Report D_R-ENS_1.5.3

September 2010

MATCH

Products, Quality and Background information

Contributors :

Lennart Robertson (SMHI, #17)

Editors :

Hendrik Elbern, R-EDA leader (FRIUUK, #21)
Vincent-Henri Peuch, R-ENS leader (METEO-FR, #15)
Laurence Rouïl, R-EVA leader (INERIS, #23)

Contact :
Lennart.Robertson@smhi.se

SEVENTH FRAMEWORK PROGRAMME

MATCH facts sheet

1. Products portfolio

Name	Description	Freq.	Available	Species	Time span
FRC	Forecasts at the surface, 500m, 1000m, 3000m and 5000m above ground	Daily	20 UTC the day before first valid day	O_3, NO, NO_2, CO, SO_2, $PM_{2.5}$, PM_{10}	0-72h, hourly

2. Performance statistics

The model did not produce useful results in some days in October, November and December 2009. In the first two occasions this was attributed to upgrades of the GRIB package that had to be corrected. In December the model was switched of for restart that was accidently not made.

3. Availability statistics

Quarter 4 (March, April, May 2010) : 100% of the MATCH forecasts have been provided in time.

Quarter 5 (June, July, August 2010) : 95.6% of the MATCH forecasts have been provided in time.

MATCH data were missing for 3, 6, 7 and 8 August 2010.
The reason is most likely a routine crash most due to a too small dimension of a work area needed for reading MACC emissions at high resolution. Action is made by setting default size so that recompilation always returns working executables.

4. Assimilation and forecast system: synthesis of main characteristics

Assimilation and Forecast System	
Horizontal resolution	0.5°
Vertical resolution	40 levels
Gas phase chemistry	Based upon EMEP (Simpson et al., 1993), with modifications for isoprene production (Carter, 1996; Langner et al., 1998).
Heterogeneous chemistry	No
Aerosol size distribution	0.02-0.1, 0.1-1, 1-2.5, 2.5-10
Inorganic aerosols	Sulphate, Nitrate, Ammonium
Secondary organic aerosols	No

Aqueous phase chemistry	Sulphur oxidation by H2O2 and O3
Dry deposition/sedimentation	Resistance approach/size dependent sedimentation velocity, in link with PBL parameterisations
Mineral dust	Yes
Sea Salt	Yes
Boundary values	MOZART IFS forecast for the day before
Initial values	MATCH 24h chemical forecasts from the day before
Anthropogenic emissions	TNO (2000) inventory with 0.25°x0.125° resolution (0.5°x0.25° resolution for shipping emissions)
Biogenic emissions	Simpson et al., 1993
Forecast System	
Meteorological driver	12:00 UTC operational IFS forecast for the day before (0.5°, 40 levels)
Assimilation System	
Assimilation method	3Dvar, (EnKF)
Observations	NRT, (satellite)
Frequency of assimilation	Performed once a day for the previous day and hour by hour
Meteorological driver	IFS forecast and analyses for the same day (0.5°, 40 levels)

Evolutions in the MATCH suite

2010/10/30 : The 3Dvar code is operational
2010/10/22 : Chi-2 test included in the 3Dvar code
2010/06/15 : Pygrib developed for postprocessing purposes
2010/09/03 : Variational quality control in 3Dvar included
2010/03/12 : Speeded up reading of NETCDF files
2010/02/10 : Code for BUFR handling revised
2009/12/14 : The 3D variational scheme in a first parallel version
2009/12/11 : Observations in BUFR format implemented I 3D variational code
2009/12/10 : Delivered hind casts for August 2007
2009/12/01 : Export of results to MARS at SMHI
2009/12/01 : Start of background error calculations
2009/12/01 : New general GRIB1/GRIB2 code developed
2009/11/01 : Adaptations for full global boundaries
2009/10/10 : Updated aerosol module with optical characteristics
2009/10/10 : Aerosol optical parameters made available to end user
2009/07/04 : Installation of version 4.8.6
2009/06/01: Start of MACC pre-operational forecasts

MODEL background information

1. Forward model

The Multi-scale Atmospheric Transport and Chemistry model (MATCH) is an off-line model with a rather flexible design, account for various forcing weather data on different resolutions and projections, and different sets of schemes for deposition and chemistry.

In MACC MATCH is forced by weather data from ECMWF MARS archive. The chemistry schemes used is a modified version proposed by Simpson et al. (1993).

1.1 Model Geometry

The model geometry is taken from the input weather data. The vertical resolution is therefore identical with the ECMWF operation model and thus in hybrid vertical coordinates. The horizontal geometry is defined when retrieving the weather data from the MARS system (currently a lat-long grid with 0.5 degree resolution). There are presently 30 layers used.

The domain covered is 17.5° W to 35° E and 32° N to 71.5° N. The grid is an Arakawa C-grid with staggered wind components.

1.2 Forcing and boundary values

The model is initialised with fixed boundary conditions, for most of the 70 species accounted for, based on climatology. Some of the fixed boundaries are overridden by dynamical boundaries from the global CTM MOZART at intervals of every 3 hour. At present dynamic boundaries are taken for O3, CO, HCHO, NO, NO_2, SO_2, HNO_3, HO_2NO_2, PAN, CH_4, CH_3CHO, C_5H_8 and OXYLENE. The dynamic boundary fields are redistributed in the vertical in a mass-conservative way to fit into the vertical hybrid coordinates used by ECMWF. The model top boundary is taken as the top layer of the redistributed boundaries.

The procedures handling of boundary conditions imply that the model will run with decent boundaries at moments when the dynamic boundaries may fail.

The GEMS emissions in 11 SNAP classes are used. Ship emissions from EMEP are added.

Land-use information are based on 4 classes, sea, low vegetation, forest and no vegetation.

1.3 Dynamical core

Mass conservative transport schemes are used for advection and turbulent transport. The advection is described by Bott-like scheme (Bott [1989], Robertson et al. [2007]). A second order transport scheme is used in the horizontal as well as the vertical. The vertical diffusion is described by an implicit mass conservative first order scheme where the exchange coefficients for neutral and stable conditions are parameterized following Holtslag and

Moeng (1991). In the convective case the turbulent Courant number is directly determined from the turn-over time in the ABL.

Part of the dynamical core is the initialization of wind components. This is a very important step to ensure mass conservative transport. The initialization is based on a procedure proposed by Heiman and Keeling (1989), where the winds are adjusted by means of the difference between the input pressure tendency and the calculated pressure tendency assumed to be an error in the divergent part of the wind field.

1.4 Physical Parameterisations

Boundary layer parameterization is based description of surface heat and vapour fluxes described van Ulden and Holtslag (1985) for land surfaces and Burridge and Gadd (1977) for sea surfaces. The boundary layer height is based on formulations proposed by Zilitinkevich and Moronov (1996) for the neutral and stable case and on Holtslag et al. (1995) for the convective case. These parameterizations drive the formulations for dry deposition and vertical diffusion.

1.5 Chemistry

The photochemistry scheme is based on the EMEP chemistry scheme (Simpson et al., 1993), with some updates where a modified production scheme for isoprene is the most notable based on the so-called Carter-1 mechanism (Carter, 1996; Langner et al., 1998).

2. Assimilation system

3D variational data assimilation scheme (nor yet in operation)
Ensemble Kalman Filter (EnKF) is available in a serial version (not yet operational).

3. Development plan

- Scrutinization of the photo chemical chemistry scheme.
- Increasing model resolution
- Adding emissions from biomass burning

3.1 Development in progress or finilized

- Implementation of routines for radiation transfer for visibility products (finalized).
- Parallelization of the 3D data assimilation scheme (finalized).
- Adding code for handling of observations for data-assimilation at ECMWF (finalized).
- Technical update with a general GRIB1/GRIB2 package (finalized).
- Calculations of background error statistics for 3Dvar (finalized)
- Implementing software for a wider range of observation types in BUFR format (finalized)
- Parallelization of the EnKF data assimilation scheme (in progress).

The next development steps for the MACC version of MATCH are :

- scrutinization of the photo chemical chemistry scheme.
- parallelization of the EnKF data assimilation scheme.

4. References

Carter, W.P.L, Condensed atmospheric photo oxidation mechanism for isoprene. *Atmosph. Env.* **30,** 4275-4290, 1996.

Heiman, M. and Keeling, C.D., A three-dimensional model of CO_2 tramsport based on observed winds. Model description and simulated trace experiment. In *aspects of Climate Variability in the Pacific and Western Americas* (ed. D.H Peterson). American Geophysical Union, Washington, DC, pp. 237-275, 1989.

Holtslag. A.A.M., Meigaard, E. van and De Rooy, W.C., A comparison of boundary layer diffusion schemes in unstable conditions over land. *Boundary Layer Met.*, **76**, 69-95, 1995.

Langner, J., Bergström, R. and Pleijel, K., European scale modeling of sulphur, oxidized nitrogen and photochemical oxidants. Model dependent development av evaluation for the 1994 growing season. SMHI report, RMK No. 82, Swedish Met. And Hydrol. Inst., Norrköping, Sweden, 1998.

Robertson, L. , Langner, J. and Engardt, M., An Eulerian limited-area atmospheric transport model. *J. Appl. Met.* **38**, 190-210, 1999.

Simpson, D. Andersson-Sköld, Y. and Jenkin, M.E., Updating the chemical scheme for the EMEP MSC-W oxidant model: current status. EMEP MSC-W Nore 2/93, 1993.

van Ulden, A.P and Holtslag, A.A.M., Estimation of atmosphericv boundary layer parametrers fort diffusion applications. *J. Climate. Appl. Met.*, **24**, 1196-1207, 1975.

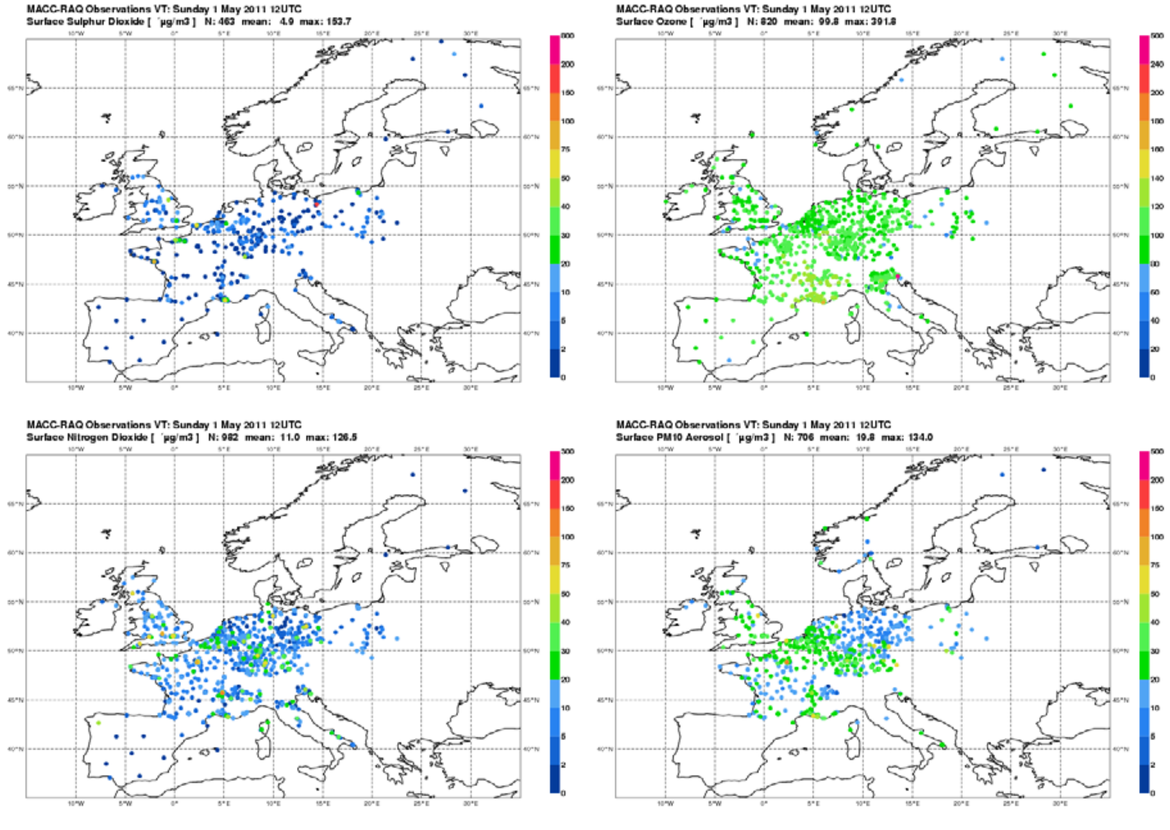
Zilitinkevich, S. and Mornom, D.V., A multi-limit formulation for the equilibrium depth of a stable stratified boundary layer. *Max-Planck-Institute for Meteorology*. Report No. **185**, ISSN 0397-1060, 30 pp., 1996.

Verification report for quarter #4

This verification report covers the period March/April/May 2010. For this report, average skill scores (bias, root mean square error, correlation) for the MATCH model are successively presented for three pollutants : ozone, NO_2 and PM10. The skill is shown for the entire forecast horizon 0 to 72h (3-hourly values), allowing to evaluate the entire diurnal cycle and the evolution of performance from day 1 to day 3.

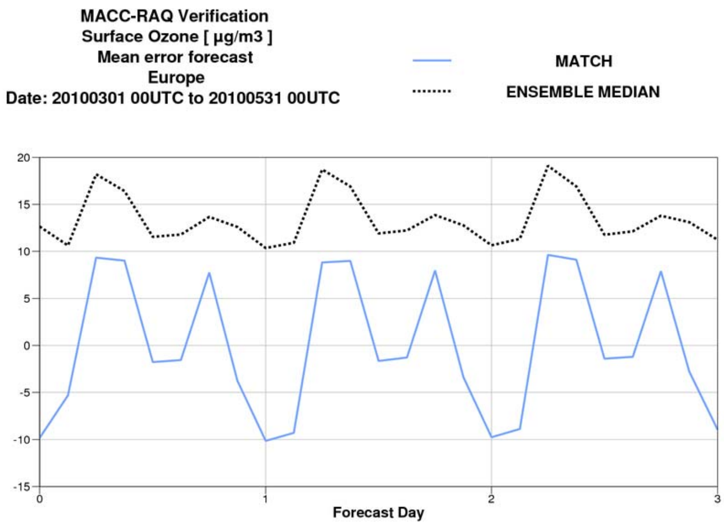
For this verification period, as was the case on the MACC website, verifications are performed against all available data in Near-Real-Time (NRT) for the following countries: Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece (Athens area only), Italy (not all regions), Netherlands, Norway, Poland, Spain, Sweden and the United Kingdom. The total number of sites is typically up to: 900 for ozone, 1200 for NO_2, 550 for SO_2, 300 for CO, 900 for PM10. As an example, the data coverage for Sunday May 1st 2011 12UTC is depicted below. For many countries, all observational data are discarded after 7 days, following the terms in the corresponding Memoranda of Understanding or the other form of agreement. Within MACC, the D-INSITU subproject is working with the European Environment Agency (EEA) to set up a new, more extensive and robust Near-Real-Time dataflow. Once this will be set up, the MACC/R-ENS verification will switch to this new data. According to plans, it is expected that the range of countries will be extended, while work is on-going to ensure that the (large) number of sites with available data in some countries will be kept to a similar level. Data will arrive also sooner, typically less than 3 hours after measurement: currently, hourly data from the day before are available to MACC between 2 and 10 UT approximately everyday, depending on the country.

Site typology is not taken into account in the NRT overall skill scores computations, as there is currently no uniform description available for all regions and countries. Within MACC, work has been carried out [Joly and Peuch, 2011] to build an objective classification of sites, based on the past measurements available in Airbase (EEA). This classification will be used in the last months of MACC, and beyond in MACC-II, to limit verification to the sites that have a sufficient spatial representativeness. Currently, about half of the sites used in the verification are labelled “urban” and are, in principle, not representative enough for comparing against forecast models that have horizontal resolutions of approximately 0.2° over Europe, which thus induces expected biases. However, our first findings indicate that a significant fraction of sites labelled “urban” or “suburban” has actually little local character (as seen in past time series of measurements) and could be in fact more representative than what could be inferred from the metadata information. A remaining issue in the current verification procedure is the large variation in site densities depending upon the different countries: some countries (particularly Germany and France) have much higher site densities and the overall skill scores presented here, and on the MACC website as well, are governed to a large extent by the behaviour of the models in these “data intensive” countries. In the future, our verification procedures will be more segmented (by regions, by countries), which will allow to point out more specificities of model skill in other relevant areas: Northern Europe, Eastern Europe, Mediterranean area…



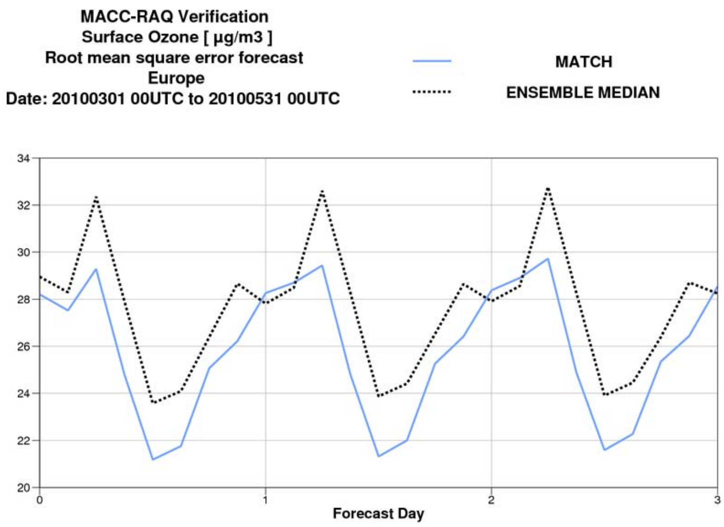
Typical Near-Real-Time data coverage (May 1st 2011, 12 UTC)

MATCH: ozone skill scores, period #4 (March, April, May 2010)

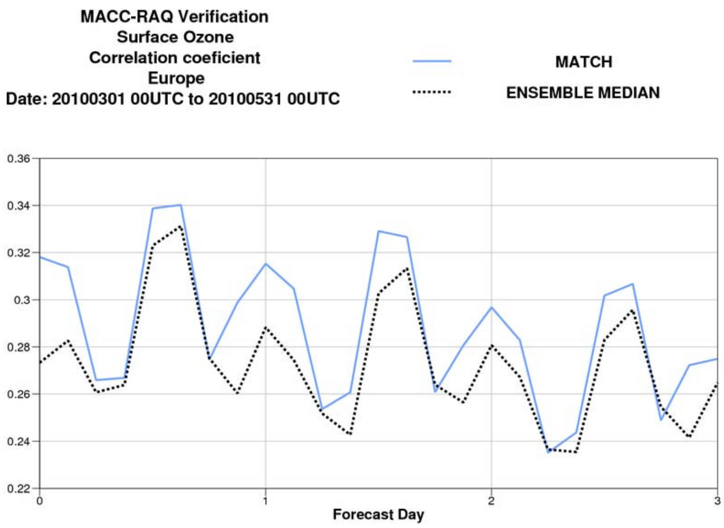


MATCH has a stronger reduction of ozone levels than the ensemple median that is especially pronounced during night time. The deposition velocities are most likely too large during night time.

Test with alternative depositions values are in progress.

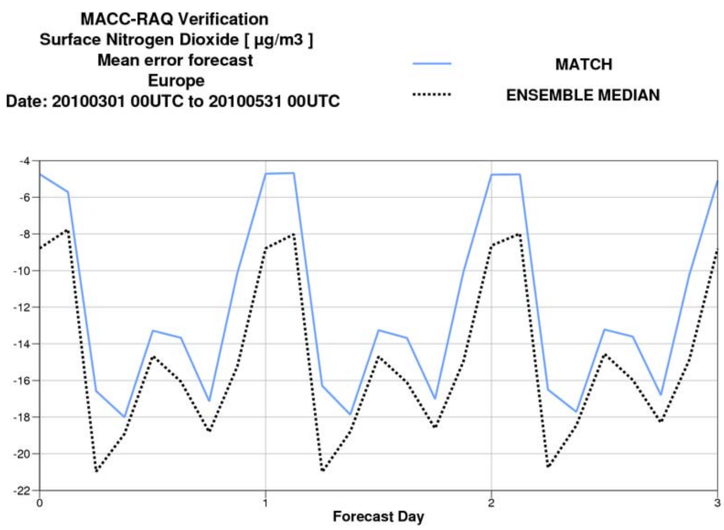


This very much an image of the bias error with the same reason as above.

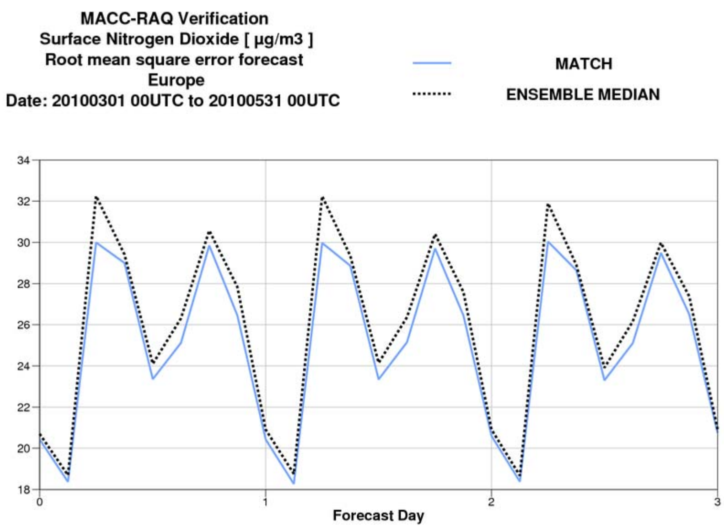


The correlation is just slightly better than the ensemble but both are on a rather low level.

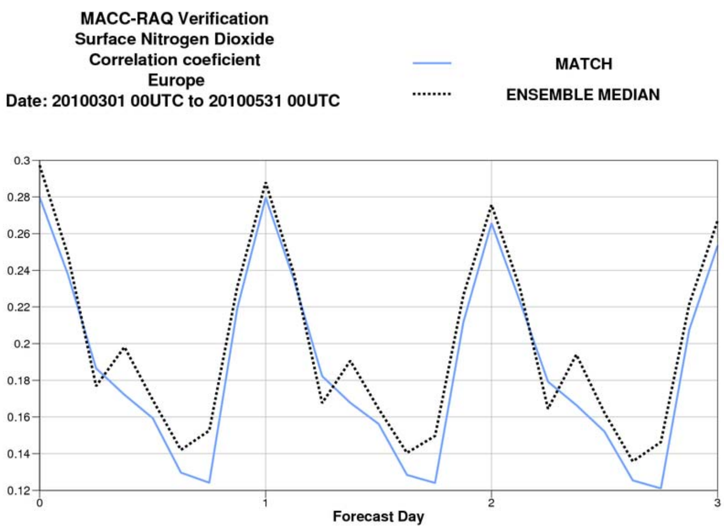
MATCH: NO_2 skill scores, period #4 (March, April, May 2010)



MATCH describes a slightly higher NO_2 bias that should be expected as the ozone levels are lower than the ensemble.

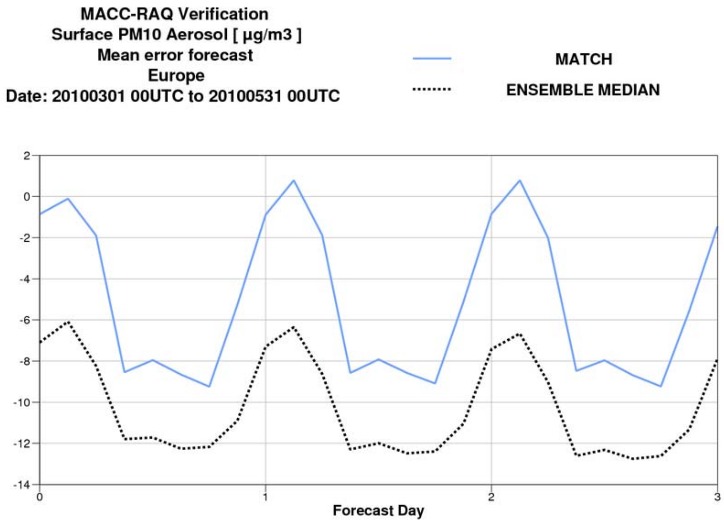


The difference from the ensemble is very small.



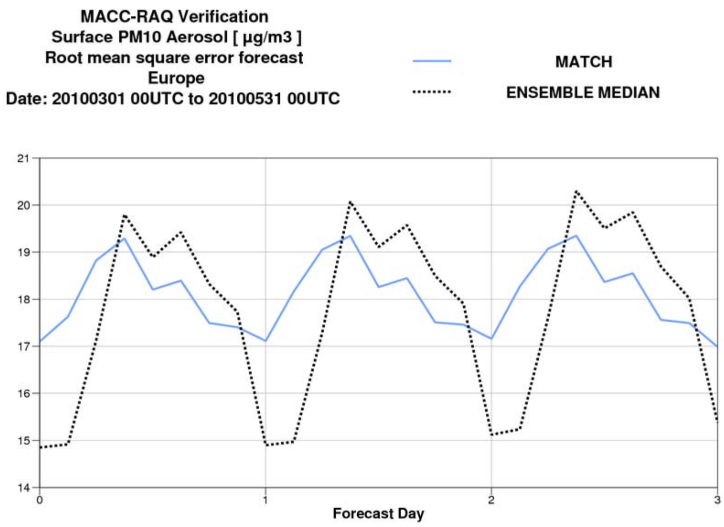
The difference from the ensemble is very small also.

MATCH: PM10 skill scores, period #4 (March, April, May 2010)

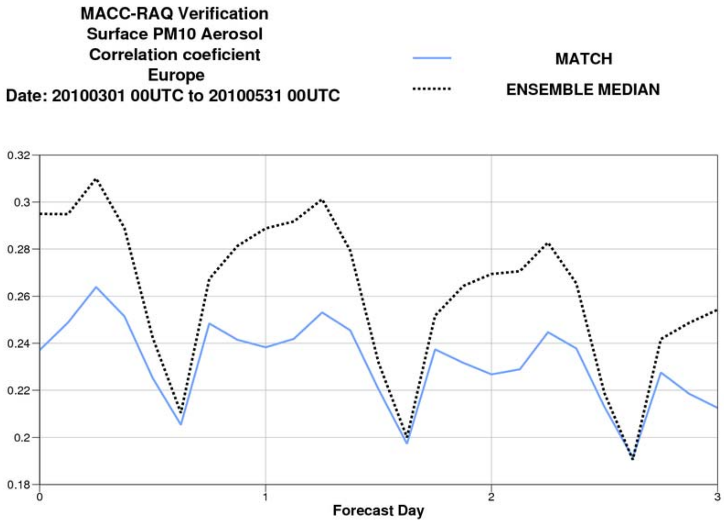


The diurnal variation as almost the same as for the ensemble but with a less negative bias and with larger amplitude.

We have no clear explanation MATCH appears to be better here.



MATCH and the ensemble median display a similar shift in the peak hour, from late night to late morning. RMSE for MATCH is rather constant over the forecast days,



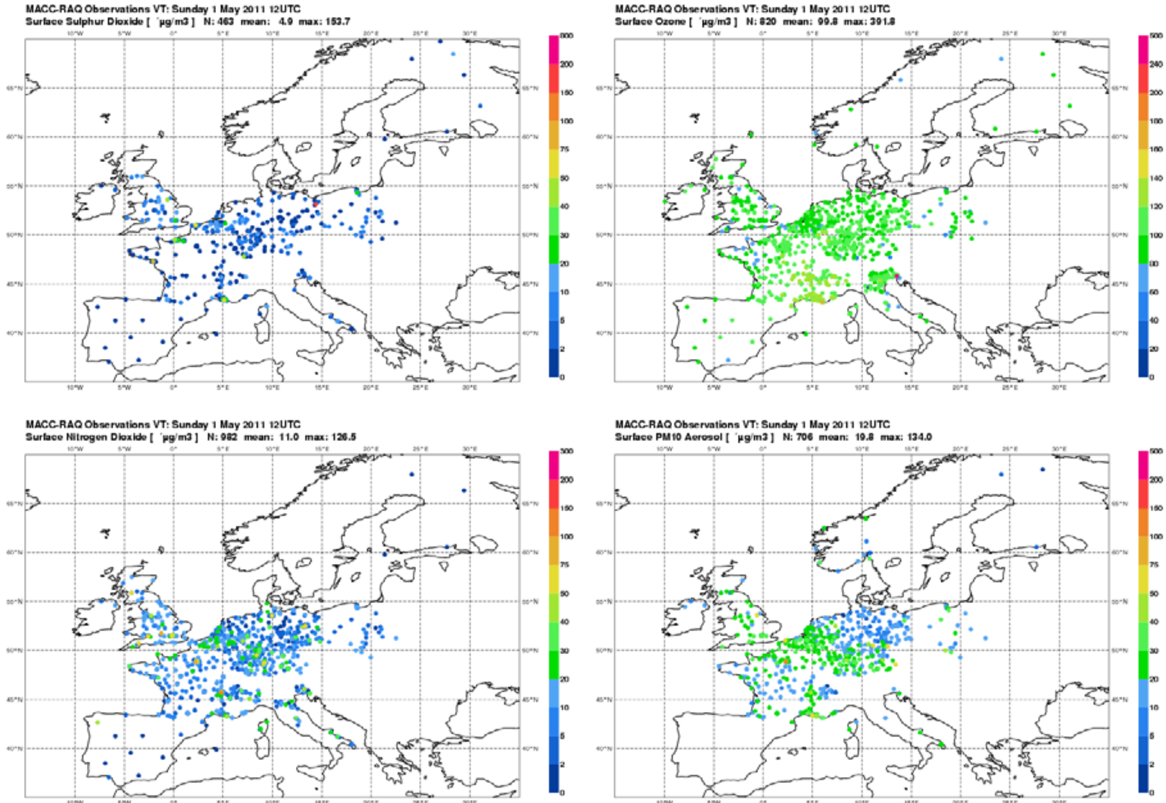
Similar variation pattern and a similar trend of decreasing correction by each forecast day, even though MATCH appears to have a poorer correlation.

Verification report for quarter #5

This verification report covers the period June/July/August 2010. For this report, average skill scores (bias, root mean square error, correlation) for the MATCH model are successively presented for three pollutants : ozone, NO_2 and PM10. The skill is shown for the entire forecast horizon 0 to 72h (3-hourly values), allowing to evaluate the entire diurnal cycle and the evolution of performance from day 1 to day 3.

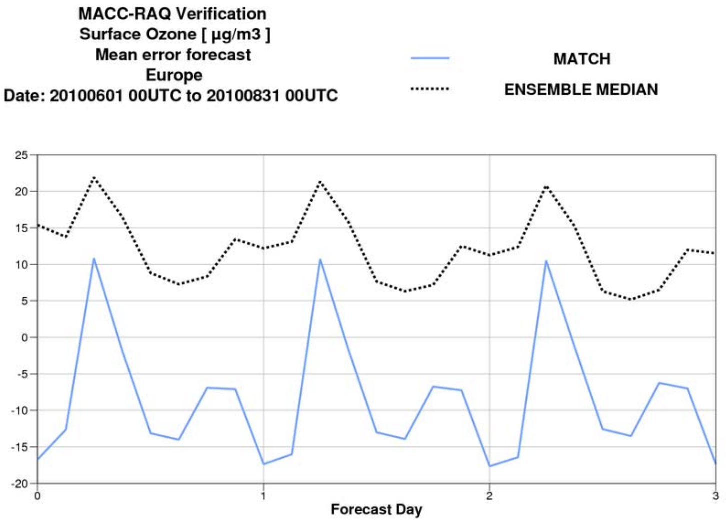
For this verification period, as was the case on the MACC website, verifications are performed against all available data in Near-Real-Time (NRT) for the following countries: Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece (Athens area only), Italy (not all regions), Netherlands, Norway, Poland, Spain, Sweden and the United Kingdom. The total number of sites is typically up to: 900 for ozone, 1200 for NO_2, 550 for SO_2, 300 for CO, 900 for PM10. As an example, the data coverage for Sunday May 1^{st} 2011 12UTC is depicted below. For many countries, all observational data are discarded after 7 days, following the terms in the corresponding Memoranda of Understanding or the other form of agreement. Within MACC, the D-INSITU subproject is working with the European Environment Agency (EEA) to set up a new, more extensive and robust Near-Real-Time dataflow. Once this will be set up, the MACC/R-ENS verification will switch to this new data. According to plans, it is expected that the range of countries will be extended, while work is on-going to ensure that the (large) number of sites with available data in some countries will be kept to a similar level. Data will arrive also sooner, typically less than 3 hours after measurement: currently, hourly data from the day before are available to MACC between 2 and 10 UT approximately everyday, depending on the country.

Site typology is not taken into account in the NRT overall skill scores computations, as there is currently no uniform description available for all regions and countries. Within MACC, work has been carried out [Joly and Peuch, 2011] to build an objective classification of sites, based on the past measurements available in Airbase (EEA). This classification will be used in the last months of MACC, and beyond in MACC-II, to limit verification to the sites that have a sufficient spatial representativeness. Currently, about half of the sites used in the verification are labelled “urban” and are, in principle, not representative enough for comparing against forecast models that have horizontal resolutions of approximately 0.2° over Europe, which thus induces expected biases. However, our first findings indicate that a significant fraction of sites labelled “urban” or “suburban” has actually little local character (as seen in past time series of measurements) and could be in fact more representative than what could be inferred from the metadata information. A remaining issue in the current verification procedure is the large variation in site densities depending upon the different countries: some countries (particularly Germany and France) have much higher site densities and the overall skill scores presented here, and on the MACC website as well, are governed to a large extent by the behaviour of the models in these “data intensive” countries. In the future, our verification procedures will be more segmented (by regions, by countries), which will allow to point out more specificities of model skill in other relevant areas: Northern Europe, Eastern Europe, Mediterranean area…



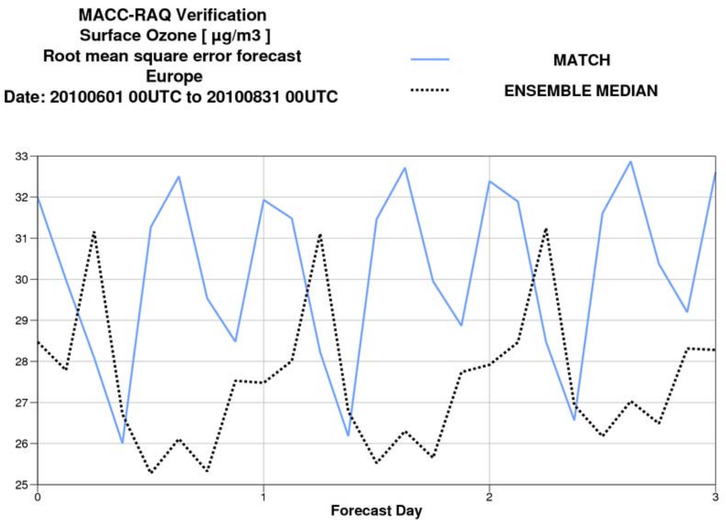
Typical Near-Real-Time data coverage (May 1st 2011, 12 UTC)

MATCH: ozone skill scores, period #5 (June, July, August 2010)

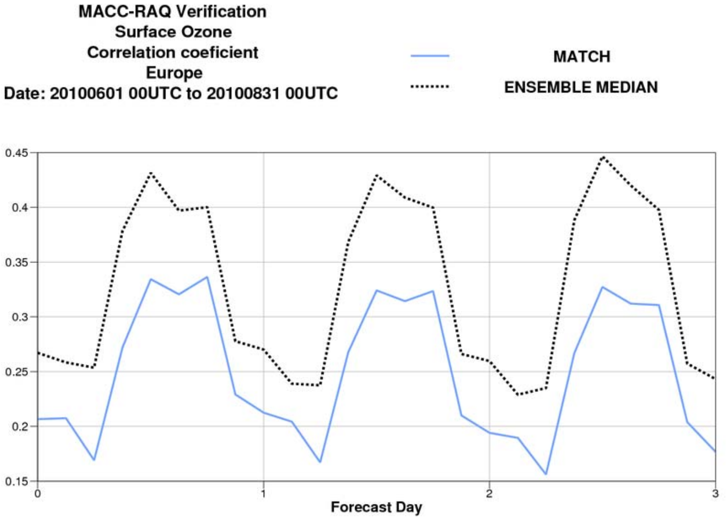


The MATCH bias is on the average about -10 ug/m3 while the corresponding mean for the ensemble is about +10. MATCH appears to underestimate the titration by high NO emissions in the morning hours leading to positive bias.

The resolution of 0.5 degree and a lot of urban measurements may contribute to this bias pattern.

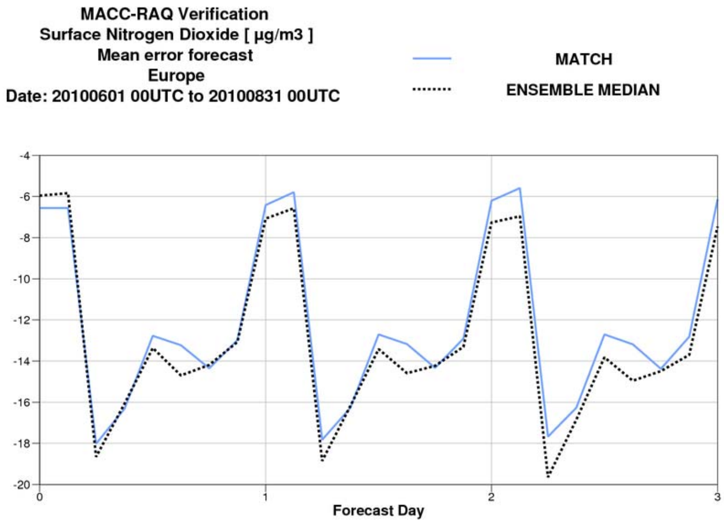


It is in this case really hard to produce any good explanation the difference between MATCH and the ensemble.

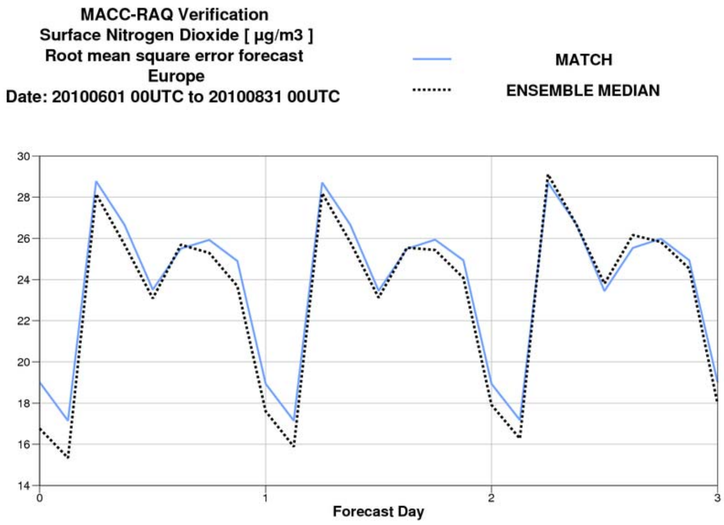


Similar variation of the correlation but with a somewhat lower level for MATCH.

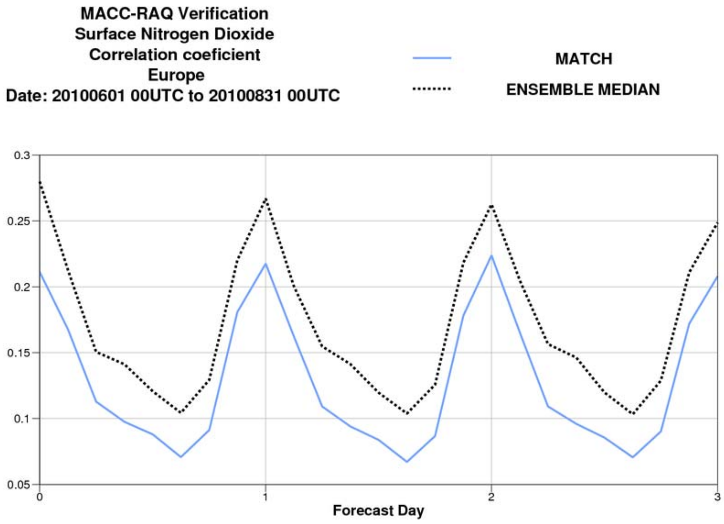
MATCH: NO_2 skill scores, period #5 (June, July, August 2010)



The MATCH model appears to be within the main stream of the ensemble for NO_2.

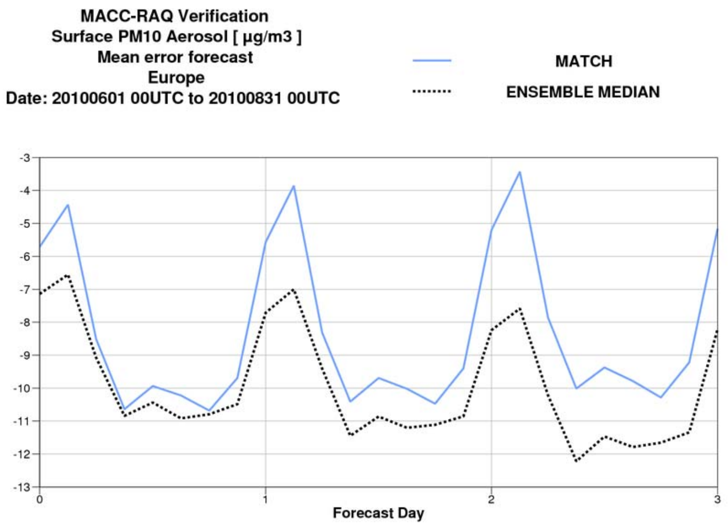


Same as for NO_2 bias.

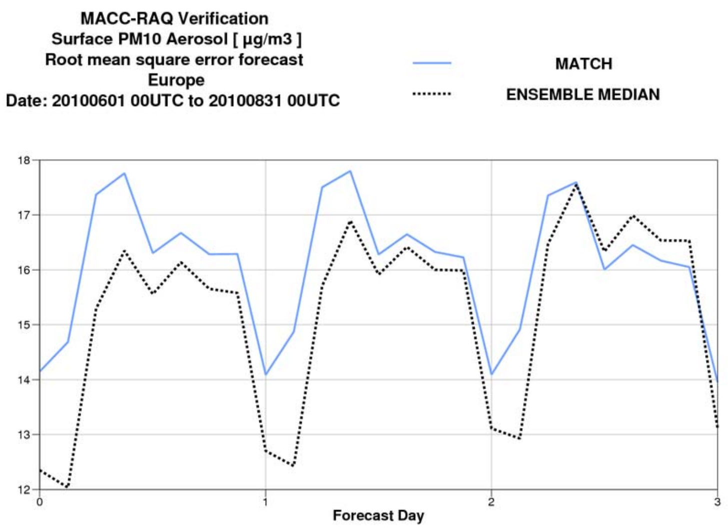


Same variation of the correlation in MATCH and the ensemble but MATCH performs a bit worse.

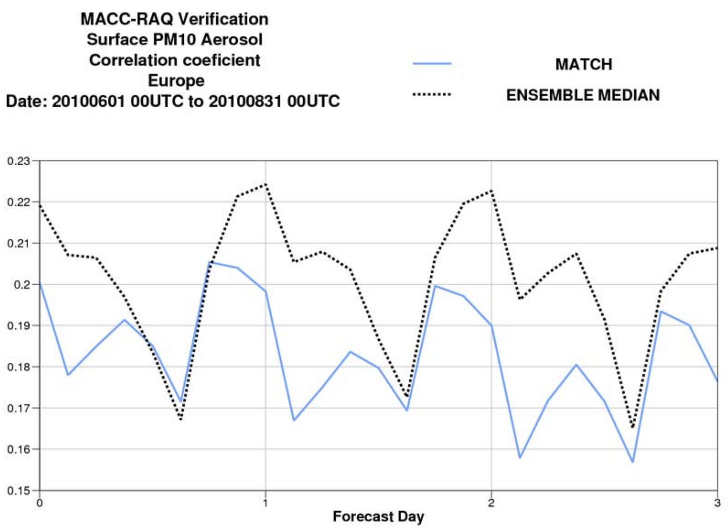
MATCH: PM10 skill scores, period #5 (June, July, August 2010)



This is very similar to period #4. At the moment MATCH do not include fire emissions that would be a part of the explanation but could not explain the diurnal variation.



The MATCH appears to the same RMSE patterns an level over the three forecast days, while the ensemble RMSE is increasing day bay day and is at day three very much the same as MATCH.



The ensemble appears to be a bit better than MATCH but the correlations are not that impressive.