Recommendations on Rapid Data Delivery
GAW - MACC GRG subgroup meeting
UFS Schneefernerhaus, Grainau
7 October 2009

List of the discussion session participants:
Martin Schultz (FZJ), Claire Granier (UPMC-SA), Nellie Elguindi (CNRS), Werner Thomas (DWD), Harald Flentje (DWD), Paolo Cristofanelli (GAW site Monte Cimone, CNR), Oksana Tarasova (WMO) and Anja Werner (DWD).

Topics of the discussion:
1) Timing of the data delivery
2) Data Quality Objectives for model validation
3) Data format
4) Data policy
5) Mechanism of data submission
6) Feedbacks to the data providers/services
7) Agreements between GAW stations and MACC project

Rapid delivery strategy

The emphasis of the current GAW Strategy is mainly on surface monitoring of atmospheric composition. However, the mandate of the GAW programme includes the integration of satellite and aircraft observations with surface measurements, as well as the integration of chemical data and numerical models. Chemical reanalysis and data assimilation in the chemical weather forecast models are among the methodologies that integrate observations and models. In response to requests to integrate chemical and meteorological data and models, several projects have recently been initiated, and MACC: "Mapping Atmospheric Composition and Climate" is among the most successful ones. The MACC consortium of about 40 partner institutes will further develop a global operational medium range forecast/assimilation capability for atmospheric dynamics and composition, using all available satellite and in-situ data.

Near-real-time (NRT) data availability from GAW stations is a key element in validation of the model forecasts for reactive gases and aerosols and in the improvement of weather forecasts. As it has been shown within the GEMS project, intensive validation is very helpful for verifying geographical regions and also for process descriptions where the models need to be improved. Continuation and further enhancement of this service is therefore vital for making this task successful.

The Global Atmosphere Watch (GAW) network was important for the success of the GEMS project and will become a key element within MACC, provided that GAW data will be available in NRT. Currently several GAW stations provide observations in NRT mode, including Zugspitze/Hohenpeissenberg (Germany), Mt. Cimone (Italy), Izaña and Santa Cruz (Spain), Cape Point (South Africa), Tamanrasset/Assekrem (Algeria), Neumayer (Antarctica) and Moussala (Bulgaria). The results of this initiative were presented during the session earlier day.

Participation of the other GAW stations in NRT data delivery will strongly enhance the effectiveness of GAW as a key in-situ atmospheric surveillance network.

The Group Reactive Gases (G-RG) coordinator of the MACC project and other G-RG
team members confirmed the possibilities to return a valuable service to the GAW station managers, which can help to motivate the stations to contribute to the validation of the MACC global reactive gases forecasting system.

An offer was made to set up a web site where the GAW members can see plots of their data in comparison to the MACC model results and in relation to other stations. Nevertheless, the main incentive for GAW station managers should be to have MACC and the future atmospheric service as a key user of their data and the better integration of the GAW stations into the global earth observing system.

It was pointed out that clear requirements and rules must be formulated from both sides. A more formal service level agreement or at least a memorandum of understanding could also help GAW station managers to secure national funding for the measurements needed in MACC G-RG and the rapid data delivery.

As an outcome of the discussion the following requirements and obligations have been proposed.

**Recommendations**

1. **Terminology**
   A discussion on the terminology of data delivery timeliness took place. Near-real time delivery means that the data are accessible within one model step (less than 12 hours). Taking into consideration the Quality Assurance system which is applied to the measurements within the GAW programme data delivery of reactive gases is for the most part only possible with a longer delay.

   | Near-real time: | submission within 12 hrs a |
   | Delayed mode:   | submission later than 12 hrs |
   | Rapid delivery: | 1 day threshold (level 1) up to 1 week (level 2) |

   a data suitable for assimilation

   It was agreed that data submission of GAW station data on reactive gases should take place in the regime of a “Rapid Delivery” (RD), in the ideal case within 1 day, thus, immediate validation of forecasts is possible. Data submitted within one week can be used for fast model validation and is also valuable.

   Data submission later than 1 week should be done in the standard way to the World Data Centre for Greenhouse Gases.

   RD delivery has different requirements to the data quality and data flagging in comparison with a standard submission. Submitted in the regime of RD data are considered as preliminary while the data stored in the WDCGG are considered as high quality final data.

2. **Substances of interest**

   As an initial step, surface O₃ and CO will be the substances of interest for RD. These substances are measured at many GAW stations and have a well established system of quality control. However, in the near to midterm future SO₂, NOₓ and VOCs are desirable. These species are also very valuable but can initially be delivered with longer delays (focus on the evaluation of reanalyses).
3. Sites of interest

All continuously operating GAW stations are requested to submit data in RD mode to the MACC GAW-Validation group. It is desirable that the set of stations give global coverage. As a first step it is recommended that Global GAW and selected Regional stations participate in the initiative.

4. Data resolution and Data quality requirement

Only continuous measurements are considered for participation in the RD regime. Data are expected to be submitted as 1hr averages. No intensive data quality control has to be performed except the standard checks of the measuring equipment according to the Standard Operating Procedures (SOPs) or Measurements Guidelines (MGs) for the respective gases.

For submission acceptable uncertainties are:
- for surface ozone uncertainty of 15% (±5ppb) is acceptable
- for CO uncertainty of 15% (±10ppb) is acceptable.

5. Flagging of data

A simple flagging system should be applied to the data. “Good” data are to be flagged with a “0” (zero) and possible questionable data (any problem) are flagged with an “1”. Missing values are labelled with “-999”.

6. Data format

For submission the format of the World Data Centre for Greenhouse gases (WDCGG) is applied (see WDCGG guide, version 10, Annex 2, p.11).

7. Data Policy

The following main principles are applied to the data access/use submitted to the data portal:

a. Data are exclusively used for fast validation of MACC forecast products.

b. The access to the data server is restricted. Data files are only accessible by the respective partner site and the MACC GAW validation team (MGvt).

c. One month after submission, the data can be deleted from the server by request of the respective partner site.

d. The MACC validation team will regularly check WDCGG for final or updated data sets. In case final or revised data sets are available, the “fast delivery” data set will be deleted in agreement with the respective GAW partner site. These fully calibrated and quality controlled data sets shall replace the “fast delivery” data sets for additional validation of archived forecast or reanalysis products.

e. Validation products (plots) will be available for the GAW and MACC community.

f. In the case of data use for analysis of “interesting episodes” or any other scientific application, co-authorship on the respective publication should be offered to the data providers.

8. Data server

Submission is done by ftp transfer to the following controlled DWD ftp server:
A valid email address is necessary to register for the GAW FTP account; the registration procedure and terms of use are described at the above given link. Once registered, the upload of data is possible.

9. Availability of data on WMO’s GTS/WIS

The meta data on the data submitted to the DWD ftp site should also be made available on WMO’s Global Telecommunication System (GTS) and in the future on the WMO Information System (WIS). It should be made available in such a way that the data is searchable and discoverable but not accessible, unless permission has been given by the data originator. DWD makes sure to translate the submitted data files to a suitable WIS-compliant format such as CREX or BUFR and to get it posted on the GTS/WIS so that it can be discovered.

10. Tasks of the MACC GAW validation team (MGvt) and provided services

The main task of the MGvt is the use of GAW data for validation of MACC G-RG NRT forecasts produced with IFS – Mozart, IFS-TM5 or MOCAGE and later with the C-IFS system currently under development.

The data will be processed with some simple statistical filters to remove outliers. The results of this automated quality check (and the software tool to perform these checks) will be made available to the respective partner site.

For the partners submitting the data the following products will be provided:
- global/regional forecasts
- validation products (plots)
- model datasets for the GAW station location

The details of the product layout and product delivery need to be discussed and should be iterated with interested GAW station managers.

11. Letter of commitment

A letter of agreement will be signed between a representative of each participating GAW site and the coordinator of the MACC project. With the signature both parties agree to follow these rules and the respective GAW site becomes a contributing partner of the MACC project.

Upon request the WMO Secretariat will provide a letter of support to the respective GAW station participating in this initiative to help securing funds for continuation of measurements and quality assurance.