WP 1.1 High-altitude climatic observation system and climate station network

D1.1F: High-altitude station network: full operation of the new measurement programmes, evaluation and quality control services and near-real time data delivery/early warning.

Contents

1. Introduction

The monitoring of trace atmospheric constituents in the lower troposphere still represents a fundamental activity to assess their long- and short-term variability, to investigate the underlying processes and to assess the impact of natural and anthropogenic sources into the atmosphere.

The Global Atmosphere Watch program of the World Meteorological Organization (WMO/GAW) coordinates a global network of surface stations to understand and control the increasing influence of human activity on the global atmosphere (WMO, 2017). The backbone of the WMO/GAW is a global network of more than 400 surface stations, performing routine observations of atmospheric constituents in the troposphere/stratosphere. Atmospheric stations belonging to the WMO/GAW network are requested to adopt standard operating procedures (SOP), to perform quality assurance/quality check actions (QA/QC) and routinely submit data to specific World Data Centers (WDCs) covering 6 focal areas (i.e., atmospheric aerosols, greenhouse gases, selected reactive gases, ozone, UV radiation, precipitation chemistry).

Nextdata was aimed at creating a network in mountain and remote areas, based on atmospheric observatories for the monitoring of atmospheric composition and ancillary data (meteorological parameters and solar radiation). The main goal of this network is to investigate the processes which influence the variability of air pollutants and climate-altering compounds to contribute towards a better assessment of related impacts to mountain ecosystems and climate in the Mediterranean basin. The network comprises five high-mountain atmospheric observatories: Monte Cimone (CMN, northern Apennines, the only WMO/GAW global station in Italy; 2165 m a.s.l.), Plateau Rosa (PRS, western Alps, WMO/GAW regional station; 3480 m a.s.l.), Col Margherita (MRG, eastern Alps; 2550 m a.s.l.), Monte Portella – Campo Imperatore (CMP, central Apennines; 2401 m a.s.l.), and Monte Curcio (CUR, southern Apennines, WMO/GAW regional station; 1796 m a.s.l.). In addition to these observatories, the WMO/GAW regional stations Capo Granitola (CGR, south-western Sicily) and Lampedusa (LMP, central Mediterranean Sea), provide complementary information on the background conditions of the Mediterranean basin marine boundary layer.

Six of these observatories are already part of international projects/research programs for the monitoring of Essential Climate Variables (ECVs). More specifically, in the framework of the WMO/GAW activities, observations of greenhouse and reactive gases are carried out at Plateau Rosa, Mt. Cimone, Mt. Curcio, Capo Granitola and Lampedusa (Table 1). Moreover, Plateau Rosa and Lampedusa also started the labelling processes to be included in the European Research Infrastructure ICOS (Integrated Carbon Observation System, http://www.icos-ri.eu, see Hazan et al., 2016), while Mt Cimone (thanks to the commitment of Nextdata) was the first atmospheric station in Italy which obtained the ICOS label. Measurements of physical properties of atmospheric aerosols are performed at Mt. Cimone, Mt. Curcio, and Capo Granitola. Mt. Cimone and Capo Granitola are part of ACTRIS-2 (Aerosols, Clouds and Trace gases Research Infrastructure, https://www.actris.eu/, see FMI 2017) which aims in consolidating high-quality observations of aerosols, clouds and trace gases in Europe. Col Margherita and Mt. Curcio were part of GMOS (Global Mercury Observation System), a global observational network providing comparable data on mercury levels in ambient air and deposition (see Cinnirella et al., 2014, Sprovieri et al. 2017). Within GMOS a web-based system for QA/QC has been developed in order to check raw data related to atmospheric mercury (D'Amore et al. 2015). During 2017-2020, both observatories Col Margherita and Mt. Curcio are involved in the iGOSP (Integrated Global Observing Systems for Persistent Pollutants, http://www.igosp.eu), strand 3 of ERA-PLANET project (Tsinganoset al., 2017) funded in the frame of the Horizon2020.

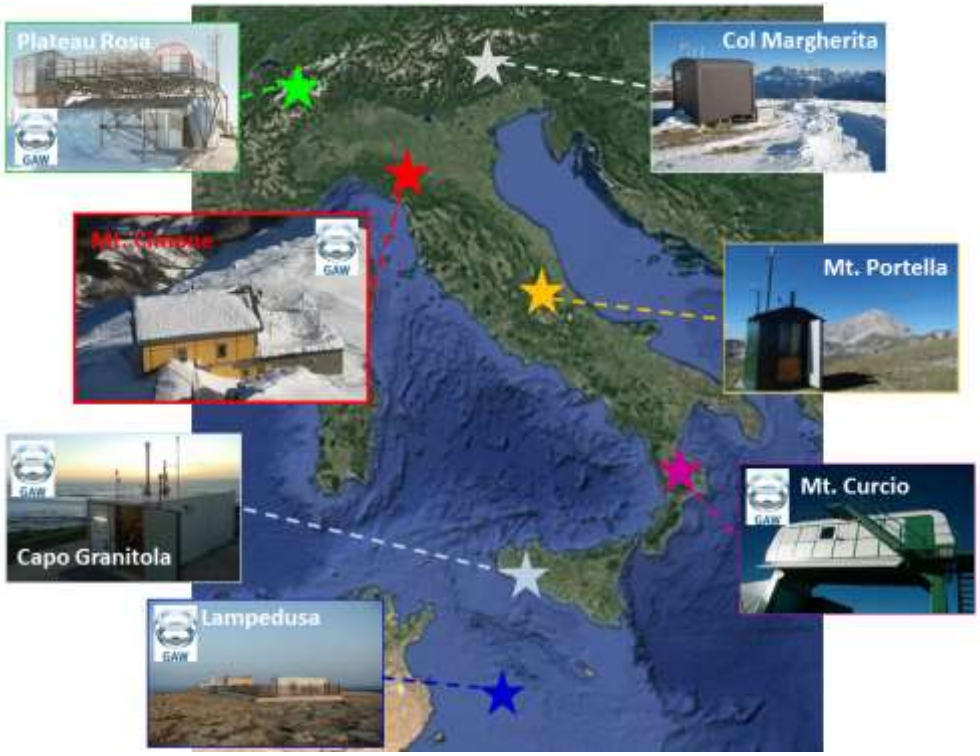

Figure 1. *Geographical location and pictures of the monitoring stations belonging to the atmospheric background observational network supported by the NextData project.*

Within NextData, the integration of the atmospheric background observational network passed by a number of activities:

- *The implementation of new measurement programmes at the background stations;*
- *The definition of guidelines to harmonize the measurement methodologies, protocols, QA/QC procedures, including the adoption of common reference calibration scales (e.g. GAW-WMO, ICOS, ACTRIS);*
- *The implementation of automatic procedures for submitting essential climate variables (ECVs) recorded at Italian Atmospheric Observatories to WMO/GAW data centres;*
- *The implementation of NRT data delivery and early warning services*

New measurement programmes for the monitoring of near-surface ozone were implemented at LMP and MRG, while a common calibration scale (referred to the WMO SRP#15 standard) have been implemented for the whole network. At CMN, new measurement of CO, CO_2 and CH_4 were upgraded by the adoption of a Cavity Ring Down Spectrometer (CRDS). The implementation of the related calibration and quality check procedures allowed CMN to be officially included in the ICOS-RI network on November 2018. At CMN, the DOAS system was upgraded to obtain atmospheric profiling capacity by using the MAX-DOAS technique, while the near-surface aerosol measurements were implemented by acquiring a new Scanning Mobility Particle Sizer system ACTRIS-compliant. Moreover, the PFR solar photometer for the determination of AOD during summer season was calibrated by the WMO/GAW central reference laboratory in Davos (PMOD). Measurements of NO and SO_2 at CGR and CMN were linked to the WMO scale by using certified air standard mixtures from NPL (UK). Meteorological observations were implemented at MRG, CMN, Chieti University and LMP.

Here we provide a final overview of the operational system for the evaluation and quality control of the ECVs recorded at the Nextdata observatories together with a description of the NRT data delivery and early warning services based on the continuous observations of these variables.

2. Automatic processing of ECVs: the evaluation and quality control (EQC) system

To contribute to the implementation of the observatory network, one action carried out by NextData was to set-up a suite of routines for the automatic processing of the data observed at these measurement sites, thus integrating the single measurement site in a common framework of data evaluation and creation, thus supporting the participation of these atmospheric observatories in the WMO/GAW activities. Besides making the data creation process faster and favoring a timely data submission, the adoption of standardized validation procedures will also assure a more subjective flagging of data, as well as the possibility to trace back the actions which led to data validation (i.e., data revisions will be easier). In particular, the ECVs covered by this work primarily focus on near-surface trace gases, aerosol properties and (ancillary) meteorological parameters which are under the umbrella of the World Data Center for Greenhouse Gases (WDCGG, see https://ds.data.jma.go.jp/gmd/wdcgg/), World Data Center for Reactive Gases and World Data Center for Aerosol (WDCRG and WDCA, see http://ebas.nilu.no).

The goal of this activity is to integrate the measurement network by supporting a more efficient data creation process, which is a pre-requisite for a fast and efficient data publication. Indeed, due to the large amount of data recorded by the measurement sites, it is not efficient to perform the data validation/flagging by the so-called "visual inspection" and by manual manipulation of data files. Moreover, there is a widely recognized need for the provision of ECVs in near-real time or real-time mode for a number of applications (data assimilation, atmospheric model verification, early warning systems, see, e.g., Wagner et al., 2015), which imply the delivery of quality-assessed data with well-defined data formats.

Table 1. *List of Essential Climate Variables (ECVs) collected at the atmospheric background monitoring observatories part of NextData. Grey-shaded cells indicate the ECVs for which automatic processing is already active. 1AOD measurements are processed in the framework of AERONET and GAW-NRT programs. In green, the measurement for which the automatic evaluation and quality control system was activated, are reported.*

ECV	**Reference programs**		**Monitoring stations**					
		PRS	***MRG***	***CMN***	***CMP***	***CUR***	***CGR***	***LMP***
CO_2	**WMO/GAW**	X		X		X	X	
CH_4		X		X		X	X	
CO				X		X	X	
O_3		X	X	X	X	X	X	X
SO_2				X		X	X	
NO	**WMO/GAW ACTRIS**			X		X	X	
NO_2				X		X	X	
Particle scattering				X		X	X	
Particle absorption				X		X	X	
Particle size distribution (by SMPS)				X		X		
Particle concentration number				X		X	X	
Coarse particle size distribution (by OPC)				X	X	X	X	
AOD[1]				X				X
Meteorological parameters	**WMO/GAW**	X	X	X	X	X	X	X
Solar radiation			X	X	X	X	X	X

It must be clearly stated that this action would not overcome the quality assurance/quality check activity carried out at the topical/thematic centers of integrated initiatives (e.g., ICOS, ACTRIS), nor at WMO/GAW WDCs.

Currently, the automatic processing of data is active for a subset of ECVs and measurement sites (see Table 2). In particular, the CMN observatory (Cristofanelli et al., 2018) was selected as a "proof-of-concept site", due to the large number of ECVs observed and the large variety of data formats produced by the measurement systems. Then, the procedures developed for CMN have been adapted to the other atmospheric observatories included to this exercise. Specific procedures have been developed for each ECV and for specific instruments (i.e., one processing chain for each instrument type) used by the considered stations, see Table 2.

The core of the present activity was the set-up of the automatic procedures but a centralized prototype system was implemented with the dual aim of demonstrating the routine effectiveness and supporting Italian observatories in the process of data production and submission. On a daily basis, raw data from measurement sites are transferred to a server located at CNR-ISAC HeadQuarters (HQs) in Bologna, for automatic data processing and storage. For these reasons, specific data delivery services have been activated from the Nextdata stations. The automatic data processing encompasses a preliminary harmonization of file formats, which is the pre-requisite for the subsequent data flagging, data aggregation (to common temporal frames: 1 and 60 minutes), and final harmonization of files according to WMO/GAW WDCs formats. Further details about the data model adopted, the management of near-real time data delivery from station to the elaboration server and about data flagging are provided in Deliverable D1.1B.

The automatic data processing also encompasses the creation of many data products (updated daily), which provide an overview of the instruments and data behavior, to support both the quality control of data, as well as the data inspection for scientific or operational purposes (i.e., the identification of events of interest or to perform preliminary data analysis); for more details see Deliverable D1.1B and the AMTD paper by Naitza et al. (2018).

Table 2. *Summary of automated procedures developed for the automatic processing of ECVs as a function of specific instruments and data file format for submission to WMO/GAW WDC.* Automatic processing still not operative but running "on-demand". For explanation of "data file formats" see Section 2.4*

Class	ECV	Instruments	Data file format
Trace gases (near-surface)	Ozone	Thermo 49i, Thermo 49c	NASA-AMES
	NO, NO_2	Thermo 42iTL	NASA-AMES
	SO2	Thermo 43iTL	NASA-AMES
	*CO	Picarro G2401	Ascii (WDCGG)
	*CO_2, CH_4	Picarro G2401	Ascii (WDCGG)
Aerosol (near-surface)	Aborption coefficient	MAAP 5012	NASA-AMES
	Aerosol scattering	TSI 3563	NASA-AMES
	Number particle concentration	TSI 3775	NASA-AMES
	Size distribution	OPS (Grimm 1.108)	NASA-AMES
Meteorology	T, RH, P, WD, WS Global, UV	Various	NASA-AMES

At the current stage, the automatic routines are active for Monte Cimone (CMN), Col Margherita (MRG), Monte Portella–Chieti University (CMP), Capo Granitola (CGR) and Lampedusa (LMP), see Table 1. Institutions in charge of ECV observations at Plateau Rosa (PRS) and Mt. Curcio (CUR), albeit involved in the development of EQC routines, did not operationally implement the system.
A dedicated web site (https://nextdata.bo.isac.cnr.it/) was implemented for supporting measurement PIs in easily accessing the whole suite of generated data products (as a function of the different measurement sites and ECVs).

2.1 Workflow of EQC system

To increase the inter-operability of the automatic procedures, the sequential steps of the workflow are the same for all the different ECVs and specific modules or functions have been developed to be inserted in the programming code to increase flexibility of usage:

1) data file collection from observatories
2) data formatting of raw files coming from stations;
3) data check and flagging;
4) data correction (if needed);
5) data aggregation (time averaging) and flagging;
6) data formatting;
7) creation of data products.

By this 7-steps process, three different data levels are produces according with WMO/GAW WCDRG and WCDA data reporting guidelines (see also https://ebas-submit.nilu.no/Submit-Data/Data-Reporting):

- Level-0: annotated raw data; format instrument specific; contains all parameters provided by the instrument; contains all parameters/info needed for processing to final value; "native" time resolution;
- Level-1: data processed to final parameter (calibration and correction implemented to data series), invalid data and calibration episodes removed, "native" time resolution, normalization to standard temperature and pressure (i.e., 273.15 K, 1013.25 hPa) if necessary;
- Level-2: data aggregated to hourly averages, atmospheric variability quantified by standard deviation or percentiles.

In the present prototype system, the automatic processing is executed by a set of four scripts specifically designed for each instrument/ECV and for each measurement site:

- “P10” is the script devoted for the formatting/homogenization of raw data files
- “P20” is the script devoted to the production of Level-0 data files
- “P21” is script devoted to the production of Level-1 and Level-2 data files
- “P22” is the script devoted to the generation of data products .

These scripts are run sequentially basing on the above order list starting at 2:00 UTC of each calendar day. For each observatory and ECV, the elaboration is started by executing the “P10” scripts. Then, after an appropriate lag time, scripts “P20”, “P21” and, finally, “P22” are executed for each single observatory and ECVs. They elaborate data file produced at the atmospheric observatories since the begin of the current year (i.e. January, 1st) until the day before. Finally, a diagnostic is launched to check the correct execution of the script. In the current configuration (24 ECVs from 5 observatories), all the processing chains are completed in 4 hours.

2.2 Automatic processing of ECVs: prototype of a centralized system

All the developed routines are virtually stand-alone and any hypothetical user, after installing an "R" environment, can use them on its own PC (both Linux or Window) or server for automatic and on-demand application. However, with the purpose of demonstrating the effectiveness of the proposed routines and supporting Italian observatories in the process of data production and submission, a centralized prototype system was implemented for sustaining the automatic processing of ECVs at the atmospheric observatories involved within the Nextdata project.

While, as detailed below, the delivery of data file from the atmospheric observatories is a duty in charge of observatory personnel/Institutions, the automatic operation of the routines is in charge of a small group of people at the CNR-ISAC HQs in Bologna, involving an IT expert and two routine developers. In the case the raw data format are not changed, the system is expected to be robust and not frequent technical intervention are needed. When, for whatever reason, data flow from observatories is interrupted the routines are still running and data are considered as missing. The case of delayed data delivery does not represent an issue, since the routines are designed to elaborate all the data file generated in the current year.

A specific product (called "Health Status Report") to check the correct execution of the routines (from "P10" to "P20") was designed to be accessed by measurement PIs or CNR-ISAC personnel. This product is generated on a daily basis and the plots indicate, for each observatory, the correct execution of elaboration routines (see Figure 3 for an example): if a routine worked successfully, a bar is drawn (in case of routine failure, the related bar is not plotted).

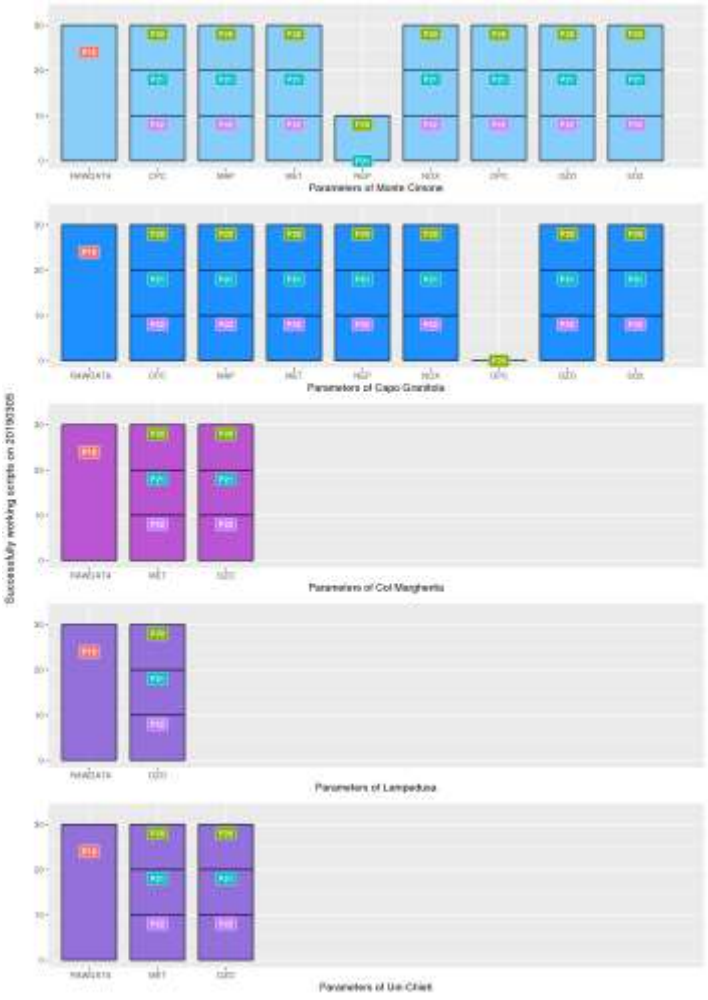

Figure 2. Example of "Health Status Report" describing the execution status of the EQC routines. For each observatory (CMN, CGR, MRG, LMP and Chieti University) each bar denotes the correct execution of routine "P10" (script for the generation of formatted raw data), "P20" (script for the generation of Level-0 data), "P21" (script for the generation of Level-1 and Level-2 data), and "P22" (script for the generation of graphic products). As shown in this example, the "P21" and "P22" routines failed for CMN on 2019, March 5th (the date is reported on y-axis).

2.3 Data product

A dedicated web site (https://nextdata.bo.isac.cnr.it/) was implemented for supporting measurement PIs in easily accessing the whole suite of generated data products (as a function of the different measurement sites and ECVs).

This web site is presenting all the data products generated by the EQC system (see deliverable D1.1B for a complete list and description of these products). The goal of the website is dual: firstly, to provide easy and direct access to information related to the correct execution of measurements, secondly to support reporting activity related to the description of current atmospheric conditions at the measurement sites and about occurrence of special events.

The web interface is directly linked with the elaboration server where the data products are generated and stored, allowing the automatic update of visualized data products. Data products are organized as a function of measurement sites, ECV, specific variable and product classes (Figure 3 and 4).

Figure 3. Home of the EQC system web site, where data product for single stations can be accessed.

Figure 4. Data product "search" interface. For each station (e.g. Lampedusa) ECV, variable, and specific data product can be selected (e.g. quarterly diurnal variability of near-surface ozone).

2.4 Routine availability

All the procedures have been implemented in "R" Language and Environment for Statistical Computing (R Development Core Team, 2015). These routines are freely downloadable from the Nextdata data repository (http://geonetwork.igg.cnr.it.): a total of 24 scripts were coded for 8 considered ECVs). Moreover, 3 scripts for the management of the numflag according with nasa-ames and GAW-WDCRG, GAW-WDCA policies, have been provided.

3. Implementation of NRT data delivery and early warning services

The activation of the automatic procedures for the flagging and formatting of ECV data files from Nextadata observatories, provided the opportunity to implement the near-real time services already presented in the deliverables Deliverable D1.1A and Deliverable D1.1B.

In particular, two different data streams were implemented for:

(1) near-surface aerosol (i.e., eqBC and total particle number concentration) in the framework of ACTRIS/ GAW NRT

(2) reactive gases (i.e., O_3, NO, NO_2, SO_2) in the framework of CAMS-Copernicus.

The two data streams are based on different routines and delivery strategies which depend on the specific requirements from the two programs. For ACTRIS/GAW NRT, a specific routine (based on "R" scripts) was implemented to elaborate (as an instance a specific data header is created) the raw data files produced from several instruments (i.e., NEPH, CPC, and MAAP) and manages the automatic delivery to a SFTP server every 30 minutes.

The data delivery service was activated towards CAMS-Copernicus for the purpose of on-line model verification. In this case, a "bash" script was used to extract the requested information from the level-2 files (i.e. hourly averages) of reactive gases and to organize them in the formats requested from CAMS-Copernicus for the submission. In particular, CAMS-Copernicus requests that daily files containing hourly averages and related standard deviation of the ECVs of interest are delivered with (at least) a daily frequency. The same script manages both the appending of the data header as well as the daily delivery of data files to a dedicated SFTP server.

Currently, the ACTRIS/ GAW NRT data stream is active for CMN station, while the CAMS-Copernicus data stream is active for CMN, CGR, PRS, LMP and Chieti University.

4. References

Carslaw, D.C. and K. Ropkins, (2012). openair — an R package for air quality data analysis. Environmental Modelling & Software. Volume 27-28, pp. 52–61

Cinnirella S., D'Amore F., Bencardino M., Sprovieri F., Pirrone N. (2014) The GMOS cyber(e)-infrastructure: advanced services for supporting science and policy. Environmental Science and Pollution Research, 21(6): 4193-4208. doi: 10.1007/s11356-013-2308-3.

D'Amore F., Bencardino M., Cinnirella S., Sprovieri F., Pirrone N. (2015) Data quality through a web-based QA/QC system: implementation for atmospheric mercury data from the global mercury observation system. Environmental Science: Processes & Impacts, 17: 1482-1491. doi: 10.1039/C5EM00205B

Finnish Meteorological Institute, ACTRIS Stakeholder Handbook 2017, ISBN 978-952-336-049-5 (pdf version), 2017.

Hazan, L., Tarniewicz, J., Ramonet, M., Laurent, O., and Abbaris, A.: Automatic processing of atmospheric CO2 and CH4 mole fractions at the ICOS Atmosphere Thematic Centre, Atmos. Meas. Tech., 9, 4719-4736, https://doi.org/10.5194/amt-9-4719-2016, 2016.Naitza, L., Putero, D., Marinoni, A., Calzolari, F., Roccato, F., Busetto, M., Sferlazzo, D., Aruffo, E., Di Carlo, P., Bencardino, M., D'Amore, F., Sprovieri, F., Pirrone, N., Dallo, F., Gabrieli, J., Vardè, M., Barbante, C., Bonasoni, P., and Cristofanelli, P.:

Automatic procedures for submitting essential climate variables (ECVs) recorded at Italian Atmospheric Observatories to WMO/GAW data centers, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-245, in review, 2018.

Sprovieri, F., Pirrone, N., Bencardino, M., D'Amore, F., Angot, H., Barbante, C., Brunke, E.-G., Arcega-Cabrera, F., Cairns, W., Comero, S., Diéguez, M. D. C., Dommergue, A., Ebinghaus, R., Feng, X. B., Fu, X., Garcia, P. E., Gawlik, B. M., Hageström, U., Hansson, K., Horvat, M., Kotnik, J., Labuschagne, C., Magand, O., Martin, L., Mashyanov, N., Mkololo, T., Munthe, J., Obolkin, V., Ramirez Islas, M., Sena, F., Somerset, V., Spandow, P., Vardè, M., Walters, C., Wängberg, I., Weigelt, A., Yang, X., and Zhang, H.: Five-year records of mercury wet deposition flux at GMOS sites in the Northern and Southern hemispheres, Atmos. Chem. Phys., 17, 2689-2708, https://doi.org/10.5194/acp-17-2689-2017, 2017

Tsinganos K. et al, 2017. ERA-PLANET, a European Network for Observing Our Changing Planet, Sustainability 9(6), 1040; https://doi.org/10.3390/su9061040

Wagner, A., Blechschmidt, A.-M., Bouarar, I., Brunke, E.-G., Clerbaux, C., Cupeiro, M., Cristofanelli, P., Eskes, H., Flemming, J., Flentje, H., George, M., Gilge, S., Hilboll, A., Inness, A., Kapsomenakis, J., Richter, A., Ries, L., Spangl, W., Stein, O., Weller, R., and Zerefos, C.: Evaluation of the MACC operational forecast system – potential and challenges of global near-real-time modelling with respect to reactive gases in the troposphere, Atmos. Chem. Phys., 15, 14005-14030, https://doi.org/10.5194/acp-15-14005-2015, 2015.