Data Fusion Analysis of Sentinel-4 and Sentinel-5 Simulated Ozone Data

CECILIA TIRELLI, SIMONE CECCHERINI, NICOLA ZOPPETTI, SAMUELE DEL BIANCO, MARCO GAI, FLAVIO BARBARA, AND UGO CORTESI

Istituto di Fisica Applicata Nello Carrara del Consiglio Nazionale delle Ricerche, Florence, Italy

JUKKA KUJANPÄÄ

Finnish Meteorological Institute, Helsinki, Finland

YU HUAN

Royal Belgian Institute for Space Aeronomy, Brussels, Belgium

ROSSANA DRAGANI

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

(Manuscript received 29 April 2019, in final form 4 February 2020)

ABSTRACT

The complete data fusion method, generalized to the case of fusing profiles of atmospheric variables retrieved on different vertical grids and referred to different true values, is applied to ozone profiles retrieved from simulated measurements in the ultraviolet, visible, and thermal infrared spectral ranges for the Sentinel-4 and Sentinel-5 missions of the Copernicus program. In this study, the production and characterization of combined low Earth orbit (Sentinel-5) and geostationary Earth orbit (Sentinel-4) fused ozone data is performed. Fused and standard products have been compared and a performance assessment of the generalized complete data fusion is presented. The analysis of the output products of the complete data fusion algorithm and of the standard processing using quality quantifiers demonstrates that the generalized complete data fusion algorithm provides products of better quality when compared with standard products.

1. Introduction

Global and continuous measurements of ozone vertical profile are essential to monitor the evolution of the atmospheric ozone from the surface up to the mesosphere. Instruments developed over the last decades to monitor ozone from space exploit a variety of observation geometries and spectral regions (Quesada-Ruiz et al. 2020; Heue et al. 2016; Hassler et al. 2014; Nirala 2008); however, due to the inherent limitations of each measurement technique, none of the existing systems is able to provide ozone observations that cover the entire vertical profile from the surface up to the top of atmosphere. The advantages of a multispectral approach for observing ozone profiles from space have been demonstrated by using simulated data (Landgraf and Hasekamp 2007; Worden et al. 2007; Natraj et al. 2011; Hache et al. 2014; Costantino et al. 2017) and real measurements (Fu et al. 2013; Cuesta et al. 2013). Moreover, two review papers on this subject are Lahoz et al. (2012) and Timmermans et al. (2015). In the next decades, the instruments aboard Copernicus atmospheric Sentinel missions, that is, Sentinel-5 Precursor (S5P), Sentinel-4 (S4) and Sentinel-5 (S5) (ESA 2016, 2017, 2000), can be exploited to monitor the profile of ozone concentration in Earth's atmosphere with unprecedented accuracy and timeliness. Although not included as part of the operational processing of the atmospheric Sentinels measurements, synergistic approaches to data analysis preserve high priority in the investigation of scientific and technological advancements required to achieve the upgrading from research to operational algorithms. In this framework, the development

Denotes content that is immediately available upon publication as open access.

Corresponding author: Cecilia Tirelli, c.tirelli@ifac.cnr.it

DOI: 10.1175/JTECH-D-19-0063.1

of innovative techniques to better exploit the synergy between ozone measurements covering a wide range of spectral regions is crucial to reduce the quantity of data and improve their quality in terms of improved accuracy and vertical resolution. The data fusion approach is ideal for this purpose. In this technique the observations of the different instruments are used to retrieve from each one an independent vertical profile and, a posteriori, an algorithm is implemented to combine into a single estimate the profiles retrieved from the observations acquired by the different instruments. Advanced Ultraviolet Radiation and Ozone Retrieval for Applications (AURORA) is a 3-yr project supported by the European Union in the frame of its Horizon 2020 Call (EO-2-2015) for "Stimulating wider research use of Copernicus Sentinel Data" (Cortesi et al. 2018). The primary goal of the project is to exploit the complementary measurement capabilities of the instruments on board the S4 and S5 missions, operating on sun-synchronous polar low Earth orbit (LEO) and on geostationary orbit (GEO), respectively, for near-real-time monitoring of the ozone vertical profile with unprecedented accuracy. Within the AURORA project, the complete data fusion (CDF) algorithm (Ceccherini et al. 2015), generalized to the case of fusing profiles retrieved on different vertical grids and referred to different true profiles (Ceccherini et al. 2018), is used to combine the information associated to the operational products of the LEO instruments, as well as to the ones on the GEO mission. The fused ozone profiles resulting from this first step will be subsequently merged into assimilation models, to integrate the combined products from LEO and GEO measurements in a short-term ozone-forecasting model. This paper provides a description of the implementation of the generalized CDF to the S4 and S5 simulated ozone datasets, as well as a quality assessment of the fused products also compared to the standard ones. The paper is structured as follows. Section 2 describes the simulation activity and the definition and implementation of the coincidence algorithm for the LEO–LEO, GEO–GEO and LEO–GEO observations. Section 3 shows the implementation and improvement of the CDF method. Section 4 describes the production and characterization of the fused data (GEO–GEO, LEO–LEO and LEO–GEO analysis) with assessment of the fused data quality. Conclusions are drawn in section 5.

2. Simulation and coincidence criteria

a. Simulation

This work was carried out before the launch of the atmospheric Sentinels; thus, synthetic level 2 (L2; i.e., the geophysical products retrieved from the measured radiances) Sentinels data in the thermal infrared (TIR), visible (VIS), and ultraviolet (UV) spectral ranges are provided by simulators developed within the AURORA project.

The S5P mission was launched on 13 October 2017. It carries the Tropospheric Ozone Monitoring Instrument (TROPOMI) with observing capability spanning the ultraviolet to shortwave infrared spectral band. At present, the S5P data are not included in the archive of synthetic data of the AURORA project. The S4 mission consists of a ultraviolet–visible–near-infrared (UVN) imaging spectrometer and will rely on the utilization of subsets of data from EUMETSAT's thermal Infrared Sounder (IRS), both embarked on EUMETSAT's geostationary MTG-S platforms. The instrument of the S5 mission is an ultraviolet–visible–near-infrared shortwave (UVNS) imaging spectrometer and the mission will rely on data from Infrared Atmospheric Sounding Interferometer Next Generation (IASI-NG), both on board EUMETSAT's MetOp Second Generation (SG).

The spectral bands and the available products specifications used in this study for the simulation of the operational and nonoperational ozone products of the S4 and S5 missions are summarized in Table 1. Simulations of standard ozone L2 data were carried out considering the ozone products requirements for Sentinels missions instruments: information are extracted from the S4 and S5 mission requirements document (MRD) (ESA Mission Science Division 2007) and mission requirements traceability document (MRTD) (ESA Mission Science Division 2017) and from Ingmann et al. (2012). Information about the infrared sounders products specifications were extracted from the Post–EUMETSAT Polar System (EPS) MRD (EUMETSAT 2010) for IRS and from Crevoisier et al. (2014) for IASI-NG. Operational ozone L2 data of S4 and S5 missions are derived from TIR and UV bands, but do not include ozone retrieval products from measurements acquired in the visible band. However, in this work the simulation of S4 and S5 measurements in the visible band (from 425 to 497 nm) have been used to retrieve the ozone total column to be fused with ozone profiles from the UV and TIR spectral regions.

Simulations of various ozone total columns and profiles are carried out in different spectral ranges for selected atmospheric scenarios defining the state of the atmosphere and providing information on meteorology, atmospheric composition and surface albedo. The MERRA-2 reanalysis (Gelaro et al. 2017) was selected as the most complete data source for the required

TABLE 1. Sentinel-4 and Sentinel-5 missions—products specification.

Mission	Instrument			Utilization of data from[a]		
Sentinel-4 (GEO)	UVN Spectrometer on board MTG Sounder Geo coverage: 30°W–45°E [at 40°N], 30°–65°N Pixel: 8 km at 45°N 0° lon Revisit time: 60 min			IRS on board MTG Sounder Geo coverage: Four LAC zones (local area coverage)[b] Pixel: 4 km nadir Revisit time: 60 min (full disc)		
	Band	Product	Uncertainty	Band	Product	Radiometric resolution
	305–330 nm (UV1)	Tropospheric O_3	10% PBL column, 25% tropospheric column	1030–1080 cm^{-1} (IRS-3a)	O_3 vertical profile	NEΔT at 280 K: 0.25/0.14/0.07 threshold(T)/breakthrough(B)/ objective(O)
	325–337 nm (UV2)	Total O_3	3%–5%			
	425–497 nm (VIS)	O_3 column	Not available			
Sentinel-5 (LEO)	UVNS Spectrometer on board MetOp-SG Geo Coverage: global Pixel: 15 km(T), 5 km(B) Revisit time: daily (no night)			IASI-NG on board MetOp-SG Geo Coverage: global Pixel: 12 km nadir Revisit time: daily		
	Band	Product	Uncertainty	Band	Product	Radiometric resolution
	270–330 nm (UV1/UV2-VIS)	O_3 vertical profile	10% PBL column, 25% tropospheric column	1030–1080 cm^{-1} (IAS-4)	O_3 vertical profile	NEΔT at 280 K: 0.195
	325–337 nm (UV1/UV2-VIS)	Total O_3	3%–5%			
	425–497 nm (UV2-VIS)	O_3 column	Not available			

[a] https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-4.
[b] https://www.eumetsat.int/website/home/Satellites/FutureSatellites/MeteosatThirdGeneration/MTGDesign/index.html#irs.

fields. In addition, the ozone climatology of McPeters and Labow (2012) was selected as a priori in the different retrieval algorithms. The Sentinel-4 instrument will monitor Earth's radiance within the so-called geographic coverage area (GCA), which covers Europe, parts of North Africa, and parts of the Atlantic from 30° to 65°N in latitude and from 30°W to 45°E in longitude. The UVN instrument has an instantaneous field of regard of 4.0° that covers the north–south range of the GCA. For the east–west range a scan mirror is used, that will scan continuously from east to west over a range of about ±4.5° with a fixed scan duration of 60 min. The size of the simulated pixels in each scan line is $8 \times 8\,km^2$ for UV, $9 \times 12\,km^2$ for VIS, and $15 \times 15\,km^2$ for TIR. Sentinel-5 will be operating in nadir looking push broom mode from sun synchronous low Earth orbit. The wide across-track field of view (FoV) of 180° will provide a wide swath of about 2670 km on Earth and thus almost globally allows for daily coverage of Earth's surface. The size of the simulated pixels in each scan line is $15 \times 15\,km^2$ for UV, $7 \times 7\,km^2$ for VIS, and $12 \times 12\,km^2$ for TIR. As S4 and S5 missions are still in the preparatory phase, the simulated orbits and pixels are based on specifications obtained from ESA. The geolocations, observation times and observation geometry angles were generated for a period of four months (1 April–31 July 2012) for both the S4 and S5 measurements. In this work only the first week of April is used for the data fusion analysis. In the project, the amount of considered pixels had to be limited because of the maximum number of pixels per day that could be ingested by both the data assimilation systems (DASs), considering the computing resources available. First of all, only clear-sky conditions (defined as the pixels with a cloud fraction ≤1%) were considered. Moreover, for each spectral range of S4, 1 in every 10 scan lines was sampled, of which 1 in every 10 pixels was selected. In the case of S5 different selection criteria were used for the three spectral ranges:

- TIR range: 1 in every 5 scan lines and 1 in every 4 pixels were sampled.
- VIS range: 1 in every 7 scan lines and 1 in every 7 pixels were sampled.
- UV range: All pixels were simulated.

The TIR simulator is based on the line-by-line radiative transfer model (RTM) Kyoto Protocol Informed Management of the Adaptation (KLIMA) (Cortesi et al. 2014) and uses an optimal estimation retrieval approach (Rodgers 2000) to simulate the ozone profiles, covariance matrices (CMs) and averaging kernel matrices (AKMs) required for the assimilation. The simulation in the VIS wavelength range is performed through a spectral fit using a differential optical absorption spectroscopy (DOAS) approach (Platt 1994). For the simulation of Sentinel VIS radiance spectra, the GOME direct fitting (GODFIT) algorithm was used, in its mode for forward calculation of synthetic radiance spectra. This algorithm directly adjusts simulated radiances to measured ones in a relevant fitting window. The RTM at the core of the model is the linearized discrete ordinate radiative transfer (LIDORT) scattering code. The VIS simulator's outputs are the total ozone columns with their associated uncertainty and AKMs. Finally, the outputs of the UV simulator (i.e., ozone profiles, CMs, and AKMs) are derived using the KNMI Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval (DISAMAR) inversion package, based on the optimal estimation approach, and the Layer Based Orders of Scattering (LABOS) algorithm, as radiative transfer model. TIR and UV products have been simulated not performing the retrieval but smoothing the true profile with the AKM and adding a random error consistent with the CM [see Eq. (3.12) in Rodgers 2000]. For the simulations only cloud-free scenes are assumed for both S4 and S5 and the effect of aerosols is ignored.

b. Coincidence criteria

A study to define the coincidence algorithm for observations provided by instruments on GEO and LEO satellite platforms was performed. The aim of this algorithm is to select the sets of simulated ozone measurements to be fused. As mentioned above, in the project the CDF solutions (fused ozone products) are assimilated by state-of-the-art DASs to provide accurate ozone analyses and forecasts. In AURORA, two DASs are used: the ECMWF Integrated Forecasting System (IFS) and the KNMI TM5 Data Assimilation Model (TM5DAM). The IFS is a comprehensive Earth-system model to simulate the atmospheric dynamics and the physical processes that occur in the terrestrial atmosphere. Observations, including those for ozone, are assimilated in 12-hourly time windows with a four-dimensional variational data assimilation scheme (Rabier et al. 2000) formulated in terms of increments (e.g., Courtier et al. 1994). In the AURORA project, the IFS assimilation system will be used in two configurations: the first is based on that running operationally at ECMWF and that also serves as the atmospheric core used for its reanalysis productions; the second one (referred as C-IFS) is also based on the same dynamical and assimilation system but in this case the model has been extended to include atmospheric composition. The TM5DAM is based on the TM5 (Krol et al. 2005; Huijnen et al. 2010), a global chemistry-transport model that simulates the concentrations of atmospheric trace gases including greenhouse gases (GHG),

such as carbon dioxide and methane, chemically active species (e.g., ozone), and aerosols. The study of the coincidence algorithm has to take into account both the characteristics of the simulated data and those of the fusion and assimilation processes. Three data fusion experiments were performed in this study:

- GEO–GEO fusion: TIR, UV and VIS simulated ozone data from the GEO platform (S4) are fused.
- LEO–LEO fusion: TIR, UV and VIS simulated ozone data from the LEO platform (S5) are fused.
- LEO–GEO fusion: TIR, UV and VIS simulated ozone data from both GEO and LEO platform (S4 and S5) are fused.

The coincidence algorithm was designed to guarantee the greatest generality and flexibility, in order to be adapted to the user's main objectives and requirements and it is based on

- an indexing mechanism to assign a unique identifier for each pixel involved in the coincidence selection;
- a coincidence cell defined by latitude, longitude and time thresholds; and
- a coincidence manager to support a query system for pixel mapping, selection and storing.

The spatial and time thresholds for the definition of the coincidence criteria generally depend on the available amount of data and their distribution. The operational products of S4 and S5 are expected to provide very good coverage over the geographical areas they are designed to sample but, since not all the pixels were simulated in the project, these thresholds have to be adapted to deal with the filtering criteria used in AURORA. Moreover, the development of the coincidence algorithm has to take into account the characteristics of the assimilation process and the horizontal resolution of the assimilation grid. The horizontal resolution will be 40 km for IFS, 80 km for C-IFS, and 100 km for TM5. To maximize the information in input to the assimilation, the coincidence grid cell should have size of the same order of the assimilation grid cell. Based on the constraints mentioned above, a fixed grid has been chosen for the determination of the coincidence cells. Two or more products will be considered coincident if they fall in the same spatial coincidence cell and their acquisition times are within a predefined time interval of 1 h. The time dimension is relevant only in case of LEO–GEO coincidences, since LEO–LEO and GEO–GEO neighboring measurements of the same orbit are virtually simultaneous. Cells of various grid size were tested considering values of 0.5°, 0.25°, and 0.125° in latitude and 0.625°, 0.3125°, and 0.156 25° in longitude. These tests determined the number of coincidences on the base of the spatial distribution of all the orbit pixels. Cell sizes were tested using simulated pixels of the first week of April 2012, to quantify the resulting number of coincidences. These tests were carried out taking into account the decimation applied during the pixel selection in preparation of the simulation phase. A significant number of coincidences is guaranteed only by the grid with cell size of 0.5° in latitude and 0.625° in longitude: 67% of the cells hold two or more S5 pixels and 3% of them at least one S4–S5 coincidence. The selection of the cell size results from the compromise between the DAS spatial resolution and the number of coincidences, a key aspect for the data fusion process, and it depends on the selection criteria used to simulate the S4 and S5 pixels for the three spectral ranges (see section 2a).

3. CDF algorithm

The simulated L2 ozone products (profiles or columns with the associated AKMs and CMs) for TIR, VIS, and UV spectral ranges that fall into the coincidence cells are used as input for the generalized CDF algorithm (Ceccherini et al. 2018). The CDF (Ceccherini et al. 2015) is a generalization of the weighted mean in the case of AKMs different from identity matrices and is named complete for its capability to take into consideration all the features of the measurements that are being combined. It is based on the assumption to have N independent and simultaneous measurements of the vertical profile of an atmospheric target referred to the same space–time location. The N state vectors $\hat{\mathbf{x}}_i$ ($i = 1, 2, \ldots, N$) retrieved using the optimal estimation method (Rodgers 2000) are here assumed to provide estimates of the profiles on a common vertical grid. The vectors $\hat{\mathbf{x}}_i$ are characterized by the CMs $\mathbf{S}_i$ and the AKMs $\mathbf{A}_i$ (Ceccherini et al. 2003; Ceccherini and Ridolfi 2010; Rodgers 2000).

The CDF solution for the considered profiles is given by

$$\mathbf{x}_f \equiv \left(\sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1}\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \boldsymbol{\alpha}_i + \mathbf{S}_a^{-1} \mathbf{x}_a\right), \tag{1}$$

where

$$\boldsymbol{\alpha}_i \equiv \hat{\mathbf{x}}_i - (\mathbf{I} - \mathbf{A}_i)\mathbf{x}_{ai}, \tag{2}$$

$\mathbf{x}_{ai}$ is the a priori profile used in the ith retrieval, $\mathbf{I}$ is the identity matrix, and $\mathbf{x}_a$ and $\mathbf{S}_a$ are the a priori profile and

its CM used to constrain the data fusion. The CM of the CDF solution, obtained propagating the errors of $\hat{\mathbf{x}}_i$ into $\mathbf{x}_f$, is given by

$$\mathbf{S}_f = \left(\sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1}\right)^{-1} \sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i \times \left(\sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1}\right)^{-1}, \quad (3)$$

and the AKM obtained taking the derivative of $\mathbf{x}_f$ with respect to the true profile is expressed by the following equation:

$$\mathbf{A}_f = \left(\sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1}\right)^{-1} \sum_{i=1}^{N} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i. \quad (4)$$

The CDF formula requires a summation of terms that have a common vertical grid referred to as the fusion grid. Thus, when the fusing profiles are represented on different vertical grids, a resampling of the AKMs is needed (Calisesi et al. 2005). The resampling is obtained as explained in Ceccherini et al. (2016):

$$\mathbf{A}_i' = \mathbf{A}_i \mathbf{R}_i, \quad (5)$$

where $\mathbf{A}_i$ is the original square matrix (its dimensions are defined by the number of levels of the ith measurement) and $\mathbf{A}_i'$ is the transformed AKM, a rectangular matrix (with the number of columns defined by the number of levels of the fusion grid, the final grid in which the profiles are fused). The $\mathbf{R}_i$ is the generalized inverse matrix of the linear interpolation matrix $\mathbf{H}_i$, which interpolates the profiles obtained on different grids on the fusion grid. The application of the CDF method to vertical profiles obtained with different instruments on different retrieval grids and observing different true profiles was analyzed in Ceccherini et al. (2018). An interpolation error is present when the vertical grids of the fusing profiles differ from the fusion grid and an interpolation of the AKMs is needed. The fusing profiles are, in general, not exactly collocated in space and time, and therefore, they refer to different true profiles; thus, a coincidence error is introduced. The CDF formula was therefore modified and generalized to account for both the interpolation and coincidence errors by replacing $\boldsymbol{\alpha}_i$ with

$$\tilde{\boldsymbol{\alpha}}_i = \boldsymbol{\alpha}_i - \mathbf{A}_i(\mathbf{C}^{(i)} - \mathbf{R}_i \mathbf{C}^{(f)})\mathbf{x}_a, \quad (6)$$

where $\mathbf{C}^{(i)}$ and $\mathbf{C}^{(f)}$ are sampling matrices from a fine grid [including all the levels of the fusion grid (f) and of the N fusing grids (i)] to the grids (i) and to the grid (f), respectively.

Besides, $\mathbf{S}_i$ is replaced with

$$\tilde{\mathbf{S}}_i = \mathbf{S}_i + \mathbf{S}_{i,\mathrm{int}} + \mathbf{S}_{i,\mathrm{coin}}. \quad (7)$$

The interpolation and the coincidence errors are characterized by the CMs $\mathbf{S}_{i,\mathrm{int}}$ and $\mathbf{S}_{i,\mathrm{coin}}$, respectively:

$$\mathbf{S}_{i,\mathrm{int}} = \mathbf{A}_i(\mathbf{C}^{(i)} - \mathbf{R}_i \mathbf{C}^{(f)})\mathbf{S}_a(\mathbf{C}^{(i)} - \mathbf{R}_i \mathbf{C}^{(f)})\mathbf{A}_i^{\mathrm{T}}, \quad (8)$$

$$\mathbf{S}_{i,\mathrm{coin}} = \mathbf{A}_i \mathbf{C}^{(i)} \mathbf{S}_{\mathrm{coin}} \mathbf{C}^{(i)\mathrm{T}} \mathbf{A}_i^{\mathrm{T}}. \quad (9)$$

The $\mathbf{S}_{\mathrm{coin}}$ accounts for the dispersion of the true profiles and, therefore, depends on the coincidence criteria.

CDF with total columns

Since the retrieval of VIS measurements produces a total column and not a profile, the fusion of total columns with profiles is needed when a VIS measurement falls in a coincidence cell. The total column is provided with a CM, which corresponds to the square of the error, and an AKM that consists of a row vector giving the derivative of the retrieved column with respect to the true profile of volume mixing ratio (VMR).

The transformation of a total column in a vertical profile can be done using the CDF formula [Eq. (1)] considering in input only the quantities related to the total column:

$$\mathbf{x} = (\mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1})^{-1} (\mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \boldsymbol{\alpha}_i + \mathbf{S}_a^{-1} \mathbf{x}_a), \quad (10)$$

where $\mathbf{A}_i$ and $\mathbf{S}_i$ are the AKM (a row vector) and the CM (coinciding with the variance) of the column, respectively, and α_i is given by

$$\alpha_i = \hat{c}_i - c_{ai} + \mathbf{A}_i \mathbf{x}_{ai}. \quad (11)$$

In this case, α_i is a scalar quantity, $\hat{c}_i$ is the retrieved total column and c_{ai} is the total column corresponding to the a priori profile $\mathbf{x}_{ai}$.

The CM and the AKM of the profile obtained from the total column are given by

$$\mathbf{S} = (\mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1})^{-1} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i (\mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1})^{-1}, \quad (12)$$

$$\mathbf{A} = (\mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1})^{-1} \mathbf{A}_i^{\mathrm{T}} \mathbf{S}_i^{-1} \mathbf{A}_i. \quad (13)$$

To compare the performance of the fused product with that of the individual measurements, it is useful to transform the information embedded in the columns retrieved from the VIS measurements in vertical profiles. Indeed,

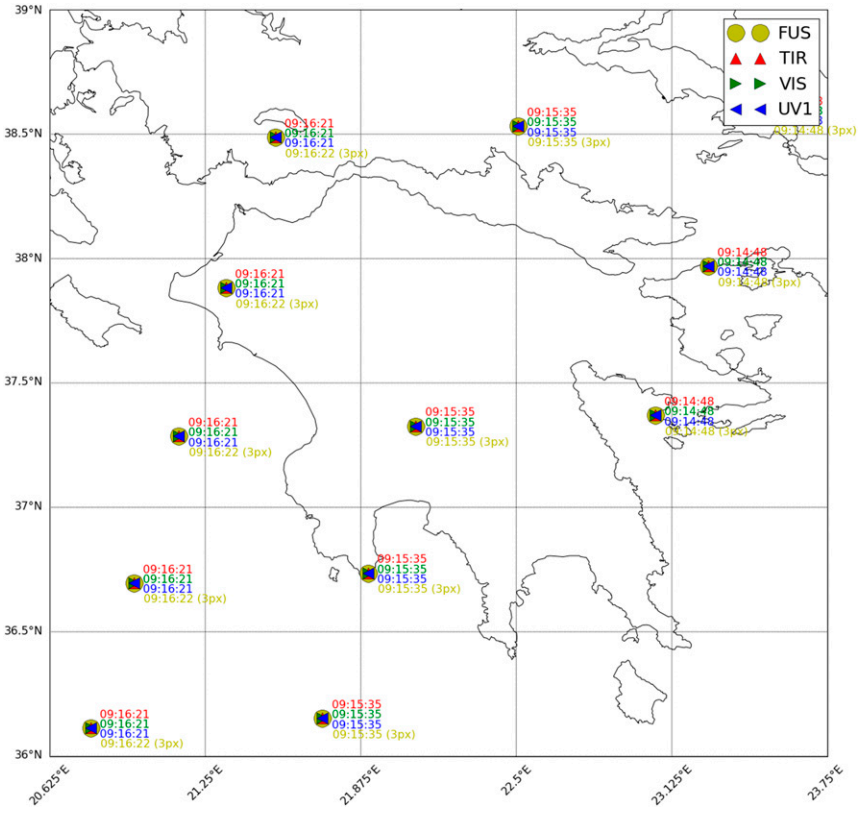

FIG. 1. Example of GEO–GEO coincidences over Greece, the black lines represent the borders of the coincidence grid cells (0.5° in latitude × 0.625° in longitude).

in this way all the products to be compared are vertical profiles and a quality assessment can be done more easily.

4. Production and characterization of the fused data with assessment of the fused data quality

The CDF generalized method has been used for the three data fusion experiments described above: the GEO–GEO, the LEO–LEO, and the LEO–GEO data fusion. For each experiment, we produced the fused data corresponding to the simulated measurements of the first week of April 2012. The quality assessment of the data fusion evaluates three elements:

- The average difference with respect to the true profile of TIR, UV, and VIS retrieved profiles and of the fused profile:

$$\Delta_{i,\mathrm{tot}} = \hat{\mathbf{x}}_i - \mathbf{x}_{\mathrm{true}}, \tag{14}$$

$$\Delta_{f,\mathrm{tot}} = \mathbf{x}_f - \mathbf{x}_{\mathrm{true}}. \tag{15}$$

- The average total errors of the ozone profiles obtained from the TIR, UV, and VIS measurements and from the data fusion. The total error is calculated as the square root of the diagonal elements of the following CM:

$$\mathbf{S}_{i,\mathrm{tot}} = (\mathbf{K}_i^{\mathrm{T}}\mathbf{S}_{yi}^{-1}\mathbf{K}_i + \mathbf{S}_{ai}^{-1})^{-1} \tag{16}$$

for the TIR, UV, and VIS measurements and

$$\mathbf{S}_{f,\mathrm{tot}} = \left(\sum_{i=1}^{N}\mathbf{A}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{A}_i + \mathbf{S}_a^{-1}\right)^{-1} \tag{17}$$

for the fused profile

- The synergy factor (SF) (Aires et al. 2012). For each pressure level (j) the SF is defined as the ratio between the minimum total error of the fusing profiles ($\sigma_{i,\mathrm{tot}}^{(j)}$) and the total error of the fused profile ($\sigma_{f,\mathrm{tot}}^{(j)}$):

$$\mathrm{SF}^{(j)} = \frac{\min_{i=1,2,\ldots,N}\sigma_{i,\mathrm{tot}}^{(j)}}{\sigma_{f,\mathrm{tot}}^{(j)}}. \tag{18}$$

When a synergy among the sources of information exists the error SF is larger than 1 (supposing that the same a priori CM is used for the individual and fused measurements).

- The average number of degrees of freedom (DOF) for TIR, UV, VIS, and fused profiles
- The values of the diagonal elements of AKMs for TIR, UV, VIS, and fused profiles

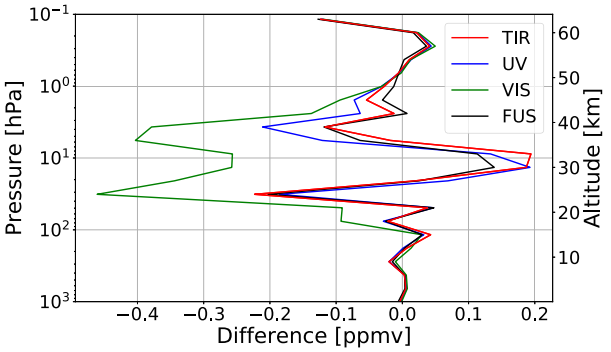

FIG. 2. Average differences between the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line) with respect to the values of the true profiles.

a. GEO–GEO fused data

In the GEO–GEO data fusion, the simulated ozone data from the single sensors (TIR, UV, and VIS) of the geostationary platform (S4) are fused. In this case, S4 measurements were simulated in the same space–time locations for the three spectral bands; therefore, it has been possible to select the profiles to fuse that correspond exactly to the same location. Since in the GEO–GEO case the distances between different simulated pixels of the same band are always larger than the dimension of the coincidence grid cell, we fused sets of three measurements corresponding to three different spectral regions related to the same space–time locations. The total number of analyzed pixels, where the three retrieved profiles have been fused, is 28 938. In Fig. 1, examples of GEO–GEO coincidences are shown. Because of the exact coincidence of the pixels locations, fused measurements are referred to the same true profile, and therefore, the fusion is performed with the coincidence error equal to zero. Moreover, since the measurements were simulated on the same vertical grid for all the three spectral regions, also the interpolation error is zero. Since both single retrieval and data fusion algorithm use the same a priori profile and the same a priori CM, we can easily compare the quality of the data fusion product with that of the products retrieved from TIR, UV, and VIS sensors. In the following analysis, the total columns retrieved from the VIS measurements have been transformed in vertical profiles with the method described in section 3.

Figure 2 shows that the differences between the fused and the true profiles are smaller or comparable with the same differences obtained considering TIR, UV, and VIS measurements instead of the fused one. Figure 3 shows that the average total error of the fused product is smaller than the average total errors of the single retrieval products at all pressure levels and we can see on the right that the average of the diagonal elements of the AKMs of the fused product is larger than the average of

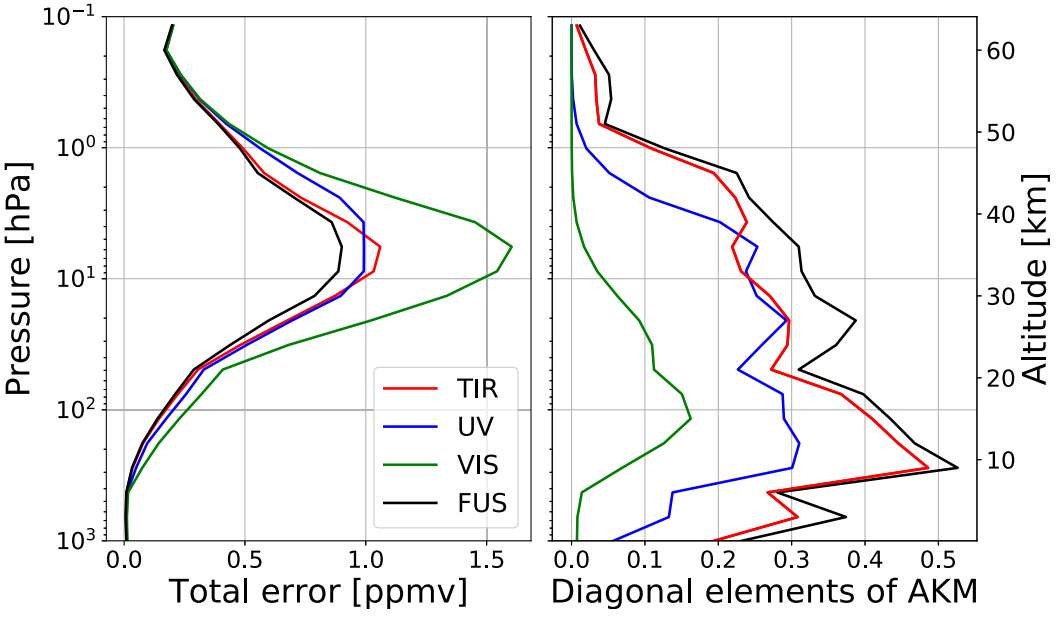

FIG. 3. (left) Average total errors of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line). (right) Average of the diagonal elements of the AKMs of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line).

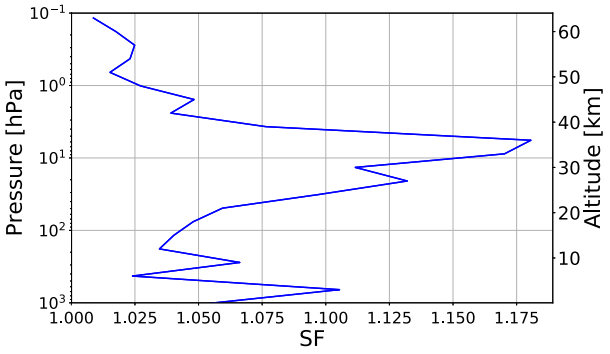

FIG. 4. Average of the synergy factor.

the diagonal elements of the AKMs of the single sensor products at all pressure levels. Moreover, Fig. 4 demonstrates that the average SF is larger than 1 at all pressure levels.

In Table 2, we report the average of the number of DOFs of the TIR, UV, VIS, and fused profiles for the case of GEO–GEO data fusion. We can see that on average the fused profile has 0.82 DOFs more than the TIR profile, 2.31 DOFs more than the UV profile, and 4.75 DOFs more than the VIS profile.

b. *LEO–LEO fused data*

In the case of LEO–LEO fusion, the S5 single sensor measurements were not simulated in the same space–time location for the three spectral bands; thus, the profiles used for the data fusion process are not in perfect coincidence. As a consequence, they refer to different true profiles and the introduction of a coincidence error is needed, as shown in section 3. The CM of the coincidence error is calculated considering an error of 5% of the a priori profile and a correlation length of 6 km. As shown in Ceccherini et al. (2019) (Fig. 2), the fused profile is slightly dependent on the value used for the coincidence error, provided that it is different from zero. Therefore, the specific choice we made for the value of the coincidence error has a small impact on the results of this study. The correlation length is used to reduce oscillations in the retrieved profile and the value of 6 km is typically used for nadir ozone profile retrieval (Liu et al. 2010; Kroon et al. 2011; Miles et al. 2015).

The comparison of the quality of the fused products with respect to the quality of the individual measurements is not as clear as in the case of the fusion of S4 measurements. For the LEO–LEO fusion the profile obtained from the CDF represents an estimate of the mean of the true profiles targets of the observations in the coincidence cell. Thus, the same quality estimators used for GEO–GEO fusion will not be explicative because these quantifiers refer to the estimation of different profiles.

TABLE 2. Average number of DOFs of the TIR, UV, VIS, and fused measurements for the three experiments of the CDF.

	CDF experiment	TIR	UV	VIS	FUS
Number of DOFs	GEO–GEO	4.90	3.41	0.97	5.72
	LEO–LEO	4.90	6.54	0.95	6.47
	LEO–GEO	4.94	5.62	0.96	6.34

As for the GEO–GEO case, both the retrievals of the single sensor measurements and the data fusion algorithm use the same a priori profiles and the same a priori CMs. The vertical grids of the retrieved profiles from the three spectral regions are the same except for the lowest point that, corresponding to the surface level, can be different in the different geolocations. As a consequence, the interpolation errors can be different from zero at the lowest altitudes. In Fig. 5, examples of LEO–LEO coincidences are shown. The total number of analyzed profiles is 46 567 for TIR, 67 864 for UV, 59 096 for VIS measurements, and the resulting fused profiles are 78 623.

Figure 6 shows that, in average, the profiles obtained from the fusion process have differences with respect to the true profiles smaller or comparable with those of the profiles obtained from the single sensor measurements as in GEO–GEO fusion.

Figure 7 demonstrates that in the LEO–LEO data fusion the smallest total error is that related to the UV measurement and not to the fused profile. In this case, the fused product is obtained fusing different combinations of the single sensor measurements and often the UV measurement is not included in the fusion. Furthermore, the introduction of the coincidence error determines an increase of the error of the fused product. However, we cannot conclude that the fused profile quality is worse than that of the UV profile, because the two retrieved profiles, as explained above, estimate different profiles. This consideration also applies to the analysis of the other quality estimators adopted in this section. Figure 7 also shows that the average value of the diagonal elements of the AKMs of the fused product is not always the largest one, as in the case of fusion of S4 measurements. At several altitudes the average values related to UV and TIR measurements are larger than that of the fused profile.

Table 2 shows the average of the number of DOFs of TIR, UV, VIS, and fused measurements. We can see that on average the fused profile has 1.57 DOFs more than the TIR profile, 0.07 DOFs less than the UV profile, and 5.52 DOFs more than the VIS profile.

Because of the introduction of the coincidence error the average of the SFs is mostly less than 1, as shown in

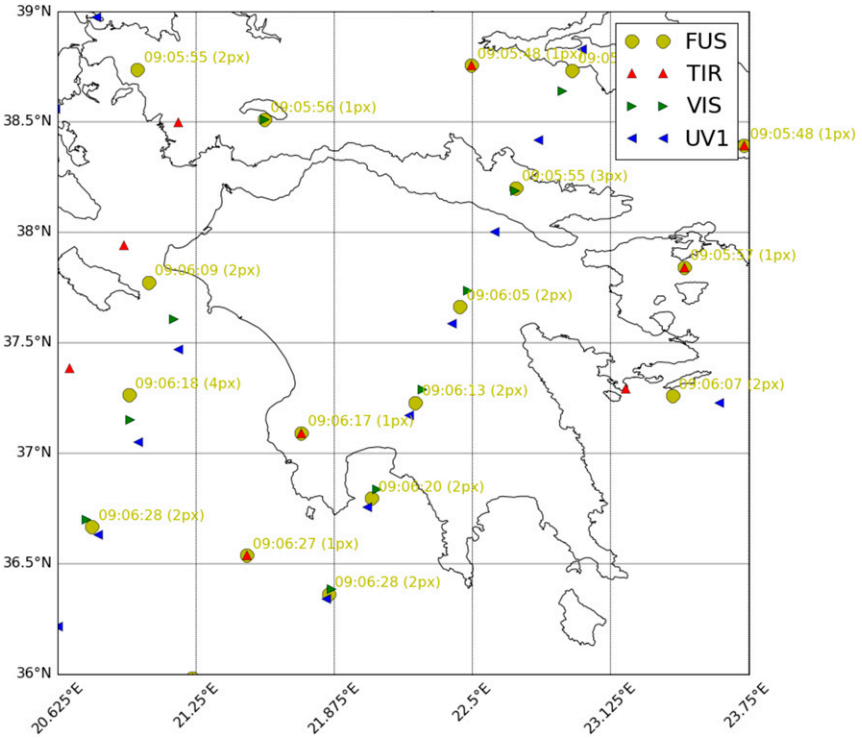

FIG. 5. Example of LEO–LEO coincidences over Greece, the black lines represent the borders of the coincidence grid cells (0.5° in latitude × 0.625° in longitude).

Fig. 8 (left panel). For this analysis, we considered only cases where a fusion really took place; that is, we excluded those where, in a coincidence cell, a single measurement occurs. Plots in Fig. 8 are obtained averaging the SF corresponding to 52 587 fused profiles. To make the quality quantifiers of the individual and fused products comparable, we decided to introduce the coincidence errors also in the individual products and to consider the individual products as estimates of the mean of the true profiles in the cell. All quantities related to each individual product were used as inputs to the modified CDF formula including the coincidence CM. In this way we obtained the corresponding product, for each individual one, representing the estimate of the mean of the true profiles within the coincidence cell.

In Fig. 8 (right panel), we show the average SF when the coincidence errors are included also in the individual products. We see that in this case the synergy factor is larger than 1 at all altitudes, showing that when the coincidence errors are included also in the individual products, the quantifiers are comparable because they refer to the same estimated profile.

c. *LEO–GEO fused data*

The CDF method has been also used to fuse TIR, UV, and VIS ozone profiles retrieved from S4 and S5 simulated measurements. In this case, no coincidence error was introduced in the fusion process if all measurements falling in a coincidence cell were in perfect time and spatial coincidence. If the fusing measurements were not in perfect coincidence, a coincidence error was introduced to all of them. As in the LEO–LEO case, the CM of the coincidence error is calculated considering an error of 5% of the a priori profile and a correlation length of 6 km. In Fig. 9, examples of LEO–GEO coincidences are shown. As in the case of

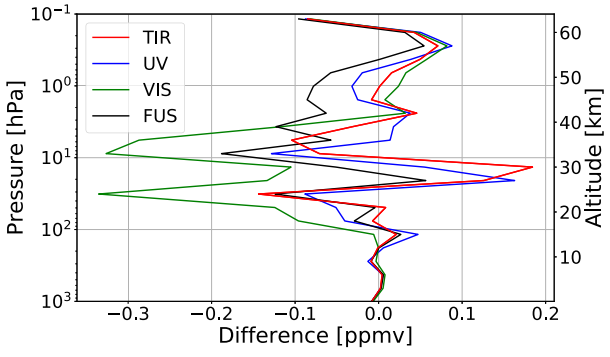

FIG. 6. Average differences between the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line) with respect to the means of the true profiles.

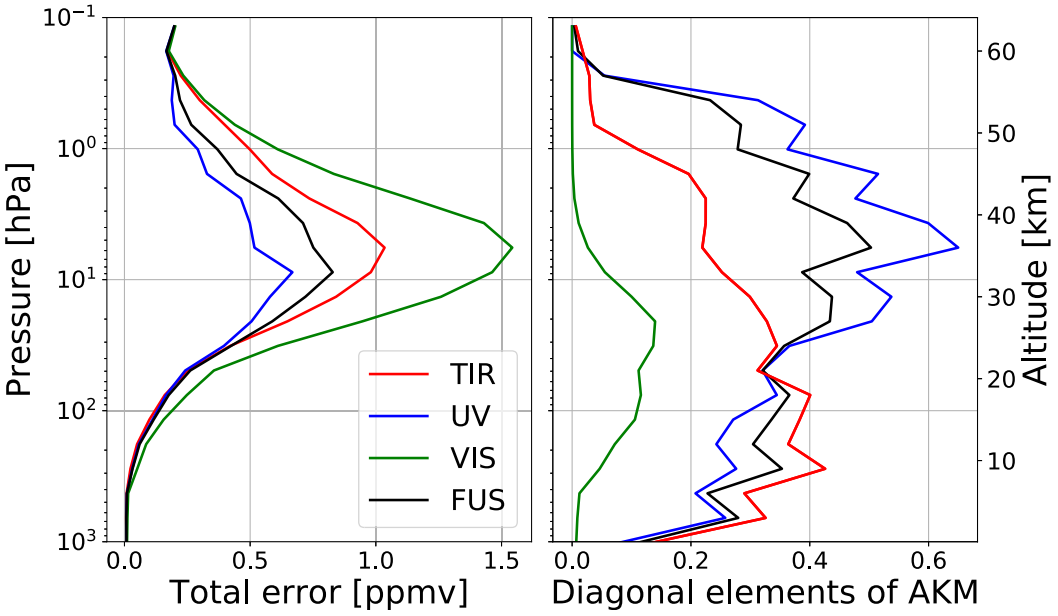

FIG. 7. (left) Average total errors of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line). (right) Average of the diagonal elements of the AKMs of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS (green line) measurements and from the data fusion (black line).

LEO–LEO fusion, the comparison of the quality of fused and individual products should take into account that the quality quantifiers refer to the estimation of different profiles. As in the case of GEO–GEO and LEO–LEO fusion, the retrievals of the TIR, UV, and VIS simulated measurements and the data fusion algorithm use the same a priori profiles and the same a priori CMs. The vertical grids of the retrieved profiles from the three spectral regions are the same with the exception of the lowest point, corresponding to the surface level varying because of the different geolocations. As a consequence, the interpolation errors can differ from zero at the lowest altitudes. The total number of analyzed profiles is 75 506 for TIR, 96 803 for

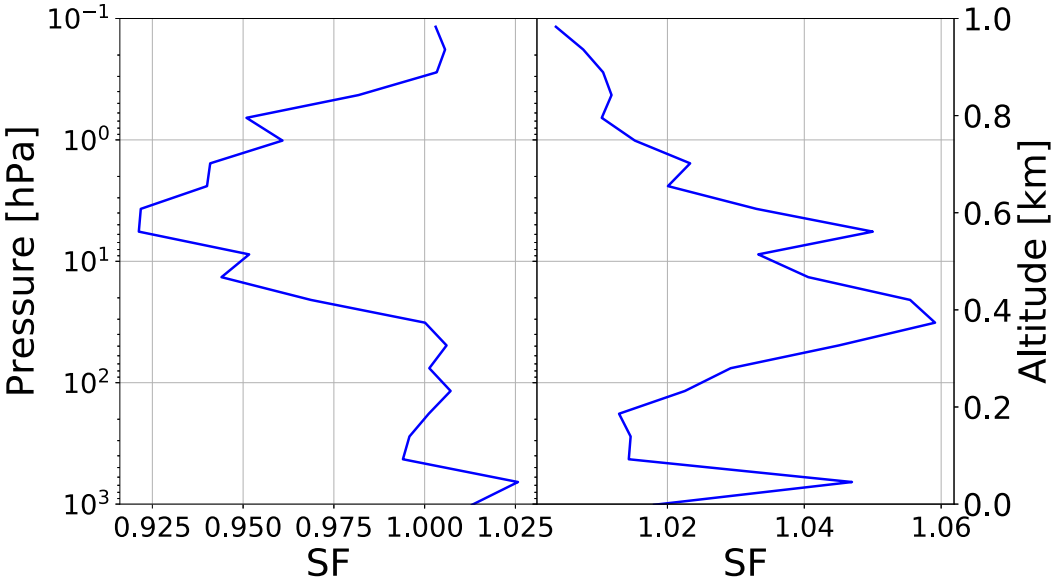

FIG. 8. (left) Average SF; (right) average SF when the coincidence errors are included also in the individual products.

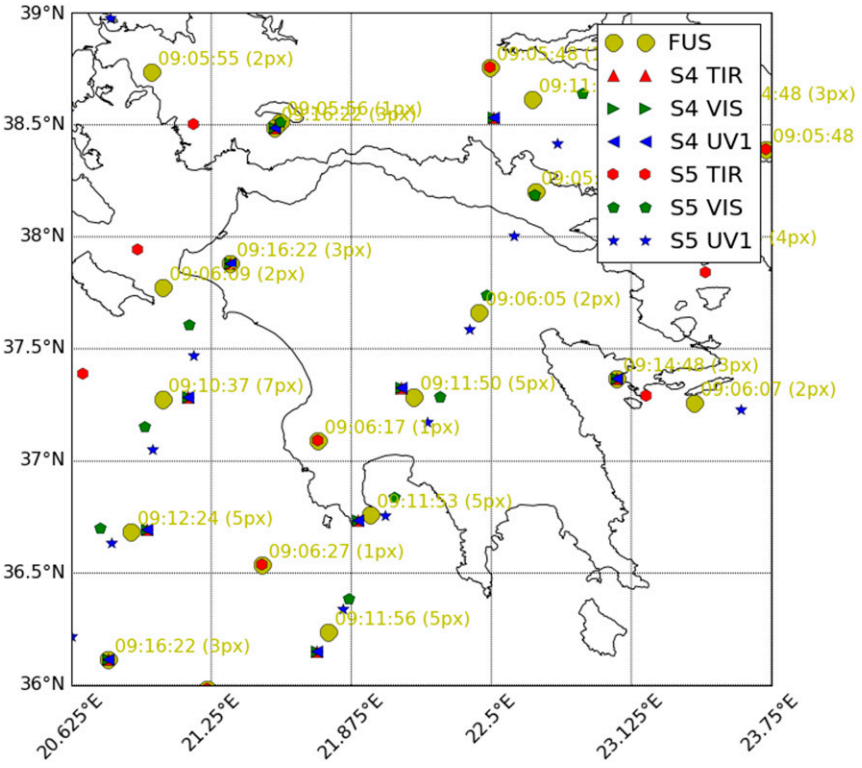

FIG. 9. Example of LEO–GEO coincidences over Greece, the black lines represent the borders of the coincidence grid cells (0.5° in latitude × 0.625° in longitude).

UV, 88 035 for VIS measurements, and the obtained fused profiles are 104 447. Figure 10 shows that, in average, the profiles obtained from the data fusion have differences with respect to the true profiles smaller or comparable with those of the profiles obtained from TIR, UV, and VIS measurements.

The behavior of the total errors observed in Fig. 11 is very similar to that observed for LEO–LEO fusion (see Fig. 7) and the same considerations made above apply also here. The average value of the AKMs diagonal elements of the fused product is not always the largest one, at several altitudes the values related to TIR and UV measurements are larger.

Table 2 summarizes the DOF average values for TIR, UV, VIS, and fused profiles. The fused profile has 1.40 DOFs more than the TIR profile, 0.72 DOFs more than the UV profile, and 5.38 DOFs more than the VIS profile.

In Fig. 12 (left panel), the average of the SFs is shown. Because of the introduction of the coincidence error the average of the SFs is mostly less than 1.

As in the LEO–LEO fusion, in order to make the quality quantifiers of the individual and fused products comparable, all quantities related to each individual product were used as inputs to the modified CDF formula including the coincidence CM. In this way, we obtained the corresponding product representing the estimate of the mean of the true profiles within the coincidence cell for each individual product.

Figure 12 (right panel) shows the average SF when the coincidence errors are included in the total uncertainty of the individual products. Now the SF is larger than 1 at all altitudes demonstrating that the quantifiers, when referred to the same estimated profile, prove the higher

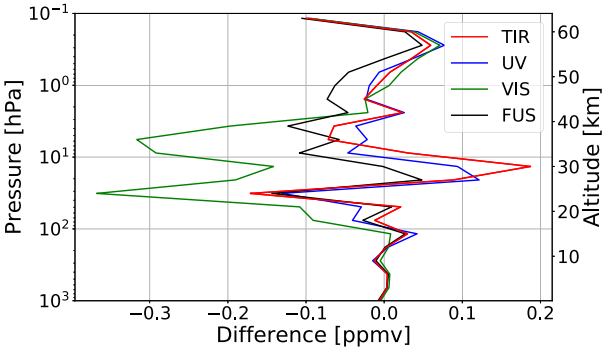

FIG. 10. Average differences between the ozone profiles obtained from the TIR measurements (red line), from the UV measurements (blue line), from the VIS measurements (green line), and from the data fusion (black line) with respect to the means of the true profiles.

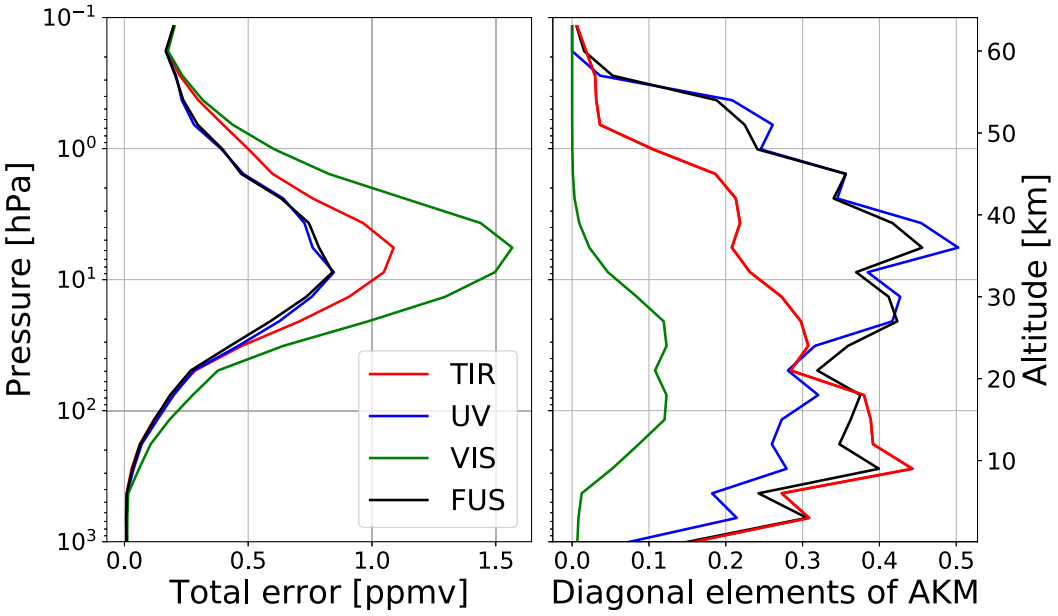

FIG. 11. (left) Average total errors of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS measurements and from the data fusion (black line). (right) Average of the diagonal elements of the AKMs of the ozone profiles obtained from the TIR (red line), UV (blue line), and VIS measurements and from the data fusion (black line).

quality of the fused product with respect to that of the individual products.

5. Conclusions

In this study, the production and characterization of combined LEO (S5) and GEO (S4) fused ozone data is performed. Fused and simulated standard products have been compared and a quality assessment of the generalized CDF is presented. The generalized CDF method has been used for three data fusion experiments: the GEO–GEO, the LEO–LEO, and the LEO–GEO data fusion. For each experimentm, we produced the fused data corresponding to the simulated measurements of

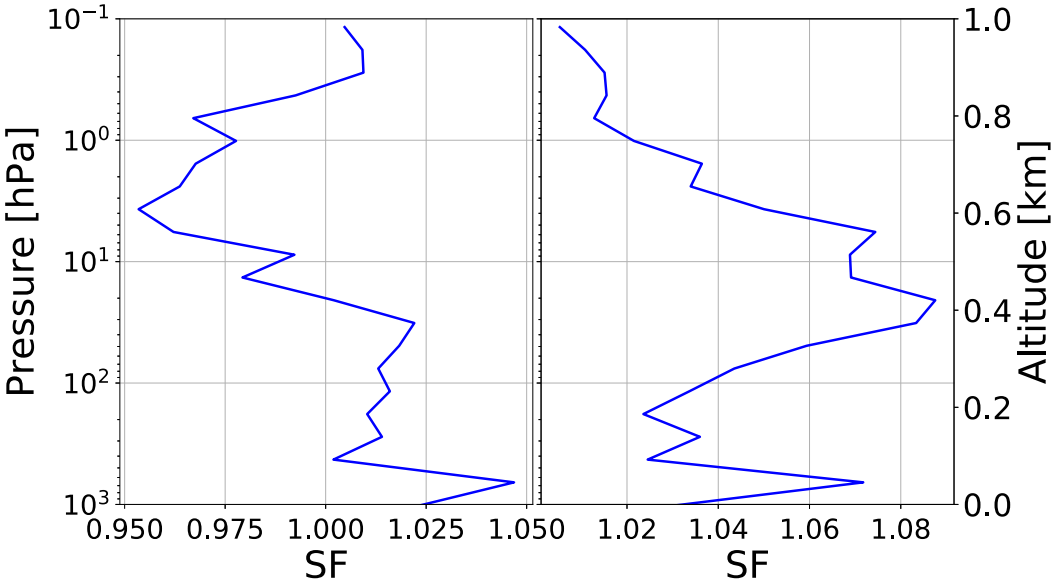

FIG. 12. (left) Average SF; (right) average SF when the coincidence errors are included also in the individual products.

the first week of April 2012. The quantifiers used to evaluate the quality of the fused data with respect to the standard products are described and a complete analysis is provided for all experiments. For the LEO–LEO and the LEO–GEO data fusion, in order to make the quantifiers of the individual and fused products comparable, the coincidence errors have been introduced also in the individual products considered as estimates of the mean of the true profiles in the cell. The analysis of the output products of the CDF algorithm by using quality quantifiers demonstrates that the generalized CDF algorithm provides products of better quality compared with that of standard products when the standard products are considered as estimates of the mean of the true profile in the coincidence cell.

Acknowledgments. The AURORA project is supported by the Horizon 2020 research and innovation program of the European Union (Call: H2020-EO-2015; Topic: EO-2-2015) under Grant Agreement 687428. The AURORA Consortium gratefully acknowledges the valuable and constant support on the many aspects of the project provided by the members of the External Expert Advisory Board: Marina Khazova (Public Health England), William Lahoz (Norwegian Institute for Air Research), Alan O'Neill (University of Reading), and Dimitris Stathakis (University of Thessaly).

REFERENCES

Aires, F., O. Aznay, C. Prigent, M. Paul, and F. Bernardo, 2012: Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A. *J. Geophys. Res.*, **117**, D18304, https://doi.org/10.1029/2011JD017188.

Calisesi, Y., V. T. Soebijanta, and R. van Oss, 2005: Regridding of remote soundings: Formulation and application to ozone profile comparison. *J. Geophys. Res.*, **110**, D23306, https://doi.org/10.1029/2005JD006122.

Ceccherini, S., and M. Ridolfi, 2010: Technical note: Variance-covariance matrix and averaging kernels for the Levenberg-Marquardt solution of the retrieval of atmospheric vertical profiles. *Atmos. Chem. Phys.*, **10**, 3131–3139, https://doi.org/10.5194/acp-10-3131-2010.

——, B. Carli, E. Pascale, M. Prosperi, P. Raspollini, and B. M. Dinelli, 2003: Comparison of measurements made with two different instruments of the same atmospheric vertical profile. *Appl. Opt.*, **42**, 6465–6473, https://doi.org/10.1364/AO.42.006465.

——, ——, and P. Raspollini, 2015: Equivalence of data fusion and simultaneous retrieval. *Opt. Express*, **23**, 8476–8488, https://doi.org/10.1364/OE.23.008476.

——, ——, and ——, 2016: Vertical grid of retrieved atmospheric profiles. *J. Quant. Spectrosc. Radiat. Transfer*, **174**, 7–13, https://doi.org/10.1016/j.jqsrt.2016.01.018.

——, ——, C. Tirelli, N. Zoppetti, S. Del Bianco, U. Cortesi, J. Kujanpää, and R. Dragani, 2018: Importance of interpolation and coincidence errors in data fusion. *Atmos. Meas. Tech.*, **11**, 1009–1017, https://doi.org/10.5194/amt-11-1009-2018.

——, N. Zoppetti, B. Carli, U. Cortesi, S. Del Bianco, and C. Tirelli, 2019: The cost function of the data fusion process and its application. *Atmos. Meas. Tech.*, **12**, 2967–2977, https://doi.org/10.5194/amt-12-2967-2019.

Cortesi, U., S. Del Bianco, M. Gai, L. M. Laurenza, S. Ceccherini, B. Carli, F. Barbara, and M. Buchwitz, 2014: Sensitivity analysis and application of KLIMA algorithm to GOSAT and OCO validation. KLIMA-ISAI Tech. Rep., Vol. 6, 153 pp.

——, and Coauthors, 2018: Advanced Ultraviolet Radiation and Ozone Retrieval for Applications (AURORA): A project overview. *Atmosphere*, **9**, 454, https://doi.org/10.3390/atmos9110454.

Costantino, L., and Coauthors, 2017: Potential of multispectral synergism for observing ozone pollution by combining IASI-NG and UVNS measurements from the EPS-SG satellite. *Atmos. Meas. Tech.*, **10**, 1281–1298, https://doi.org/10.5194/amt-10-1281-2017.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, https://doi.org/10.1002/qj.49712051912.

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.

Cuesta, J., and Coauthors, 2013: Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe. *Atmos. Chem. Phys.*, **13**, 9675–9693, https://doi.org/10.5194/acp-13-9675-2013.

ESA, 2000: Sentinel-5. ESA, accessed 4 February 2020, https://sentinel.esa.int/web/sentinel/missions/sentinel-5.

——, 2016: Sentinel-5 Precursor: ESA's Atmospheric Chemistry and Pollution Monitoring Mission. ESA Rep. SP-1332, 86 pp., https://esamultimedia.esa.int/multimedia/publications/SP-1332/SP-1332.pdf.

——, 2017: Sentinel-4: ESA's geostationary atmospheric mission for Copernicus operational services. ESA Rep. SP-1334, 92 pp., http://esamultimedia.esa.int/multimedia/publications/SP-1334/SP-1334.pdf.

ESA Mission Science Division, 2007: GMES Sentinels 4 and 5: Mission Requirements Document. ESA Rep. EOP-SMA/1507/JL-dr, 87 pp., https://earth.esa.int/c/document_library/get_file?folderId=13019&name=DLFE-794.pdf.

——, 2017: Copernicus Sentinels 4 and 5 Mission Requirements Traceability Document. ESA Rep. EOP-SM/2413/BV-bv, 166 pp., https://earth.esa.int/documents/247904/2506504/Copernicus-Sentinels-4-and-5-Mission-Requirements-Traceability-Document.pdf.

EUMETSAT, 2010: Post-EPS Mission Requirements Document. EUMETSAT Rep. EUM/PEPS/REQ/06/0043, 196 pp., https://www.dlr.de/rd/Portaldata/28/Resources/dokumente/re/pdf_peps_mrd_Post-EPS-Mission-Requirements-Document.pdf.

Fu, D., J. R. Worden, X. Liu, S. S. Kulawik, K. W. Bowman, and V. Natraj, 2013: Characterization of ozone profiles derived from Aura TES and OMI radiances. *Atmos. Chem. Phys.*, **13**, 3445–3462, https://doi.org/10.5194/acp-13-3445-2013.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Hache, E., and Coauthors, 2014: The added value of a visible channel to a geostationary thermal infrared instrument to monitor

ozone for air quality. *Atmos. Meas. Tech.*, **7**, 2185–2201, https://doi.org/10.5194/amt-7-2185-2014.

Hassler, B., and Coauthors, 2014: Past changes in the vertical distribution of ozone—Part 1: Measurement techniques, uncertainties and availability. *Atmos. Meas. Tech.*, **7**, 1395–1427, https://doi.org/10.5194/amt-7-1395-2014.

Heue, K.-P., M. Coldewey-Egbers, A. Delcloo, C. Lerot, D. Loyola, P. Valks, and M. van Roozendael, 2016: Trends of tropical tropospheric ozone from 20 years of European satellite measurements and perspectives for the Sentinel-5 Precursor. *Atmos. Meas. Tech.*, **9**, 5037–5051, https://doi.org/10.5194/amt-9-5037-2016.

Huijnen, V., and Coauthors, 2010: The global chemistry transport model TM5: Description and evaluation of the tropospheric chemistry version 3.0. *Geosci. Model Dev.*, **3**, 445–473, https://doi.org/10.5194/gmd-3-445-2010.

Ingmann, P., B. Veihelmann, J. Langen, D. Lamarre, H. Stark, and G. B. Courrèges-Lacoste, 2012: Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/-5 and -5p. *Remote Sens. Environ.*, **120**, 58–69, https://doi.org/10.1016/j.rse.2012.01.023.

Krol, M., and Coauthors, 2005: The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications. *Atmos. Chem. Phys.*, **5**, 417–432, https://doi.org/10.5194/acp-5-417-2005.

Kroon, M., J. F. de Haan, J. P. Veefkind, L. Froidevaux, R. Wang, R. Kivi, and J. J. Hakkarainen, 2011: Validation of operational ozone profiles from the Ozone Monitoring Instrument. *J. Geophys. Res.*, **116**, D18305, https://doi.org/10.1029/2010JD015100.

Lahoz, W. A., and Coauthors, 2012: Monitoring air quality from space: The case for the geostationary platform. *Bull. Amer. Meteor. Soc.*, **93**, 221–233, https://doi.org/10.1175/BAMS-D-11-00045.1.

Landgraf, J., and O. P. Hasekamp, 2007: Retrieval of tropospheric ozone: The synergistic use of thermal infrared emission and ultraviolet reflectivity measurements from space. *J. Geophys. Res.*, **112**, D08310, https://doi.org/10.1029/2006JD008097.

Liu, X., P. K. Bhartia, K. Chance, R. J. D. Spurr, and T. P. Kurosu, 2010: Ozone profile retrievals from the Ozone Monitoring Instrument. *Atmos. Chem. Phys.*, **10**, 2521–2537, https://doi.org/10.5194/acp-10-2521-2010.

McPeters, R. D., and G. J. Labow, 2012: Climatology 2011: An MLS and sonde derived ozone climatology for satellite retrieval algorithms. *J. Geophys. Res.*, **117**, D10303, https://doi.org/10.1029/2011JD017006.

Miles, G. M., R. Siddans, B. J. Kerridge, B. G. Latter, and N. A. D. Richards, 2015: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation. *Atmos. Meas. Tech.*, **8**, 385–398, https://doi.org/10.5194/amt-8-385-2015.

Natraj, V., and Coauthors, 2011: Multi-spectral sensitivity studies for the retrieval of tropospheric and lowermost tropospheric ozone from simulated clear-sky GEO-CAPE measurements. *Atmos. Environ.*, **45**, 7151–7165, https://doi.org/10.1016/j.atmosenv.2011.09.014.

Nirala, M., 2008: Multi-sensor data fusion and comparison of total column ozone. *Int. J. Remote Sens.*, **29**, 4553–4573, https://doi.org/10.1080/01431160801927202.

Platt, U., 1994: Differential optical absorption spectroscopy (DOAS). *Air monitoring by Spectroscopic Techniques*, M. Siegrist, Ed., John Wiley and Sons, 27–84.

Quesada-Ruiz, S., and Coauthors, 2020: Benefit of ozone observations from Sentinel-5P and future Sentinel-4 missions on tropospheric composition. *Atmos. Meas. Tech.*, **13**, 131–152, https://doi.org/10.5194/amt-13-131-2020.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, https://doi.org/10.1002/qj.49712656415.

Rodgers, C. D., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific, 238 pp.

Timmermans, R., W. Lahoz, J.-L. Attié, V.-H. Peuch, R. Curier, D. Edwards, H. Eskes, and P. Builtjes, 2015: Observing system simulation experiments for air quality. *Atmos. Environ.*, **115**, 199–213, https://doi.org/10.1016/j.atmosenv.2015.05.032.

Worden, J., X. Liu, K. Bowman, K. Chance, R. Beer, A. Eldering, M. Gunson, and H. Worden, 2007: Improved tropospheric ozone profile retrievals using OMI and TES radiances. *Geophys. Res. Lett.*, **34**, L01809, https://doi.org/10.1029/2006GL027806.