared sounder cloud detection using deep neural network model," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3023683.

[16] J. Lee, P. Yang, A. E. Dessler, B.-C. Gao, and S. Platnick, "Distribution and radiative forcing of tropical thin cirrus clouds," *J. Atmos. Sci.*, vol. 66, no. 12, pp. 3721–3731, Dec. 2009, doi: 10.1175/2009JAS3183.1.

[17] A. E. Dessler and P. Yang, "The distribution of tropical thin cirrus clouds inferred from *Terra MODIS* data," *J. Climate*, vol. 16, no. 8, pp. 1241–1247, Apr. 2003, doi: 10.1175/1520-0442(2003)16<1241:TDOTTC>2.0.CO;2.

[18] H. Hersbach et al., "The ERA5 global reanalysis," *Quart. J. Roy. Meteorological Soc.*, vol. 146, no. 730, pp. 1999–2049, Jul. 2020, doi: 10.1002/QJ.3803.

[19] U. Amato, G. Masiello, C. Serio, M. Viggiano, "The σ-IASI code for the calculation of infrared atmospheric radiance and its derivatives," *Environ. Model. Softw.*, vol. 17, pp. 651–667, Nov. 2002, doi: 10.1016/S1364-8152(02)00027-0.

[20] D. Cimini et al., "Spectrum synergy for investigating cloud microphysics," *Bull. Amer. Meteorological Soc.*, vol. 104, no. 3, pp. 606–622, Mar. 2023, doi: https://doi.org/10.1175/BAMS-D-22-0008.1.

[21] C. Crevoisier et al., "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. Techn.*, vol. 7, no. 12, pp. 4367–4385, Dec. 2014, doi: 10.5194/AMT-7-4367-2014.

[22] F. Vittorioso, V. Guidard, and N. Fourrié, "An infrared atmospheric sounding interferometer–new generation (IASI-NG) channel selection for numerical weather prediction," *Quart. J. Roy. Meteorological Soc.*, vol. 147, no. 739, pp. 3297–3317, Jul. 2021, doi: 10.1002/QJ.4129.

[23] C. Serio et al., "Infrared atmospheric sounder interferometer radiometric noise assessment from spectral residuals," *Appl Opt.*, vol. 54, no. 19, pp. 5924–5936, Jun. 2015, doi: 10.1364/AO.54.005924.

[24] P. Mastro et al., "Combined IASI-NG and MWS observations for the retrieval of cloud liquid and ice water path: A deep learning artificial intelligence approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3313–3322, Apr. 2022, doi: 10.1109/JSTARS.2022.3166992.

[25] K. Wyser, "The effective radius in ice clouds," *J. Climate*, vol. 11, no. 7, pp. 1793–1802, Jul. 1998, doi: 10.1175/1520-0442(1998)011<1793:TERIIC>2.0.CO;2.

[26] D. L. Mitchell and W. P. Arnott, "A model predicting the evolution of ice particle size spectra and radiative properties of cirrus clouds. Part II: Dependence of absorption and extinction on ice crystal morphology," *J. Atmos. Sci.*, vol. 51, fasc. 6, pp. 817–832, Mar. 1994, doi: 10.1175/1520-0469(1994)051<0817: AMPTEO>2.0.CO;2.

[27] A. J. Heymsfield and C. M. R. Platt, "A parameterization of the particle size spectrum of ice clouds in terms of the ambient temperature and the ice water content," *J. Atmos. Sci.*, vol. 41, no. 5, pp. 846–855, Mar. 1984, doi: 10.1175/15200469(1984)041<0846:APOTPS>2.0.CO;2.

[28] G. G. Mace, S. Benson, and E. Vernon, "Cirrus clouds and the large-scale atmospheric state: Relationships revealed by six years of ground-based data," *J. Climate*, vol. 19, no. 13, pp. 3257–3278, Jul. 2006.

[29] K. Sassen, Z. Wang, and D. Liu, "Global distribution of cirrus clouds from CloudSat/cloud-aerosol LiDAR and infrared pathfinder satellite observations (CALIPSO) measurements," *J. Geophys. Res.*, vol. 113, Oct. 2008, Art. no. D00A12, doi: 10.1029/2008JD009972.

[30] K. Sassen and J. R. Campbell, "A midlatitude cirrus cloud climatology from the facility for atmospheric remote Sensing. Part I: Macrophysical and synoptic properties," *J. Atmos. Sci.*, vol. 58, no. 5, pp. 481–496, Mar. 2001.

[31] G. G. Mace et al., "A description of hydrometeor layer occurrence statistics derived from the first year of merged CloudSat and CALIPSO data," *J. Geophys. Res.*, vol. 114, Apr. 2009, Art. no. D00A26, doi: 10.1029/2007JD009755.

[32] G. G. Mace and Q. Zhang, "The CloudSat radar-LiDAR geometrical profile product (RL-GeoProf): Updates, improvements, and selected results," *J. Geophys. Res. Atmos.*, vol. 119, pp. 9441–9462, Jun. 2014, doi: 10.1002/.2013JD021374.

[33] M. Deng, G. G. Mace, Z. Wang, and E. Berry, "CloudSat 2C-ICE product update with a new Ze parameterization in LiDAR-only region," *J. Geophys. Res, Atmos.*, vol. 120, no. 23, pp. 12–198, Dec. 2015, doi: 10.1002/2015JD023600.

[34] *AVHRR Level 1B Product Guide*. Accessed: Apr. 8, 2022. [Online]. Available: https://www.eumetsat.int/media/15351

[35] S. A. Ackerman, R. E. Holz, R. Frey, E. W. Eloranta, B. C. Maddux, and M. McGill, "Cloud detection with MODIS. Part II: Validation," *J. Atmos. Ocean. Technol.*, vol. 25, no. 7, pp. 1073–1086, Jul. 2008, doi: 10.1175/2007JTECHA1053.1.

[36] R. A. Frey et al., "Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5," *J. Atmos. Ocean. Technol.*, vol. 25, no. 7, pp. 1057–1072, Jul. 2008, doi: 10.1175/2008JTECHA1052.1.

[37] R. W. Saunders and K. T. Kriebel, "An improved method for detecting clear sky and cloudy radiances from AVHRR data," *Int. J. Remote Sens.*, vol. 9, no. 1, pp. 123–150, Jan. 1988, doi: 10.1080/01431168808954841.

[38] C. Cao, F. J. De Luccia, X. Xiong, R. Wolfe, and F. Weng, "Early on-orbit performance of the visible infrared imaging radiometer suite onboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1142–1156, Feb. 2014, doi: 10.1109/TGRS.2013.2247768.

[39] R. Godin, "Joint Polar Satellite System (JPSS) VIIRS cloud effective particle size and cloud optical thickness algorithm theoretical basis document (ATBD)," JPSS ATBD VIIRS Cloud EPS COT, NOAA, Goddard Space Flight Center, Greenbelt, MD, USA, May 2014, p. 474. Accessed: Sep. 21, 2021. [Online]. Available: https://www.star.nesdis.noaa.gov/smcd/emb/viirs_aerosol/documents/D0001-M01-S01-020_JPSS_ATBD_VIIRS-AOT-APSP_B.pdf

[40] T. Nakajima and M. D. King, "Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory," *J. Atmos. Sci.*, vol. 47, pp. 1878–1893, Aug. 1990, doi: 10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO.2.

[41] S. C. Ou, K. N. Liou, W. M. Gooch, and Y. Takano, "Remote sensing of cirrus cloud parameters using advanced very-high-resolution radiometer 3.7- and 10.9-μm channels," *Appl. Opt.*, vol. 32, pp. 2171–2180, Apr. 1993, doi: 10.1364/AO.32.002171.

[42] J. Schmetz et al., "An introduction to Meteosat second generation (MSG)," *Bull. Amer. Meteorol. Soc.*, vol. 83, pp. 977–992, Jul. 2002.

[43] *MTG-FCI: ATBD for Optimal Cloud Analysis Product*. Accessed: Apr. 12, 2022. [Online]. Available: https://www-cdn.eumetsat.int/files/2020-06/pdf_mtg_atbd_oca.pdf

[44] I. Gabur, D. P. Simioniuc, R. J. Snowdon, and D. Cristea, "Machine learning applied to the search for nonlinear features in breeding populations," *Frontiers Artif. Intell.*, vol. 5, May 2022, Art. no. 876578, doi: 10.3389/frai.2022.876578.

[45] P. Mastro et al., "On the synergic use of satellite microwave and infrared measurements for the estimation of effective radius of ice and liquid water clouds: A regression approach based on random forests," *Proc. SPIE*, vol. 12265, pp. 127–136, Oct. 2022.

[46] H. Brooks et al., *Forecast Verification Issues, Methods and FAQ*. Accessed: Mar. 24, 2022. [Online]. Available: http://www.cawcr.gov.au/projects/verification/

[47] I. Farouk, N. Fourrié, and V. Guidard, "Homogeneity criteria from AVHRR information within IASI pixels in a numerical weather prediction context," *Atmos. Meas. Techn.*, vol. 12, no. 6, pp. 3001–3017, Jun. 2019, doi: 10.5194/amt-12-3001-2019.

[48] P. Martinet, N. Fourrié, V. Guidard, F. Rabier, T. Montmerle, and P. Brunel, "Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances," *Quart. J. Roy. Meteorological Soc.*, vol. 139, no. 674, pp. 1402–1416, Nov. 2012, doi: 10.1002/qj.2046.

[49] L. Zhou, M. Divakarla, X. Liu, A. Layns, and M. Goldberg, "An overview of the science performances and calibration/validation of Joint Polar Satellite System operational products," *Remote Sens.*, vol. 11, no. 6, p. 698, Mar. 2019.

[50] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989, doi: 10.1016/0893-6080(89)90020-8.

[51] K. Levenberg, "A method for the solution of certain problems in least squares," *Quart. Appl. Math.*, vol. 2, pp. 164–168, Jul. 1944. [Online]. Available: https://www.jstor.org/stable/43633451

[52] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, Jun. 1963. [Online]. Available: https://www.jstor.org/stable/2098941

[53] A. J. Thomas, M. Petridis, S. D. Walters, S. M. Gheytassi, R. E. Morgan, "Two hidden layers are usually better than one," *Proc. Int. Conf. Eng. Appl. Neural Netw.*, 2017, pp. 279–290.

[54] G. Castellano and A. M. Fanelli, "Variable selection using neural-network models," *Neurocomputing*, vol. 31, nos. 1–4, pp. 1–13, Mar. 1999, doi: 10.1016/S0925-2312(99)00146-0.

Elisabetta Ricciardelli received the M.S. degree in physics from the University of Bologna, Bologna, Italy, in 1998, and the Ph.D. degree in methods and technologies for environmental monitoring from the University of Basilicata, Potenza, Italy, in 2008.

In 2005, she was with the Institute of Methodologies for Environmental Analysis, Tito Scalo, Italy. Her research interests include satellite-data handling for meteorological studies with a particular focus on cloud detection and classification.

Francesco Di Paola received the M.S. degree in physics from Sapienza University, Rome, Italy, in 2004, and the Ph.D. degree in geophysics from University Alma Mater Studiorum, Bologna, Italy, in 2008.

Since 2005, he has been a Researcher with the Italian National Research Council, Tito Scalo, Italy, and in 2011, he was with the Institute of Methodologies for Environmental Analysis, Tito Scalo. His research interests include rainfall estimation and nowcasting with multisensor approach combining microwave (MW)-infrared (IR) satellite data with lightning data from surface networks.

Domenico Cimini received the Laurea (cum laude) and Ph.D. degrees in physics from the University of L'Aquila, L'Aquila, Italy, in 1998 and 2002, respectively.

He is currently a Research Manager with the Institute of Methodologies for Environmental Analysis, Italian National Research Council, Tito Scalo, Italy. He has been a Research Assistant with the Cooperative Institute for Research in Environmental Sciences (CIRES), Boulder, CO, USA, and an Adjunct Professor with the Department of Electrical and Computer Engineering, University of Colorado, Boulder. He has more than 20 years of experience with ground- and satellite-based passive remote sensing, particularly microwave radiometry.

Dr. Cimini is a Life Member of the European Geosciences Union. He was a recipient of the 6th Hans Liebe Lectureship bestowed by the U.S. National Committee for the Union of Radio Scientists Internationale in 2019.

Salvatore Larosa received the Ph.D. degree in environmental science and engineering, construction and energy from the University of Calabria, Rende, Italy, in 2020, and the M.S. degree in geology from the University of Calabria, Rende, Italy, in 2002.

An experience of over ten years in the field of geographical information systems has provided him the ideal basis to face and manage the work activities concerning Geomatics. Since 2013, he has been a QGIS Core Developer. His research interests include Earth observations from space/remote sensing, meteorology, atmospheric physics and dynamics, storing and managing remotely sensed data in multidimensional databases, collecting data and analyzing for trends, using statistical techniques and programming software to create algorithms and predictive models, creating tools for data interoperability, sharing and visualizations, data management systems, information retrieval and digital libraries, data fusion, and machine learning.

Pietro Mastro (Graduate Student Member, IEEE) received the M.S. degree (summa cum laude) in computer engineering and information technology and the Ph.D. degree in engineering for innovation and sustainable development from the University of Basilicata, Potenza, Italy, in 2019 and 2023, respectively.

Since 2022, he has held a post-doctoral research grant at IREA-CNR within the following research line: "Development of algorithms for processing and analysis of synthetic aperture RADAR signals for environmental remote sensing." His primary research focuses on remote sensing with active sensors, specifically developing advanced DInSAR algorithms for monitoring surface deformation induced by subsidence, volcanoes, and earthquakes. He is particularly interested in noise filtering problems and the development of advanced methodologies based on machine learning for SAR data processing and change detection (CD) analysis. His research activities also extend to the field of remote sensing with passive sensors, specifically in the domain of Fourier spectroscopy, for studying Earth's atmosphere and land processes.

Dr. Mastro was a recipient of the GRS29-CNI 2019 Best Master's Degree Thesis in Geoscience and Remote Sensing Award in Rome, Italy, of the Best Poster Award Presentation in Solid Earth and Disaster Risk Reduction Terrain Motion and Urban/Infrastructure Assessment at ESA-MOST, DRAGON 4 2019 Symposium in Lubiana (Slovenia), in Coastal Zones and Oceans at ESA-MOST DRAGON 5 2022 MID-TERM RESULTS SYMPOSIUM, and the M.Sc Student Award financed by ESA within the ESA-CNR project entitled "INtegrated analysis of the combined Risk of groUND subsidence sea level rise and nAtural hazards in coasTal delta rEgions (INUNDATE)," Public IREA-CNR Competition n.126.273.PL.001/2017. In 2023, he was one of the recipients of the CNR Short Term Mobility (STM) Program.

Guido Masiello teaches at the School of Engineering, University of Basilicata, Potenza, Italy, where he is currently an Associate Professor of Earth and atmospheric physics. He works on applied spectroscopy topics, developing methods and instruments for the retrieval of the thermodynamic state of the Earth's atmosphere. He is a co-investigator in various projects in the area of high spectral resolution infrared sounders from satellites (far and thermal, from 100 to 3.5 μm): the infrared atmospheric sounder interferometer (IASI of the French Space Agency), the Centre National d'Études Spatiales (CNES), and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), the Far-infrared Outgoing Radiation Understanding and Monitoring (FORUM, the 9th ESA Earth Explorer Mission), the radiation explorer in the far infrared (REFIR, a three years EU project supported within the 4th framework program), the Meteosat Third Generation Infrared Sounder (MTG-IRS, a joint program ESA/EUMETSAT), and the next generation of IASI (IASI-NG of CNES and EUMETSAT).

Dr. Masiello was a reviewer for several journals, including *International Journal of Optics* and *Journal of the Atmospheric Sciences*, and for Italian and international research funding agencies.

Carmine Serio was born in July 1954. He received the M.S. degree in physics from the University of Naples, Naples, Italy, in 1978.

He is currently a Full Professor in Earth physics. He is the Director of the School of Engineering, University of Basilicata, Potenza, Italy, where he teaches environmental and atmospheric physics courses. Since 1990, he has been performing research in the field of Fourier spectroscopy, methods, and experimental techniques, applied to the study of environment, Earth atmosphere, land processes, and remote sensing of atmospheric and surface parameters. His experience with Fourier transform infrared spectrometers, or FTIR instruments, includes ground-based, airborne, and satellite platforms. He currently develops cutting-edge forward/inverse models for hyperspectral infrared observations. In addition, he acted as a Principal Investigator in various projects in high spectral resolution infrared sounders from satellites. These projects include the interferometric monitoring for greenhouse gases (IMG of the Japanese NASDA), the infrared atmospheric sounder interferometer (IASI of the French Space Agency, the Centre National d'Études Spatiales (CNES), and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)], the radiation explorer in the far-infrared (REFIR, a three years EU project supported within the 4th framework program), the Meteosat Third Generation Infrared Sounder (MTG-IRS, a joint program ESA/EUMETSAT), and IASI-next-generation (IASI-NG of CNES and EUMETSAT).

Mr. Serio is a member of the core science team of the Far-infrared-Outgoing-Radiation Understanding and Monitoring mission. The mission was selected in September 2019 by the ESA board Committee to fly in 2026 as part of the Earth Explorer Mission-9, Earth Observation Envelope Programme. Because of his expertise, he has acted as a reviewer of satellite experiments and projects.

Tim Hultberg received the M.Sc. degree in computer science and mathematics from the University of Copenhagen, Copenhagen, Denmark, in 1991, and the Ph.D. degree in mathematical modeling from the Technical University of Denmark, Kongens Lyngby, Denmark, in 2001, where he developed the open-source algebraic modeling language FLOPC++.

Since 2003, he has been a remote sensing expert at the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany, working with compression and retrieval from satellite data—especially hyperspectral infrared measurements. His research interests include inverse problems and machine learning for remote sensing.

Thomas August received the M.Sc. degree in astrophysics and planetology from the University of Toulouse, Toulouse, France, in 1997, and the Ph.D. degree in astrophysics from the University of Bordeaux, Talence, France, in 2002, for a research work financed by the Centre National Etudes Spatiales (CNES) on the potential of synthetic aperture radars for the remote exploration of subsurface.

Since 2003, he has been with the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany, where he contributed to the prototyping, development, and validation of the infrared atmospheric sounding interferometer (IASI) Level 2 (L2) products processor. In 2010, he resumed responsibilities as a Mission Scientist for IASI and IASI-NG, and since 2014, he has been leading hyperspectral Level 2 product developments and applications studies, including MTG-IRS. Since 2022, he has been resuming responsibilities as a METimage scientist and coordinating cross-mission activities on water-vapor products. His research interests include the atmospheric sounding and retrieval techniques for temperature and humidity profiles, the surface temperature and emissivity, the atmospheric composition, and the cloud detection as well as their applications in weather forecasting.

Filomena Romano received the M.S. degree in physics from the University of Bologna, Bologna, Italy, in 1990.

In 1992, she joined the Institute of Methodologies for Environmental Analysis, Tito Scalo, Italy, where she got experience on experimental and theoretical studies in atmospheric remote sensing. She collaborated in studies concerning retrieval of atmospheric aerosol from solar spectra at ground level. She has currently specialized in satellite data handling for meteorological and climatological studies. Her main research interests include cloud detection, cloud clearing, and cloud microphysical retrieval of infrared and microwave radiance from spaceborne sensors.

Open Access funding provided by 'Consiglio Nazionale delle Ricerche-CARI-CARE-ITALY'
within the CRUI CARE Agreement