Ensemble operational air quality assessments in Europe - Improving modeling platforms with statistical analysis

Anthony Ung, Laure Malherbe, Frédérik Meleux, Bertrand Bessagnet, Laurence Rouil

Ensemble operational air quality assessments in Europe – Improving modeling platforms with statistical analysis

Anthony Ung, Laure Malherbe, Frederik Meleux, Bertrand Bessagnet, Laurence Rouil and MACCII modeling team
INERIS institut, Paris, France
Corresponding author: Anthony.ung@ineris.fr

Abstract

MACC-II - Monitoring Atmospheric Composition and Climate - is the current COPERNICUS (previously known as GMES Global Monitoring for Environment and Security) - Atmosphere Service throughout Europe. For regional air quality, seven regional modeling teams combine state-of-the-art atmospheric modeling with Earth observation data assimilation. All models use the same anthropogenic emissions dataset, the same meteorological forcing and the same lateral boundary conditions provided by the global model forecasts. They differ by their physical process modeling, data assimilation or statistical analysis system. All models have also very significant and specific skill, and this is used in the ensemble approach. The ensemble is currently based upon a median value approach. In this communication, we give an overview of the current status of the ensemble approach and its performance. The performance is based upon ground-based measurements which are not used in the data assimilation or statistical analysis process. MACC-II products are considered as the "best available or most realistic" representations of air pollution patterns over Europe: more relevant than interpolation of observations and more accurate than raw simulations issued from the ensemble or a unique model.

Keywords: Air quality, modeling system, ensemble approach, validation with ground based measurement.

1. Introduction

MACC-II - Monitoring Atmospheric Composition and Climate (http://gmes-atmosphere.eu) - is the current COPERNICUS (previously known as GMES Global Monitoring for Environment and Security) - Atmosphere Service throughout Europe. Two operational services based on the strength of multi-model approaches are dedicated to air quality monitoring in Europe: the first one focuses on routine and near real time products (forecasts, Near Real Time analyses) and the second one on detailed analysis of the past period. Seven regional modeling teams combine state-of-the-art atmospheric models with Earth observation data assimilation. Table 1 summarizes the fixed specifications and shared input data by all models. All models use the same anthropogenic emissions dataset, the same meteorological forcing and the same lateral boundary conditions provided by the global model forecasts. This choice has been made in order to limit the spread in the model results and allow some interpretation of the differences. Nevertheless, the variability among the models is still quite large and highlight the most sensitive geographical areas where the re-analyses are still uncertain, either because of model weaknesses (uncertainties on emission, complex topography, limit of the dynamic and chemical parameterizations...) or because of a lack of available representative observations. Each team is free with the choice of data assimilation strategy (table 2) provided that the results are relevant, reliable and improve significantly the raw simulation results. The list and the current configuration of the modeling systems are detailed in QA/QC dossiers and available on the MACC project website for each model. All models have also very significant and specific skill and functionalities which allow to derived “ensemble” model estimations. This ensemble is currently based upon a median approach to obtain an average with improved skills compared to the individual models at no additional cost. Five years are available from 2007 to 2011. Figure 1 illustrates the ensemble yearly mean of the regulatory air pollutant concentrations (O_3, PM_{10} and NO_2) at surface level for year 2010.

Characteristics	**MACC2 specifications**
Domain	GEMS (Europe) Extended domain to include Turkey and Island
Spatial Resolution	10 km
Vertical domain (height, nb of levels)	Free
Emission data	Improved MACC2 emission inventory + forests fire emissions
Boundary conditions	MACC2/GRG + MACC2/AER to include global aerosol products
Meteorological Re-analysis	IFS (ECMWF)
Chemical Pollutants Computed	O3, NO2, PM10, PM2.5
Chemical schemes	Free
Data assimilation approach	Free
in-situ data	AIRBASE (mandatory)
Earth observations	Experiments with satellite observations

Table 1: Fixed specification and shared input data by all models

Code	**Model**	**Data assimilation chain**	**Observation data assimilated or tested in MACC**
CHM	**CHIMERE**	Kriging and Optimal Interpolation at ground level	Only in-situ data from AIRBASE
EMP	**EMEP**	No operational data assimilation	
RIU	**EURAD-IM**	4DVAR	AIRBASE in situ measurements MOSAIC air borne in situ measurements NO2 tropospheric column retrievals from OMI, GOME-2, SCIAMACHY, MOPITT CO profiles PM10 SYNAER retrievals
KNM	**LOTOS-EUROS**	Ensemble Kalman filter	AIRBASE in-situ data for O3 and PM10 OMI for NO2, evaluation studies
SMH	**MATCH**	3Dvar with transform into spectral space	AIRBASE in-situ data Satellite information : retrieval ongoing from OMI, AIRS, MODIS, but not used operationall
MFM	**MOCAGE**	3Dvar operational chain	AIRBASE in-situ data Satellite NO2 (OMI) and CO (IASI from ULB/LATMOS or EUMETSAT ; MOPITT from NCAR)
FMI	**SILAM**	3Dvar operational chain	AIRBASE in situ data
ENS	**Ensemble**	Median	

Table 2: Specification of the data assimilation chains run in MACC

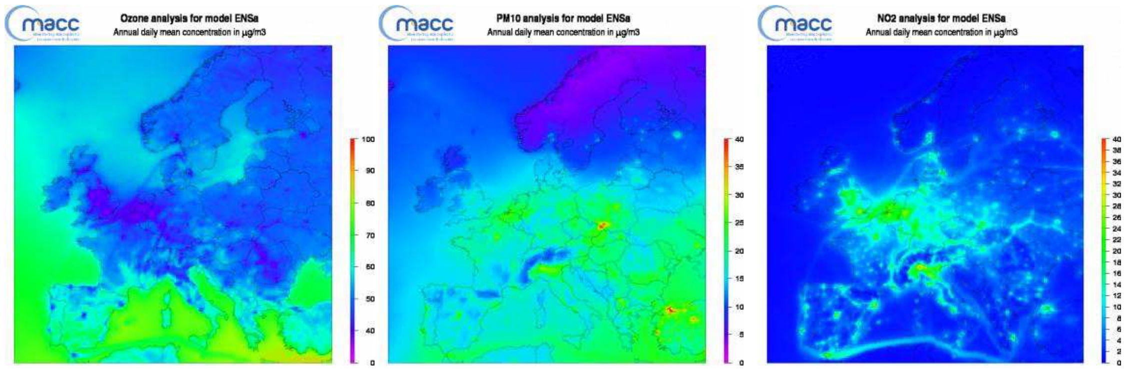

Figure 1: Ensemble 2010 yearly concentration for O_3, PM_{10} and NO_2 at surface level.

2. Evaluation of the models and the ensemble

The evaluation and extensive analysis of the performances of the modeling systems and of the ensemble to predict the concentrations are fully documented on a yearly basis in the evaluation and assessment report (Rouil L., 2011 & 2010). It is a systematic evaluation of the model, of the data assimilated results and of the ensemble model against a relevant set of dedicated observation data (from the European Airbase database).

MACC/EVA re-analyses are appropriate to report and comment on background air pollution (both rural and urban). The spatial resolution adopted in the current system does not allow catching local pollution situations, especially near busy roads or on industrial sites. Only observation data with “background” typology are considered.

This observation set is split into two sets: a set of stations for assimilation procedure and a set of stations for validation procedure (figure 2). Stations are randomly distributed and give a good spatial cover over European domain.

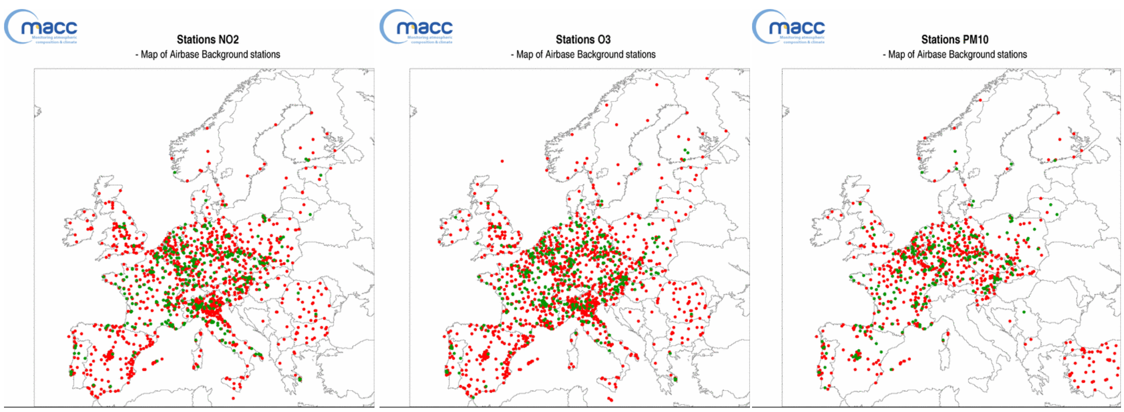

Figure 2: assimilation (red) and validation (green) stations for NO_2, O_3 and PM_{10} pollutants.

Each model performance is evaluated on the basis of classical statistical indicators which measure objectively the gap between the model results (raw data or re-analyses) and the observations at the available stations: bias with the observed mean value, root mean square error (RMSE) and correlation coefficient are the most classical. Figure 3 represents the Taylor diagrams for regulatory indicator of air quality: daily maximum concentration of ozone, daily mean concentration of dioxide nitrogen and daily mean concentration of PM_{10} particulates.

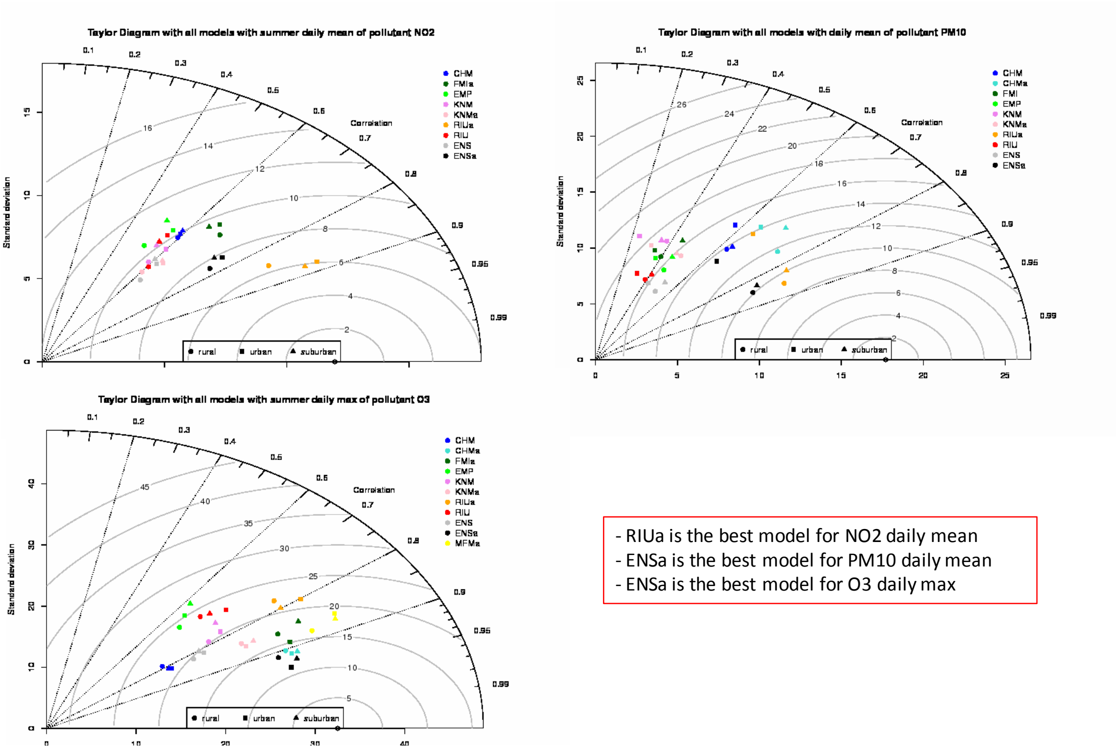

Figure 3: Taylor diagrams for all models for PM_{10}, O_3 and NO_2 with and without assimilation.

The graphics illustrate the superiority of the median ensemble model (black dots) in comparison to individual models. Each model contributes to build up the skill of the ensemble median. For NO_2, there is a limited number of available re-analysis. It results that the best model is not the ensemble but one of the member of the ensemble. Same comparison is conducted on a yearly basis since 2007 and similar results are obtained every year.

As part of the MACC-II work plan, various alternatives to the median ensemble model have been tested (Moinat P., Peuch V.H., 2012). A wide screening of methods based at exploiting the historical knowledge of which model(s) performed “best” overall over the few last days did not give convincing results. The studies have also shown that the ensemble model takes actually benefits from the different models (including the ones with less skill), and that the “best” model, in terms of proximity to observations, is in practice changing very significantly in space and time.

From one year to another, the number and the selection of the models to compute the median can evolve, but limited to 4 or 5. As a consequence, some indicators such as the yearly mean concentration of the ensemble are dependent variables.

3. Conclusion

The availability of a set of several model results allows a qualitative estimation of the model uncertainty which is linked to the variability of the model results. Southern Europe and Eastern Europe are the European sub-regions where the model responses are the most variable and uncertain, even with the data assimilation systems. This is not a surprising result; the difficulties for simulating the dynamic and chemical atmospheric processes that occur in these geographical areas are well-known from the scientific communities and are still the subject of research project. Moreover lack of measurements in Eastern Europe (compared to other geographical areas, see in annex) can also result in poor performance when modeled indicators are compared to observations. But this must be interpreted with caution.

MACC-II products are considered as the "best available or most realistic" representations of air pollution patterns over Europe: more relevant than interpolation of observations and more accurate than raw simulations or one unique modeling. In its operational stage (foreseen in 2014) the GMES/COPERNICUS atmospheric services will propose complementary and comprehensive information established on the basis of state of art chemistry-transport models run operationally by modeling teams with a long experience in the field of air pollution, and extended in-situ and satellite observation sets gathered and made available. Such information could be relevant for assessing rural and urban background concentrations and exposed areas to be reported according to the air quality Directive of the European Commission (2008/50/EC).

The work is done with contribution from the MACC regional modeling teams.

- CERFACS, FRANCE: Sébastien Massart
- CNRS/LISA, FRANCE: Matthias Beekman, Gilles Foret
- FMI, FINLAND: Mikhail Sofiev, Julius Vira
- KNMI, NETHERLANDS: Henk Eskes
- Meteo France, FRANCE : Vincent-Henri Peuch, Virginie Marécal
- Met.no, NORWAY: Alvaro Aldebenito, Michael Gauss
- FRIUKK, GERMANY: Hendrik Elbern, Elmar Friese, Achim Strunk
- SMHI, SWEDEN: Lennart Robertson
- TNO, NETHERLANDS: Arjo Segers, Lyana Curier
- INERIS, FRANCE : Laure Malherbe, Frederik Meleux, Bertrand Bessagnet

Analyses and 3-day forecasts of European air quality are provided through the MACC-II website in near-real-time: http://www.gmes-atmosphere.eu/services/raq/raq_nrt/

References

Moinat P., Peuch V.H., (2012) *Report on studies concerning alternative ensemble approaches,* report D_R-ENS_3.2 of MACC-COPERNICUS Project. http://www.gmes-atmosphere.eu/documents/deliverables/r-ens/

Rouil L. (2011, 2012). EVA assessment reports, EVA evaluation reports. http://www.gmes-atmosphere.eu/documents/deliverables/r-eva/

MACC QAQC dossiers : http://www.gmes-atmosphere.eu/documents/deliverables/rens/