HARMONISED VALIDATION SYSTEM FOR TROPOSPHERIC OZONE AND OZONE PROFILE RETRIEVALS FROM GOME TO THE COPERNICUS SENTINELS

A. Keppens(1), J.-C. Lambert(1), D. Hubert(1), T. Verhoelst(1), J. Granville(1),
G. Ancellet(2), D. Balis(3), A. Delcloo(4), V. Duflot(5), S. Godin-Beekmann(2),
T. Leblanc(6), T. Stavrakou(1), W. Steinbrecht(7), R. Stübi(8), A. M. Thompson(9)

(1)Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, arno.keppens@aeronomie.be
(2)LATMOS/IPSL/CNRS/UVSQ/UPMC, Paris, France, gerard.ancellet@latmos.ipsl.fr
(3)Aristotle University of Thessaloniki (AUTH), Greece, balis@auth.gr
(4)Royal Meteorological Institute of Belgium (KMI-IRM), Brussels, Belgium, andy.delcloo@meteo.be
(5)LACy, Université de la Réunion, Saint-Denis, France, valentin.duflot@univ-reunion.fr
(6)California Institute of Technology, JPL, Wrightwood, CA, USA, leblanc@tmf.jpl.nasa.gov
(7)Deutsche Wetterdienst (DWD), Hohenpeissenberg, Germany, wolfgang.steinbrecht@dwd.de
(8)MeteoSwiss, Payerne, Switzerland, rene.stubi@meteoswiss.ch
(9)NASA/GSFC, Greenbelt, MD, USA, anne.m.thompson@nasa.gov

ABSTRACT

This work outlines the principles and implementation procedures of a harmonised QA/validation system for assessing the quality of atmospheric composition data. The system now runs in pre-operational mode at BIRA-IASB, following a generic and fully traceable flow of operations. The system's broad applicability to virtually all (partial) column and profile datasets is demonstrated here by evaluation activities within ESA's Ozone_cci project. However, ozone data validation as envisaged in the upcoming S5PVT AO project "CHEOPS-5p" and for the future Copernicus Sentinel missions raises a number of additional challenges, foremost regarding the expected amount of data to be handled. We therefore provide a perspective on current developments of the validation system to address those challenges and make it evolve into a fast operational phase.

1. INTRODUCTION

Research addressing air pollution, the maintenance of the stratospheric ozone layer, and global climate change requires global and long-term monitoring of the vertical distribution of atmospheric ozone at ever-improving resolution and accuracy. Global tropospheric and stratospheric ozone profile measurement capabilities from space have therefore improved substantially over the last two decades, among others with new generation hyperspectral instruments measuring backscattered UV-visible sunlight (GOME, SCIAMACHY, OMI, GOME-2) and thermal emission (TES, IASI) at the nadir of the satellite. Enhanced versions of those instruments are now being developed within EU's Copernicus Earth Observation program: the UV-visible-NIR TROPOMI instrument on board of the mid-afternoon satellite Sentinel-5P, to be launched in 2016, and both UV-visible and infrared instruments of the GOME and IASI types on board of the geostationary Sentinel-4 and polar orbiting Sentinel-5, to be launched at the end of the decade. Additionally, stringent climate research user requirements like e.g. the Global Climate Observing System (GCOS) targets call for continuous quality assessment and evolution of ozone data and their associated retrieval algorithms over the whole relevant spatial domain, vertical range, and mission lifetime [1]. The fitness-for-purpose of tropospheric ozone column and ozone profile data products must thus be warily verified by means of in-depth QA/validation studies of the satellite data and associated retrieval algorithms before being used in scientific research and operational applications.

To that purpose, an extensive validation system has been developed on the heritage of various validation activities, starting in the 1990s with the first GOME ozone profile validations and progressively extending up to the current ESA Multi-TASTE Phase F data evolution and Ozone_cci production of multi-mission climate data records on ozone [2]. Currently this validation system is being further consolidated with metrological traceability practices and with generic QA guidelines established within the FP7 QA4ECV project. Within the same project, it is also being applied to other chemical species (NO2, HCHO and CO). The end-to-end approach of this system combines preliminary QA/QC procedures, data content studies, in-depth information content studies, information-content based co-location procedures, data homogenisation, and the more traditional data comparisons with respect to reference measurements acquired by ground-based networks of ozonesondes and lidars (NDACC, SHADOZ, WMO GAW). An Observing System Simulation Experiment (OSSE) with detailed metrology of the remote sensing data is thereby used to assess the propagation of errors associated with differences in horizontal smoothing and with mismatches in space and time between the various measurements [3].

In this paper we first outline the principles and implementation procedures of this QA/validation system that is now in pre-operational phase at BIRA-IASB

(Section 2). Through evaluation activities within ESA's Ozone_cci project (www.esa-ozone-cci.org), we then demonstrate its broad applicability to virtually all ozone (partial) column and profile datasets (Section 3). We finally provide a perspective on the current developments of this system (Section 5) that are required to address the specific challenges (outlined in Section 4) of the TROPOMI ozone data validation as envisaged in the upcoming S5PVT AO project "CHEOPS-5p" (Validation of Copernicus HEight-resolved Ozone data Products from Sentinel-5P TROPOMI, ID28587).

2. QA/VALIDATION SYSTEM

The QA/validation system as described in this work can be considered an extension and generalisation of the (round-robin) evaluation chain presented in [2]. Its backbone consists of (i) the characterisation of satellite and reference data (including data and information content studies, co-location studies, and data harmonisation), (ii) the metrologically traceable confrontation of profiles or (tropospheric) columns with ground-based reference data, and (iii) the derivation of Quality Indicators that enable users the evaluate the fitness for purpose of the satellite data.

The full QA/validation scheme is schematised as a ten-step flowchart in Fig. 1, involving the following operations:

(1) *Requirements* as formulated by the ozone product (end) users have to be translated into accessible evaluation requirements first.

(2) *Data selection* and post-processing of both the satellite data under study and the reference data (space-based or ground-based) for later comparison. Input from other sources that are not directly involved in the validation might thereby be required, e.g. in order to vertically extend profiles for later smoothing operations, but should be introduced with great precaution to avoid undesirable corruption of the data. Besides, it is of major importance that the datasets show a maximal agreement in settings that influence the further evaluation and comparison, hence the need for mutual flagging of the data.

(3) *Data content studies* focus on the inspection of the atmospheric state data content distributions and correlations, impact of filtering, and geographical and temporal coverage. These studies yield important figures on the validity of the comparative analysis: The mutual data screening of step (2) should be iteratively altered in order to improve the sampling agreement or to (further) align specific data properties.

(4) *Information content studies* are based on the averaging kernel matrices (AKMs) of retrieved profiles [4] and primarily include information content measures such as the Degrees of Freedom of the Signal (DFS), measurement quality quantifier (MQQ) [5], eigenvectors and eigenvalues, and vertical sensitivity. Vertical resolution and height registration (offset) estimators can be studied as well. The meridian and temporal dependences of the information content are of particular interest, and are directly related to changes in the solar zenith angle and slant ozone column. As information content metrics should be unit-independent, fractional AKMs have to be considered [2].

(5) *Selection of correlative data* from established networks of ground-based reference instruments (NDACC, SHADOZ, WMO GAW) by choosing appropriate spatial and temporal co-location criteria. Research and reference (partial) columns or profiles that obey the criteria are then extracted from the original (filtered) datasets.

(6) *Correlative data statistics* are required to justify the co-location and coincidence criteria that are chosen in step (5): Based on (a) the spatial and temporal scale of the atmospheric ozone variation and on (b) the geographical and temporal sampling statistics of the satellite-reference coincidences, these criteria should be iteratively adjusted.

(7) *Data harmonisation* in terms of both the vertical quantity (units) and the vertical grid is mandatory for later difference calculations [6]. Other ozone profile manipulations that remove error contributions from the bias and spread statistics with a view on error budget closure exercises are optional. These manipulations include prior correction, vertical resolution matching (cf. AKM smoothing) [7], horizontal resolution matching, and sampling difference correction. The two last options can be achieved by application of an OSSE containing a gridded metrology of the ozone data [3].

(8) *Comparison statistics* are determined from (height-dependent) difference histograms and assess the mean (median) and the standard deviation (68 % interpercentile) of the difference in ozone as statistical estimators of the bias and spread between the satellite and reference data, possibly including full error budget closure [8]. Comparison results can be reported as a function of several parameters of importance, e.g. time, latitude, ozone slant column density, and cloud fractional cover.

(9) *Quality Indicators are derived* from validation steps (2) to (8) and can be summarised in an overview table for clarity.

(10) *User compliance verification* by judging the listed quality indicators with respect to the evaluation requirements concludes the QA/validation flow.

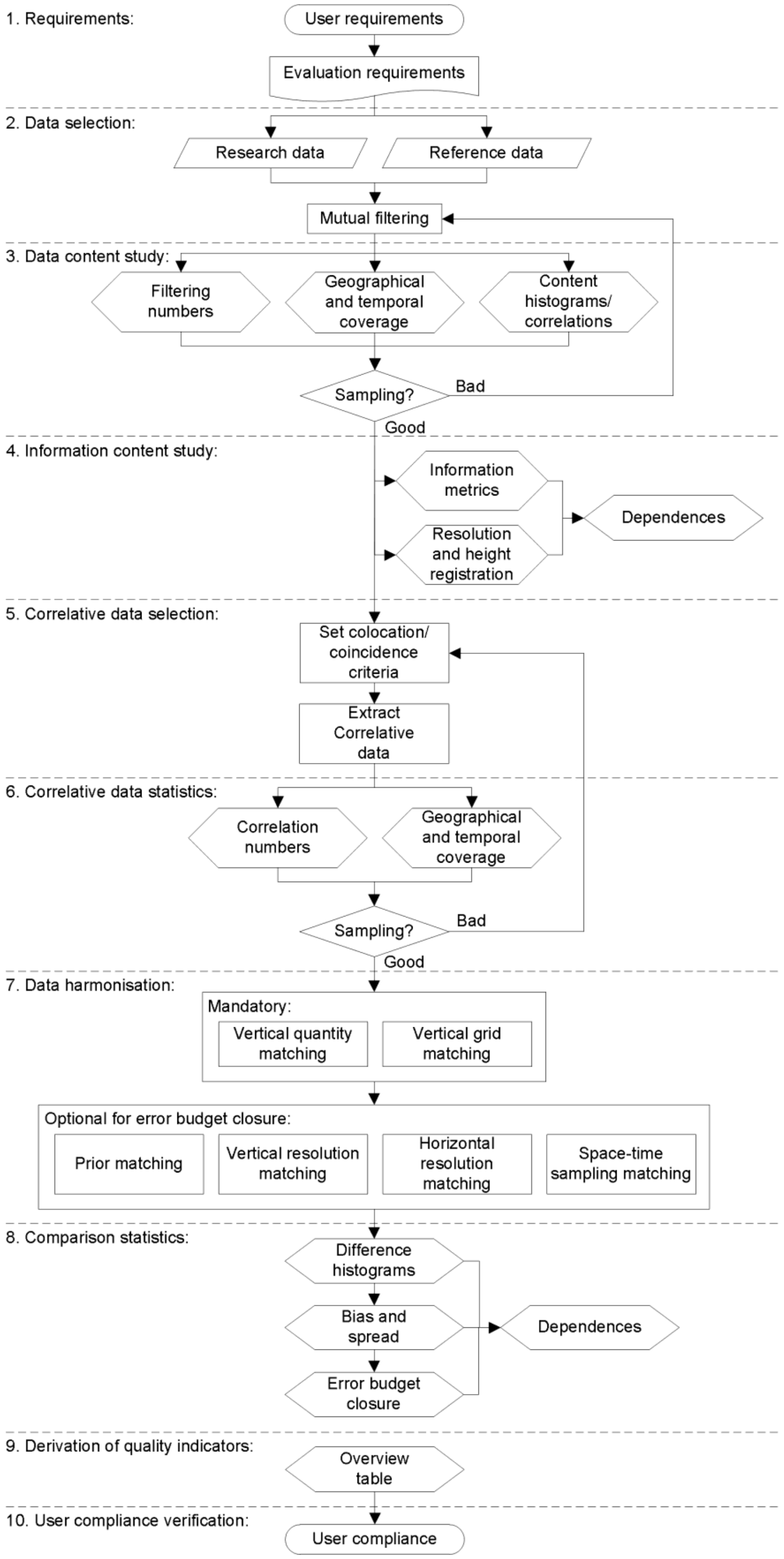
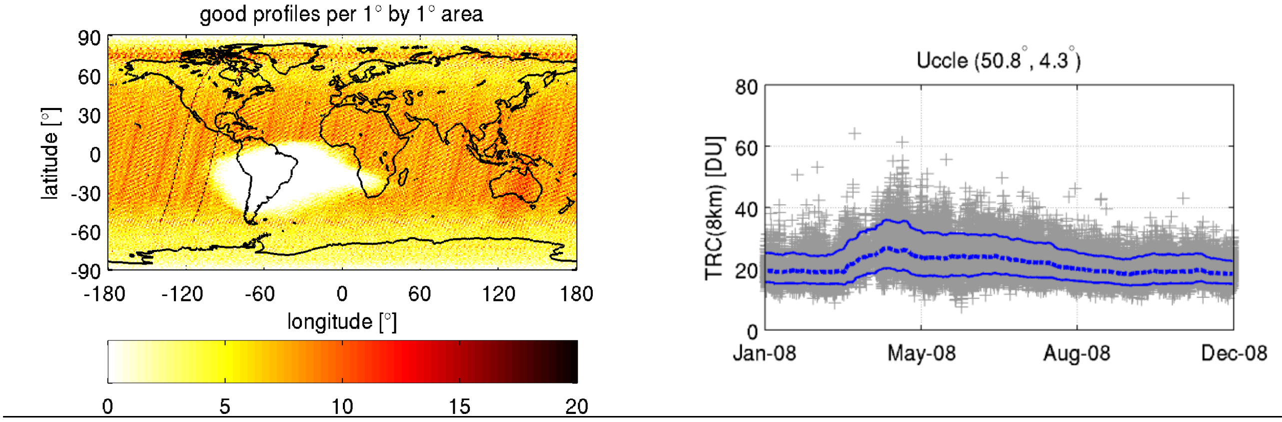
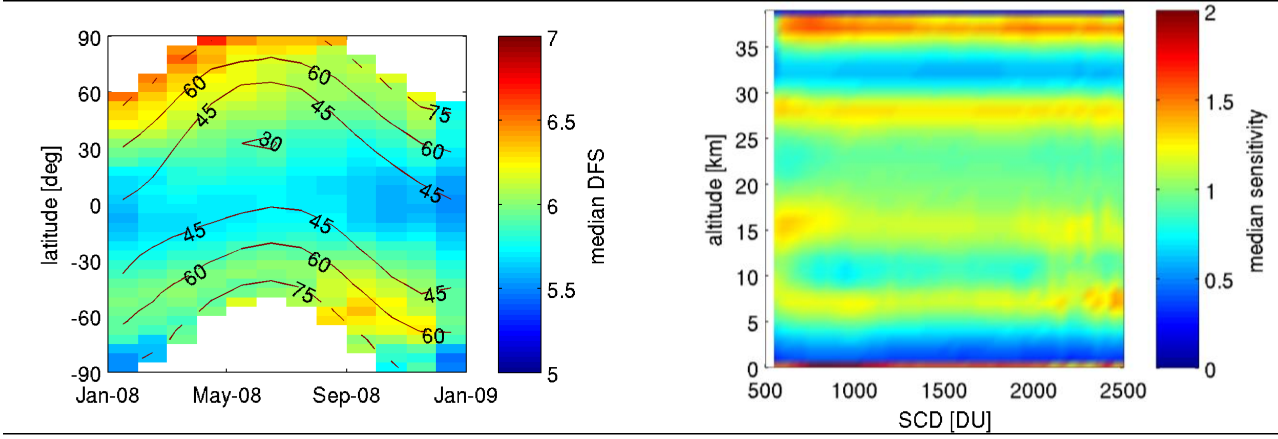


Figure 1. Ozone (partial) column and profile QA/validation system.

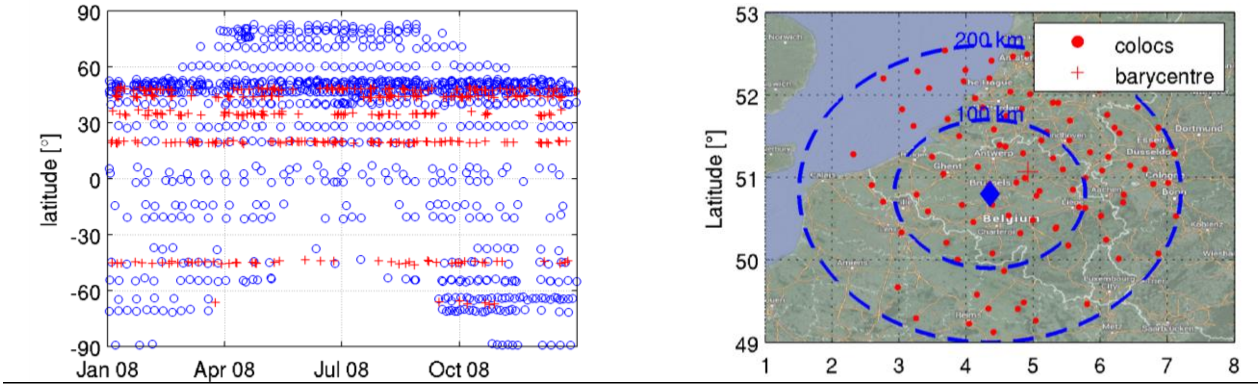
3. Data content studies



4. Information content studies



6. Correlative data statistics



8. Comparison statistics

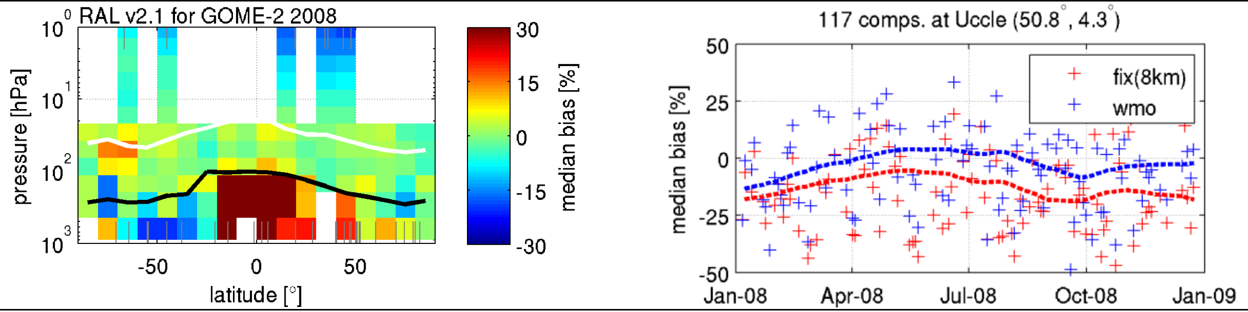


Figure 2. Exemplary output plots of the QA/validation system as obtained by its application to RAL (v2.1) retrievals of GOME-2 (globally, left) and FORLI (v20100815) retrievals of IASI (above Uccle, right), both for 2008 only.

3. SYSTEM DEMONSTRATION

Within the second phase of ESA's Climate Change Initiative on ozone (Ozone_cci), the extensive validation scheme presented in the previous section has been applied to two different satellite datasets: The first is the Rutherford Appleton Laboratory (RAL, version 2.1) retrieval of nadir ozone profiles from the GOME-2 instrument [9], while the second is the Fast Optimal Retrieval on Layers for IASI (FORLI, version 20100815) of tropospheric ozone columns [10]. Both retrieval schemes are based on the optimal estimation method [4]. The GOME-2 UV-visible and IASI infrared spectrometer instruments are located aboard the MetOp-A platform [11], launched into an 817 km altitude sun-synchronous polar orbit on 19 October 2006. In this work we focus on the measurements taken in 2008 only. Ozonesonde and lidar reference data are obtained from the Network for Detection of Atmospheric Composition Change (NDACC, www.ndsc.ncep.noaa.gov), Southern Hemisphere Additional OZonesonde (SHADOZ) program [12], and the World Ozone and Ultraviolet Data Centre (WOUDC, www.woudc.org) of the Global Atmospheric Watch (GAW).

Fig. 2 contains a selection of typical QA/validation output plots, related to steps 3, 4, 6, and 8 of the scheme presented in Fig. 1, for both retrieval datasets: Global results are shown for the RAL nadir ozone profiles, while station overpass plots for co-locations with Uccle (Belgium) are depicted for the FORLI tropospheric ozone columns. Together with a brief summary of more product-specific evaluation results, these outcomes are discussed in the next two subsections on the RAL and FORLI data validation, respectively.

3.1. Nadir ozone profile validation

The quality of the RAL v2.1 nadir ozone profile product as derived from GOME-2A spectral measurements is extensively discussed within a round-robin retrieval algorithm evaluation exercise in [2]: Data content studies reveal that RAL's initial data do not contain profiles for which the solar zenith angle exceeds 80° and that appropriate filtering removes about 10 % of the data, mainly resulting in a lack of spatial coverage around the South-Atlantic Anomaly (see Fig. 2 upper-left). Information content studies on the other hand show that RAL's output contains about 6 degrees of freedom in the signal (DFS), although a correlation with SZA can be observed (see Fig. 2 second row left). The measurement quality quantifier [5] shows a similar behaviour, while the vertical sensitivity – showing sharp oversensitive edges at about 5 and 50 km in altitude – decreases with increasing SCD (not shown). Several estimators show that RAL nadir ozone profiles show a rather poor vertical resolution, with values exceeding GCOS requirements over the complete 0-80 km altitude range. The smallest values are found around the UT/LS, where the median vertical resolution amounts to about 15 km (not shown). From statistics on the correlative data (with max. 200 km and 2 hours distance) it has been found that the RAL profiles for 2008 co-locate with 1254 ozonesonde and lidar measurements, which show a nice spatio-temporal distribution excluding the winter poles (see left in third row of Fig. 2). This homogeneity legitimates globally applied comparison statistics, revealing a high meridian and slant column density dependence of the bias, as shown in the lower-left plot in Fig. 2. The RAL ozone accuracy shows GCOS compliance around the UT/LS, but typically is too high below and slightly too low above.

3.2. Tropospheric ozone column validation

Nadir ozone profile and (tropospheric) ozone column validation for IASI as retrieved by the FORLI algorithm has only started in the current Phase 2 of the Ozone_cci project. Because of the very large amount of IASI data available for 2008, the summarised validation prospect presented here focuses on the measurements that co-locate with Uccle (Belgium): As part of the IASI data content studies, the temporal evolution of the tropospheric ozone column – obtained by integration of the nadir ozone profile from the ground up to 8 km altitude – is plotted in the upper-right corner of Fig. 2. Although the spread of the observations is high, the median tropospheric column (TRC) clearly shows an increase in local spring. Information content studies looking at the slant column density dependence of the vertical sensitivity reveal that this dependence is rather low, but also that tropospheric ozone column values have to be considered with care: The satellite ozone retrieval seems quite insensitive (< 0.5) to the lowest 3-4 km of the profile (see Fig. 2, second row on the right). Correlative data statistics for 200 km and 2 hour correlation criteria show that the co-location barycentre (median spatial location) can be off the ground station location, but is sufficiently close (50 km for the Uccle example in the third row of Fig. 2) to obtain reasonable comparison statistics. The latter in fact largely depend on the tropopause definition for ozone profile integration, as is demonstrated in the lower-right plot of Fig. 2: Although typically a negative and time-dependent IASI bias is observed at Uccle, the 8 km fixed altitude tropopause yields results that are remarkably (~10 %) lower than those obtained by integration below the WMO tropopause [13]. Tropopause definitions and related tropospheric ozone column calculations are therefore subject of further study within the ESA Ozone CCI.

4. SENTINEL CHALLENGES

The QA/validation system as outlined and demonstrated in the previous sections is currently applied in pre-operational mode. However, ozone data validation as envisioned for TROPOMI in the upcoming S5PVT AO project "CHEOPS-5p" (Validation of Copernicus HEight-resolved Ozone data Products from Sentinel-5P TROPOMI, ID28587) and for the planned Copernicus Sentinel missions raises a list of new challenges. These challenges, that will have to be rightfully addressed in ongoing and future validation system developments, can be summarised as follows:

- Large rate and amount of data;
- High horizontal resolution;
- Need for NRT quality monitoring;
- Broad scope of applications: air quality, climate...
- Need for improved error characterisation (including AK smoothing and error budget closure);
- Competing evaluation approaches and tropopause definitions for tropospheric ozone columns;
- Cloud-induced bias for height-resolved ozone;
- Different accuracies and vertical ranges of reference instruments, including ozonesondes, tropospheric lidars, stratospheric lidars, microwave radiometers, and FTIR spectrometers;
- And more challenges for other Sentinel missions, e.g. related to the Sentinel-4 geostationary orbit.

5. SYSTEM DEVELOPMENT

In order to address the validation challenges presented in the previous section and thereby make the existing system evolve towards a maximally operational phase, the following developments, linked to the validation steps of the flowchart in Fig. 1, are currently already being implemented:

(1) *Requirements* translation obviously remains a human task. The evaluation requirements however are formulated in a homogenised format to set up the suite of validation operations. This way they can act as a standard input for the next steps that are then addressed using a tick-box approach: The requirements (and users) determine which parts of the system should be dealt with.

(2) *Data selection* and post-processing of satellite and reference data is (already largely) automated, including e.g. weekly mirroring with the servers of the data providers. For very large datasets, operational (random) subset selection might be required for NRT evaluation exercises and monitoring. Mutual filtering can be semi-automated (i.e. with human validity check, also see next step) based on any flags that are present in the respective files under study.

(3) *Data content studies* are already straightforwardly automated by use of several plotting functionalities. The data sampling inspection nevertheless remains a human task that might need a manual alteration of the mutual filtering constraints afterwards.

(4) *Information content studies* largely depend on the presence and formatting of averaging kernel matrices (AKMs). The latter typically requires initial checks, but the metrics selected for study can be easily operationally determined from then on.

(5) *Selection of correlative data* (whether all or only the closest) is already automated based on tailorable co-location criteria. A fully operational selection could be based on a library of observation operators or on air mass (pixel) overlap [14]. For very large datasets as expected for TROPOMI, the correlative data selection could also be operationally performed even before all data are downloaded, i.e. by orbit prediction.

(6) *Correlative data statistics* are also operationally monitored already. Yet again, as for the data content studies, correlation criteria should be iteratively adjusted by hand if necessary.

(7) *Data harmonisation* of the vertical quantity (units) and grid is a mandatory operational step. Appropriate subcolumn definition (e.g. tropopause definition selection for tropospheric ozone column calculations) might however still call for human input. The greater difficulty for harmonisation automation is with the optional column or profile manipulations regarding error budget closure (EBC): Ideally they are applied subsequently, resulting in a full view on the known error contributions. Resolution and sampling mismatch errors can thereby be estimated using input from an Observing System Simulation Experiment like the in-house OSSSMOSE [3], but at the price of complicating this (preparatory) validation step.

(8) *Comparison statistics* are currently operationally determined from (height-dependent) difference histograms, including their dependence on (automatically) preselected parameters of importance, e.g. time, latitude, ozone slant column density, and cloud fractional cover.

(9) *Quality indicators are derived* from validation steps (2) to (8) only semi-automatically, as they require interpretation of numbers and plots.

(10) *User compliance verification* by judging the quality indicators with respect to the evaluation requirements can also only be partially automated, depending on the standardisation of the evaluation requirements definition in step (2). A human view on the user compliance obviously remains required.

6. CONCLUSIONS

Based on a ten-step data processing flowchart, this work outlines the principles and implementation procedures of an extensive and broadly applicable QA/validation system that is now in pre-operational phase at BIRA-IASB (Belgium). The system's current routine use is demonstrated for two MetOp-A instruments in the context of ESA's Ozone_cci project, namely by validation of GOME-2 nadir ozone profile retrievals by RAL and of IASI tropospheric ozone column retrievals by the FORLI algorithm. The ozone data validation that is planned for the Copernicus Sentinel missions however poses several additional challenges regarding processing in the evaluation scheme, mostly related to the expected amount of data. An overview of QA/validation system developments that are currently being implemented to address these challenges is therefore provided. Those developments are specifically linked to the ten steps of the existing validation flow and should allow the system to evolve into a fast operational phase that will first be applied to the TROPOMI ozone data validation in the upcoming S5PVT AO project "CHEOPS-5p" and future Copernicus Sentinel missions.

7. REFERENCES

[1] WMO. (2010). Implementation plan for the global observing system for climate in support of the UNFCCC. GCOS-138, Geneva, Switzerland.

[2] Keppens, A., Lambert, J.-C., Granville, J., Miles, G., Siddans, R., van Peet, J. C. A., van der A, R. J., Hubert, D., Verhoelst, T., Delcloo, A., Godin-Beekmann, S., Kivi, R., Stübi, R. & Zehner, C. (2015). Round-robin evaluation of nadir ozone profile retrievals: Methodology and application to MetOp-A GOME-2. *Atmos. Meas. Tech.* 8, pp. 2093-2120.

[3] Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J-P., Redondas, A., Van Roozendael, M. & Lambert, J-C. (2015). Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons. *Atmos. Meas. Tech.* (submitted).

[4] Rodgers, C. D. (2000). Inverse Methods for Atmospheric Sounding. World Scientific Publishing, Singapore.

[5] Ceccherini, S., Carli, B. & Raspollini, P. (2012). Quality quantifier of indirect measurements. *Opt. Exp.* 20, pp. 5151-5167.

[6] Calisesi, Y., Soebijanta, V. T. & van Oss, R. F. (2005). Regridding of remote soundings: Formulation and application to ozone profile comparison. *J. Geophys. Res.* 110(D23306), pp. 1-8.

[7] Rodgers, C. D. & Connor, B. J. (2003). Intercomparison of remote sounding instruments. *J. Geophys. Res.* 108(D34116), pp. 1-14.

[8] von Clarmann, T. (2006). Validation of remotely sensed profiles of atmospheric state variables: Strategies and terminology. *Atmos. Chem. Phys.* 6, pp. 4311-4320.

[9] Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G. & Richards, N. A. D. (2015). Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation. *Atmos. Meas. Tech.* 8, pp. 385-398.

[10] Hurtmans, D., Coheur, P.-F., Wespes, C., Clarisse, L., Scharf, O., Clerbaux, C., Hadji-Lazaro, J., George, M. & Turquety, S. (2012). FORLI radiative transfer and retrieval code for IASI. *J. Quant. Spectrosc. Ra.* 113, pp. 1391-1408.

[11] Munro, R., Siddans, R., Reburn, W. J. & Kerridge, B. J. (1998). Direct measurement of tropospheric ozone distributions from space, *Nature* 392, pp. 168-171.

[12] Thompson, A. M., Miller, S. K., Tilmes, S., Kollonige, D. W., Witte, J. C., Oltmans, S. J., Johnson, B. J., Fujiwara, M., Schmidlin, F. J., Coetzee, G. J. R., Komala, N., Maata, M., bt Mohamad, M., Nguyo, J., Mutai, C., Ogino, S.-Y., Raimundo Da Silva, F., Paes Leme, N. M., Posny, F., Scheele, R., Selkirk, H. B., Shiotani, M., Stübi, R., Levrat, G., Calpini, B., Thouret, V., Tsuruta, H., Valverde Canossa, J., Vömel, H., Yonemura, S., Andrés Diaz, J., Tan Thanh, N. T. & Thuy Ha, H. T. (2012). Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone climatology (2005-2009): Tropospheric and tropical tropopause layer (TTL) profiles with comparisons to OMI-based ozone products. *J. Geophys. Res.* 117(D23301), pp. 1-27.

[13] WMO. (1957). Meteorology – a three-dimensional science. WMO Bulletin 230(6), pp. 134-138.

[14] Balis, D., Lambert, J.-C., Van Roozendael, M., Spurr, R., Loyola, D., Livschitz, Y., Valks, P., Amiridis, V., Gerard, P., Granville, J. & Zehner, C. (2007). Ten years of OME/ERS2 total ozone data—The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8. *J. Geophys. Res.* 112(D07307), pp. 1-17.