

Interactive comment on “Automatic procedures for submitting essential climate variables (ECVs) recorded at Italian Atmospheric Observatories to WMO/GAW data centers” by Luca Naitza et al.

Luca Naitza et al.

p.cristofanelli@isac.cnr.it

Received and published: 13 December 2018

Here we provide a preliminary answer to the point raised by the reviewer #1. In the case of revision, we will consequently implement the manuscript. In the following, we provide our comments (AC) to referee's points.

Referee: The manuscript by Naitza et al. describes an automated data transmission and quality control system for a set of measurements from a set of Italian atmospheric composition monitoring stations. While this is clearly a relevant topic and I applaud

C1

the team members for having this accomplished and implemented, I don't see the manuscript worthy of publication in a peer-reviewed journal as it looks much more like a technical report.

AC: We thank the reviewer for taking his/her time in reviewing our paper. We would comment that from our point of view the manuscript meet the aims and scope of this peer-reviewed journal, considering that the development and validation of techniques concerning data processing and information retrieval from gases and aerosols measurements, constitutes an important pillar for scientific activities carried out at atmospheric observatories.

Referee: The topics of automated data processing, automated QA and flagging have been dealt with for many years, and many environmental agencies have produced heavy manuals with detailed discussion of procedures. Also WMO, which is mentioned several times in the manuscript, has produced a lot of material on such matters and operational weather centers are relying on automated procedures for daily weather forecasts. There are practically no references to any of these long-standing activities and it is rather unclear what the novel aspects of the approach described in this manuscript are.

AC: L. Naitza and co-authors are aware that in the framework of WMO many initiatives were carried out to deal with the question of near-real time data and automatic processes for quality check (among these WIGOS and WIS). However, the authors are pretty sure that the topic of automatic data handling and creation is still a challenge in the field of atmospheric composition data, especially for what concerns near-surface observations at atmospheric observatories. We participated to many national and international workshops, meetings and congresses: data quality check, data harmonization to specific formats and timely submission to reference data-sets are still a matter of strong concern for PIs, station managers and managing Institutions. In many situations, these processes still strongly rely on the manual intervention of technicians

C2

and scientists. We agree that the techniques we applied for setting-up this system are probably not innovative, but a prototype system has been created and it is currently working for a subset of Italian stations. A huge technical and scientific effort has been implemented in the 2-year activity of the project to make this system working. To our knowledge, this is one of the first systematic attempts of setting-up a series of automatic procedures for some of the most diffused instruments for trace gases, aerosol properties and meteorological parameters. It is great to know that heavy manuals with detailed discussion of procedures have been produced by WMO, but at this stage these information appeared to be not diffused (or shared) in the framework of atmospheric composition science. Leading initiatives like ICOS and ACTRIS are dealing with these issues, but a number of atmospheric observatories are not part of these initiatives. For these reasons, in the framework of the National Project NextDATA, founded by the Italian Ministry of Education, Universities and Research (MIUR), to provide a contribution to a more effective participation of Italian atmospheric observatories to WMO/GAW, we proposed the system object of this paper. The authors are confident that this initiative can be useful for many other research groups in the world working on the same research field: as a matter of fact the routines developed in this work, which are the pillars of our system, will be made available with open-access policy. On the other side, we think that it is valuable to provide a detailed description of the data-creation processes which will be adopted for a subset of ECVs for specific Italian WMO/GAW observatories. In any case, we will be glad if the reviewer would provide effective contacts to systematically implement these mentioned WMO procedures to our data-sets.

Referee: Although I know that automated and formalized workflows are still a rare commodity in atmospheric composition science (in contrast to operational air quality monitoring), and the authors of this article are therefore among the first to develop and implement such procedures consistently throughout a set of essential climate variables, this doesn't justify the lack of any deeper analysis or discussion of the procedures.

C3

To give just one example: if range checks are performed on different atmospheric variables with different frequency distributions: how are the thresholds determined and how robust are the error detection procedures in each case? Clearly, finding outliers in, for example, NO data is very different from finding outliers in ozone, CO₂, or CH₄ data.

AC: Probably we were not effective enough in describing this very important topic. In the case of a paper revision, we will be able to provide the full list of automatic checks for each ECVs and single sites, both based on the internal diagnostic of instruments and on the variability of measured components (i.e., spikes or outliers detection). For the latter issue, several specific tests were carried out to define the most appropriate threshold values (below, we report an example for CPC observations at CMN, the same was done for other aerosol parameters as well as NO_x). An example is shown by Figure 1. However, as stated in the manuscript, our procedures do not have the ambition to be "universal" (which is rather unfeasible considering the different span of atmospheric conditions that can affect measurements even in the same region), but they can be rather easily adapted to specific conditions and measurement sites. Thus, the thresholds used in the current prototype version of the system are specific for the ECVs and sites considered in our applications. Once the R procedures will be made available for the community, each single PI will modify them and adapt them to his need.

Referee: Also on the technical and data management side the paper lacks much important information, for example related to the documentation of responsibilities, resilience of the data transfer, provenance tracking and versioning.

AC: We think that all of these points are not within the scope of the paper. As mentioned in the manuscript, our system is a prototype (even if already running for a subset of ECVs in 4 observatories). For this reason, we did not provide all the information requested by the reviewer but, in the submitted version of the manuscript, we focus on the

C4

description of elaboration routines. In the case of paper revision, as requested by the referee, all these information will be provided and discussed for the running prototype version of the system.

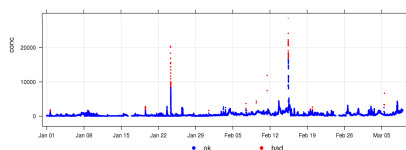
Referee: In its present form this paper should only be published as technical report at one of the institution's web site. It would have to be completely rewritten to merit publication in a scientific journal - even if this journal has a more technical scope.

AC: We are really surprised about the reviewer position. As stated in the AMT "Aims and scope", for AMT "the main subject areas comprise the development, intercomparison, and validation of measurement instruments and techniques of data processing and information retrieval for gases, aerosols, and clouds". Therefore, we think that this paper well fit with the aim of the ATM and we are confident that the explanations supplied in this letter together with the revision of the manuscript with some of the key points raised by reviewer will meet the interest of the scientific community.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-245, 2018.

C5

Reply to the Interactive comment on "Automatic procedures for submitting essential climate variables (ECVs) recorded at Italian Atmospheric Observatories to WMO/GAW data centers" by Luca Naitza et al. - Anonymous Referee #1 (Received and published: 26 October 2018)



An automatic data selection is applied to CPC raw data with a valid numflag (no calibration, no sampling failures/interruption). Following an approach similar to Huang et al. (2016, JORS). A sensitivity study was implemented to define the best agreement between manual outlier selection by skilled person and script results.

Fig. 1.

C6