

# 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 Rouil, R-EVA leader (INERIS, #23)

*Contact :*

Lennart.Robertson@smhi.se



# 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 <sub>3</sub> , NO, NO <sub>2</sub> , CO, SO <sub>2</sub> , PM <sub>2.5</sub> , PM <sub>10</sub>	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 accidentally not made.

## 3. Availability statistics

Quarter 6 (September, October, November 2010) : 93.4% of the MATCH forecasts have been provided in time.

MATCH data were missing for 16 September, 1, 19 (incomplete) and 23 October, 11 and 12 November 2010. For the first dates we have no clear explanation of missing or incomplete data. The two last dates are present and delivered the day before (20 UTC). Here most likely the transfer of data fail to MF.

Quarter 7 (December 2010, January, February 2011) : 92.2% of the MATCH forecasts have been provided in time.

MATCH data were missing for 19 December 2010, 1, 5, 7 and 16 January, 1 and 7 February 2011. Of these dates 19 December (2010), 5, 6, 16 January and 7 February (2011) are delivered at the latest 21.30 UTC before the valid date. Data transfer problems to MF may be the reason. The other dates are complete and with late deliver. No specific explanation for these failures.

## 4. Assimilation and forecast system: synthesis of main characteristics

Assimilation and Forecast System	
Horizontal resolution	0.2° (since 10 January 2011)
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 H <sub>2</sub> O <sub>2</sub> and O <sub>3</sub>
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
<b>Forecast System</b>	
Meteorological driver	12:00 UTC operational IFS forecast for the day before (0.5°, 40 levels)
<b>Assimilation System</b>	
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 (0.2°, 40 levels)

# Evolutions in the MATCH suite

## EnKF Pollen

- 2011/02/15** : Added support for general choice of projections by proj4
- 2011/01/29** : Code for pollen module proposed by FMI was added
- 2010/10/30** : The 3Dvar code is operational
- 2010/10/22** : Chi-2 test included in the 3Dvar code
- 2010/09/03** : Variational quality control in 3Dvar included
- 2010/06/15** : Pygrib developed for post processing purposes
- 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 | 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 0.2 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 O<sub>3</sub>, CO, HCHO, NO, NO<sub>2</sub>, SO<sub>2</sub>, HNO<sub>3</sub>, HO<sub>2</sub>NO<sub>2</sub>, PAN, CH<sub>4</sub>, CH<sub>3</sub>CHO, C<sub>5</sub>H<sub>8</sub> and OXYLENE. The model top boundary is defined as the mean of the horizontal boundaries at the model top (due empty global boundaries in the internal of the domain). 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 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.
- Parallelization of the EnKF data assimilation scheme.
- Calculations of background error statistics
- Implementing software for a wider range of observation types in BUFR format
- Increasing model resolution
- Adding emissions from biomass burning

#### **3.1 Development in progress or finalized**

- Implementation of routines for radiation transfer for visibility products (finalized).
- Parallelization of the 3D data assimilation scheme (in progress).
- Adding code for handling of observations for data-assimilation at ECMWF (in progress).
- Technical update with a general GRIB1/GRIB2 package (finalized).

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

- scrutinization of the photo chemical chemistry scheme.
- implementation of routines for radiation transfer for visibility products.
- parallelization of the EnKF data assimilation scheme.
- adding code for handling of observations for data-assimilation at ECMWF.
- technical update with a general GRIB1/GRIB2 package.

#### 4. References

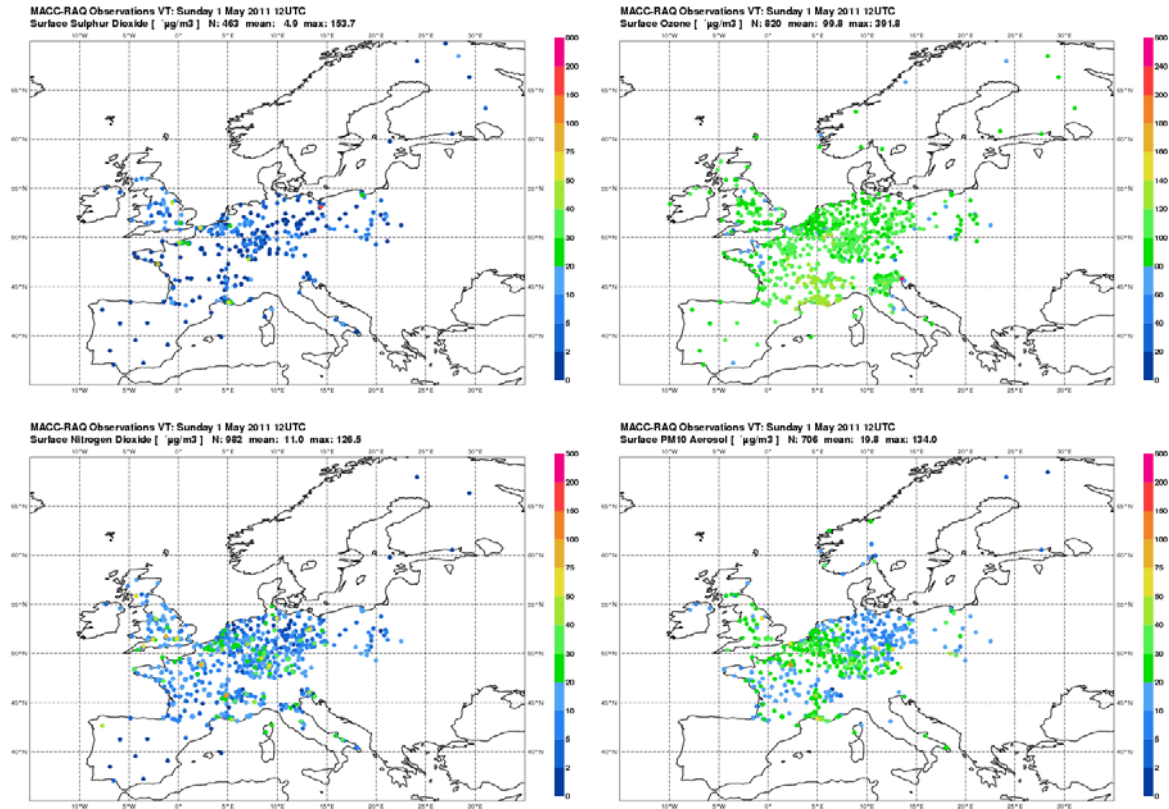
- Carter, W.P.L, Condensed atmospheric photo oxidation mechanism for isoprene. *Atmosph. Env.* **30**, 4275-4290, 1996.
- Heiman, M. and Keeling, C.D., A three-dimensional model of CO<sub>2</sub> transport based on observed winds. Model description and simulated trace experiment. In *aspects of Climate 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 and 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 atmospheric boundary layer parameters for 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 #6

This verification report covers the period September/October/November 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<sub>2</sub> 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.

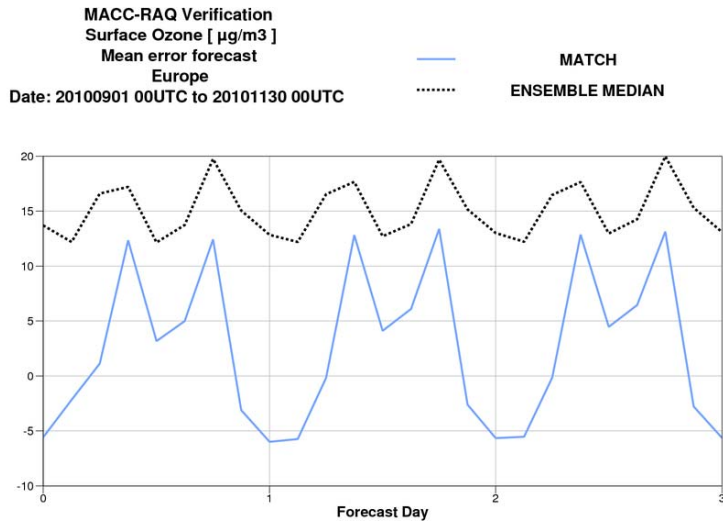
For this verification period, as was the case on the MACC website, verifications are performed against all available data in Near-Real-Time (NRT) for the following countries: Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece (Athens area only), Italy (not all regions), Netherlands, Norway, Poland, Spain, Sweden and the United Kingdom. The total number of sites is typically up to: 900 for ozone, 1200 for NO<sub>2</sub>, 550 for SO<sub>2</sub>, 300 for CO, 900 for PM10. As an example, the data coverage for Sunday May 1<sup>st</sup> 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...

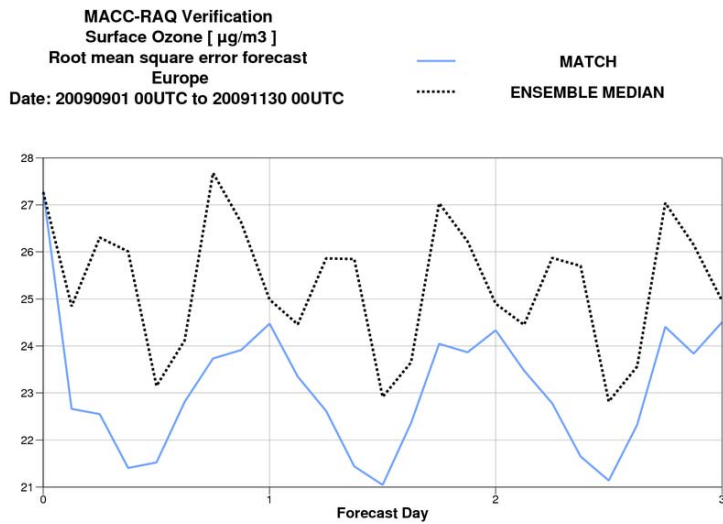


Typical Near-Real-Time data coverage (May 1<sup>st</sup> 2011, 12 UTC)

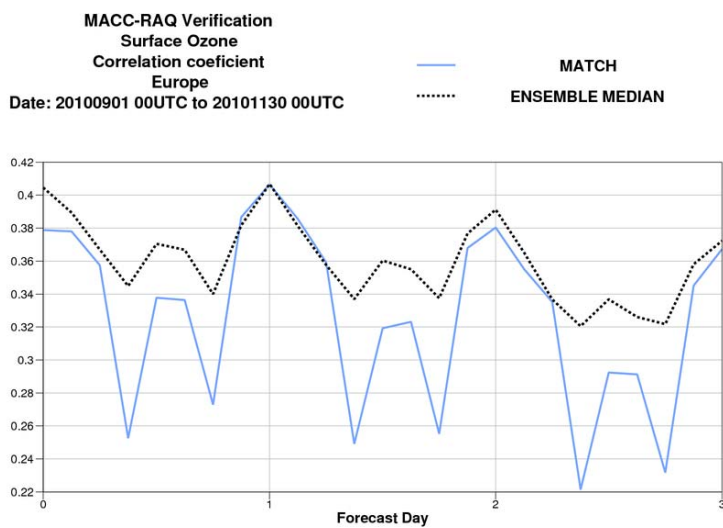
## MATCH: ozone skill scores, period #6 (September, October, November 2010)



This almost a mirror image to summer period #5 but rather similar to the spring period #4. MATCH performs generally better than the ensemble but with an amplified diurnal (but persistent) variation. The dry deposition velocities may be part of the bias seen here.

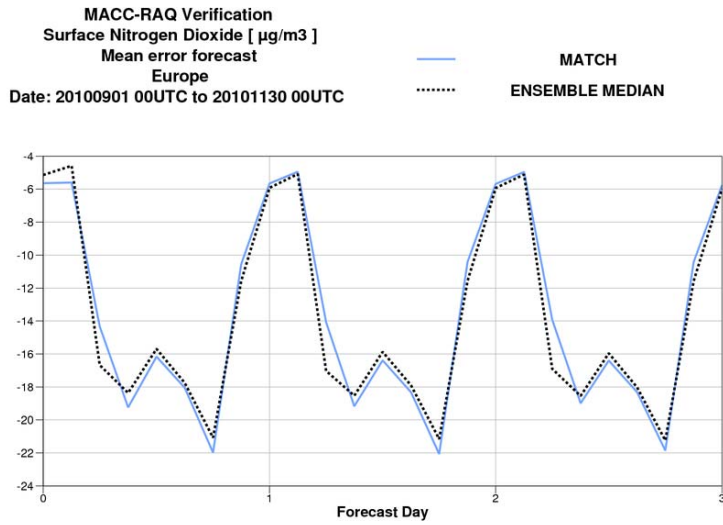


As for the spring period #4 MATCH displays a lower RMSE than the ensemble.

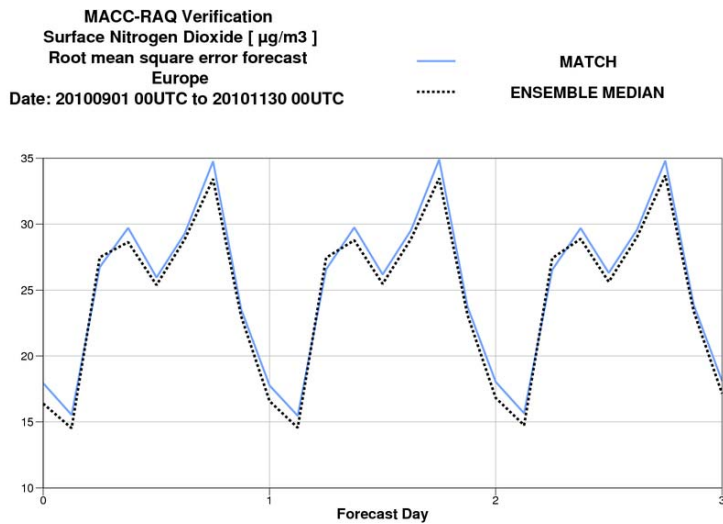


MATCH and the ensemble do have similar correlations at night time. At these points MATCH has a medium negative bias and the largest RMSE.

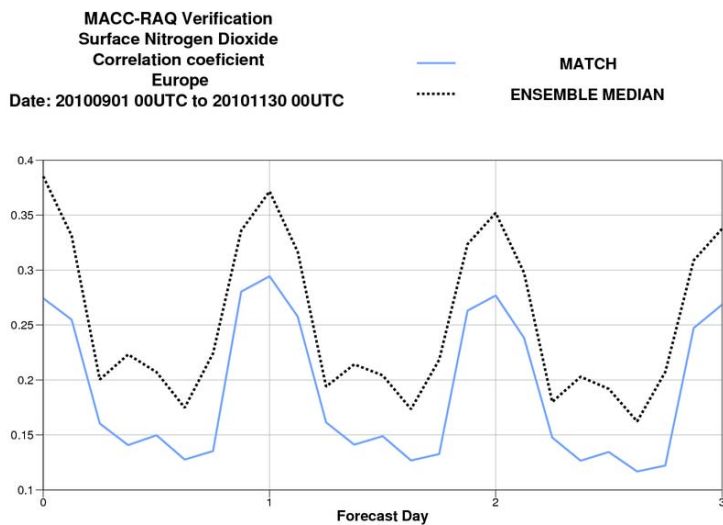
## MATCH: NO<sub>2</sub> skill scores, period #6 (September, October, November 2010)



MATCH and the ensemble appear to be almost identical.

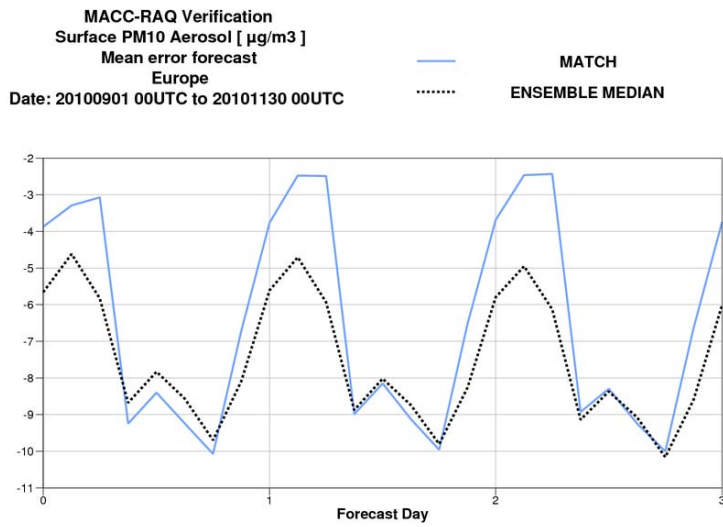


MATCH and the ensemble appear to be almost identical.

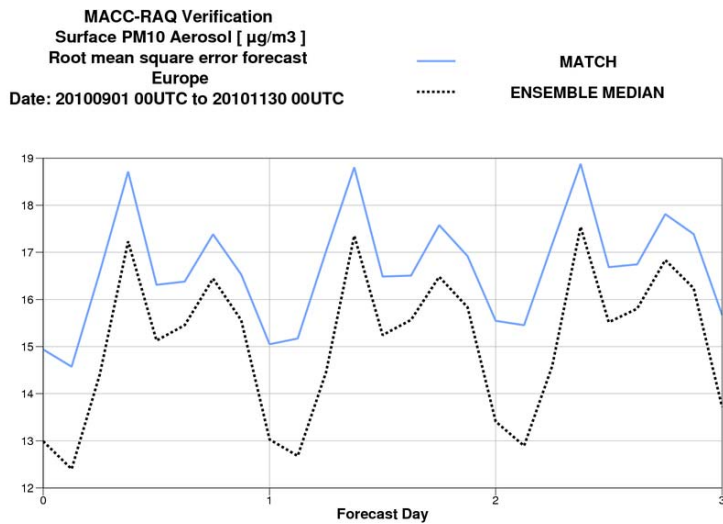


Despite the similarities in bias and RMSE MATCH displays a poorer performance with respect to the correlation. There are no obvious explanations at hand.

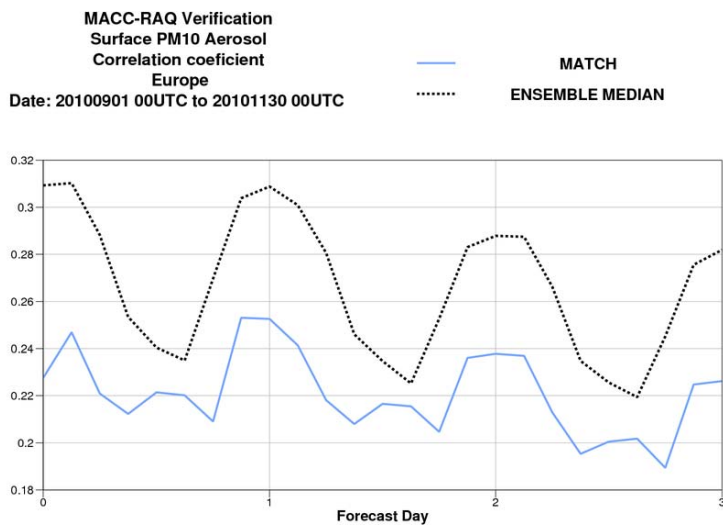
## MATCH: PM10 skill scores, period #6 (September, October, November 2010)



Since period #4 MATCH and the ensemble has narrowed, towards each other, in terms of the bias with a similar pattern where the bias is smallest night time, and with a maximum negative bias of about  $-10 \mu\text{g}/\text{m}^3$ .



There seems to be a systematic shift of about  $2 \mu\text{g}/\text{m}^2$  higher RMSE in MATCH in comparison to the ensemble.



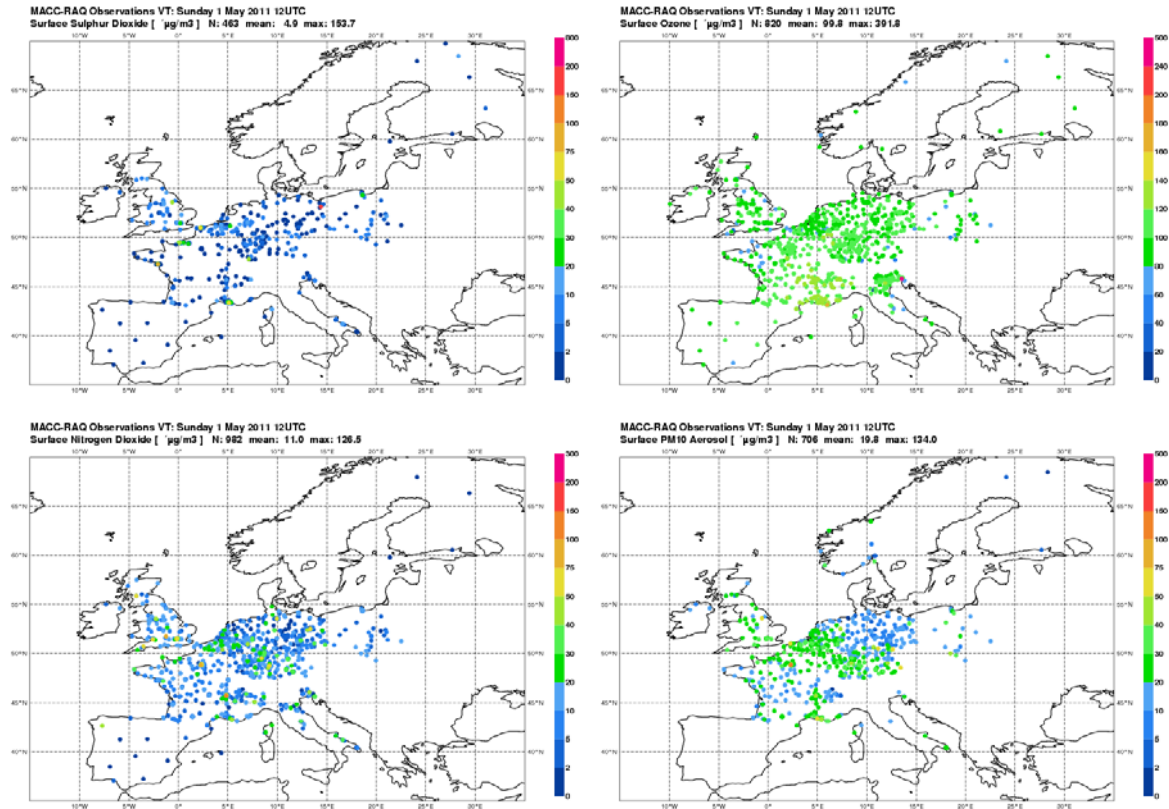
There is a general pattern that MATCH displays a lower correlation for all compounds than the ensemble.

# Verification report for quarter #7

This verification report covers the period December 2010 and January/February 2011. For this report, average skill scores (bias, root mean square error, correlation) for the MATCH model are successively presented for three pollutants : ozone, NO<sub>2</sub> 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.

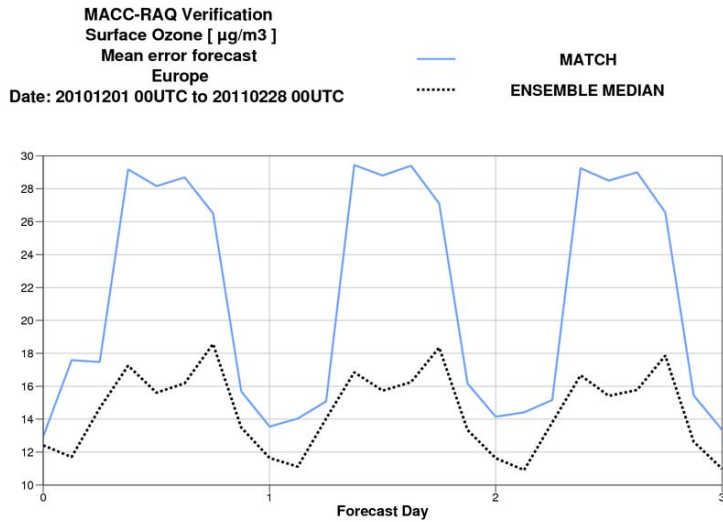
For this verification period, as was the case on the MACC website, verifications are performed against all available data in Near-Real-Time (NRT) for the following countries: Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece (Athens area only), Italy (not all regions), Netherlands, Norway, Poland, Spain, Sweden and the United Kingdom. The total number of sites is typically up to: 900 for ozone, 1200 for NO<sub>2</sub>, 550 for SO<sub>2</sub>, 300 for CO, 900 for PM10. As an example, the data coverage for Sunday May 1<sup>st</sup> 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...



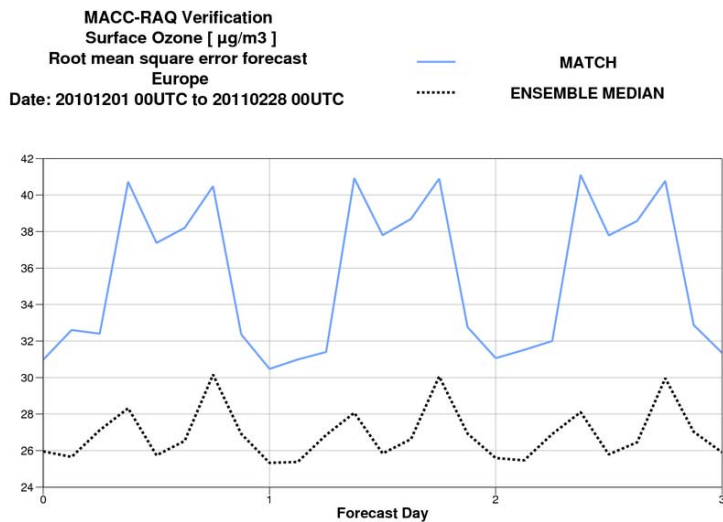
Typical Near-Real-Time data coverage (May 1<sup>st</sup> 2011, 12 UTC)

## MATCH: ozone skill scores, period #7 (December 2010, January, February)

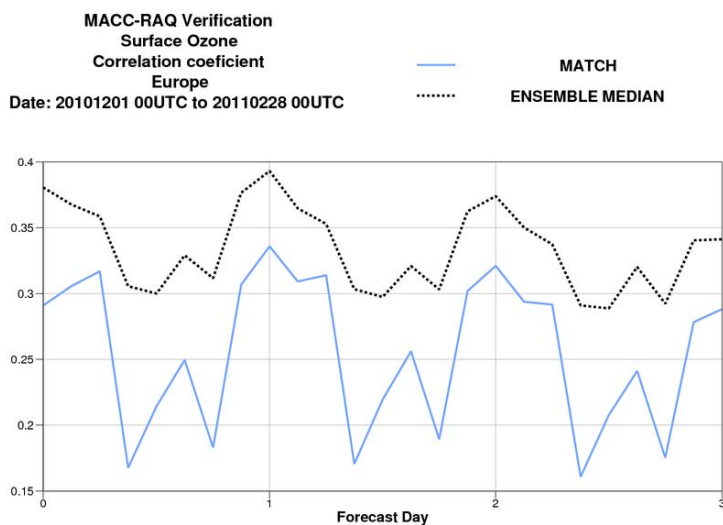


We do suspect the dry deposition over snow fields to be responsible for this large bias of MATCH during daytime in the winter period. It is probable that we overestimate the snow cover impact. We assume that the entire grid-cell is covered by snow when snow cover is present that may not be true in reality.

Since mid January the snow cover was switched off but this action did not appear to have had any impact on the statistics.

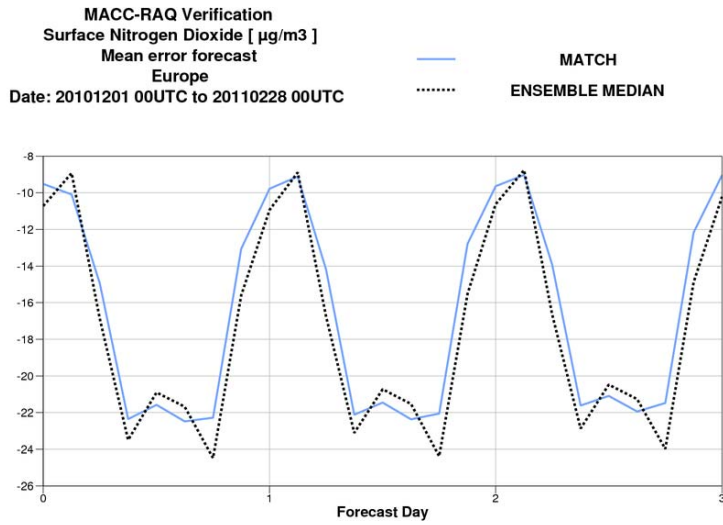


The RMSE reflects the bias error in the positive bias.

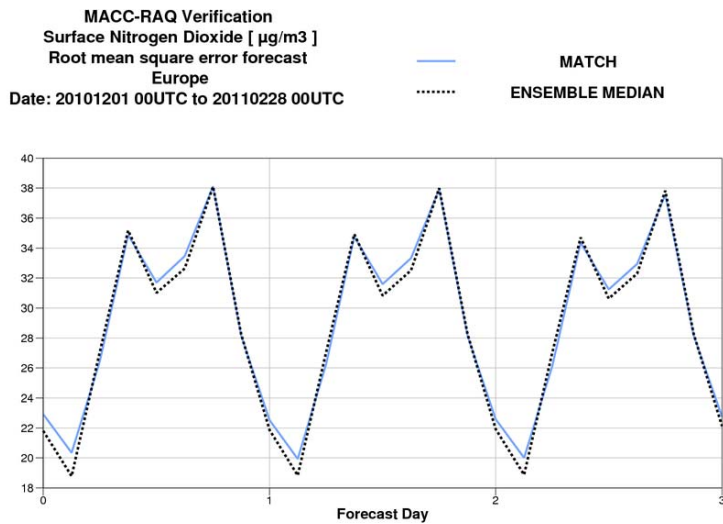


There is a general pattern that MATCH displays a lower correlation for all compounds than the ensemble.

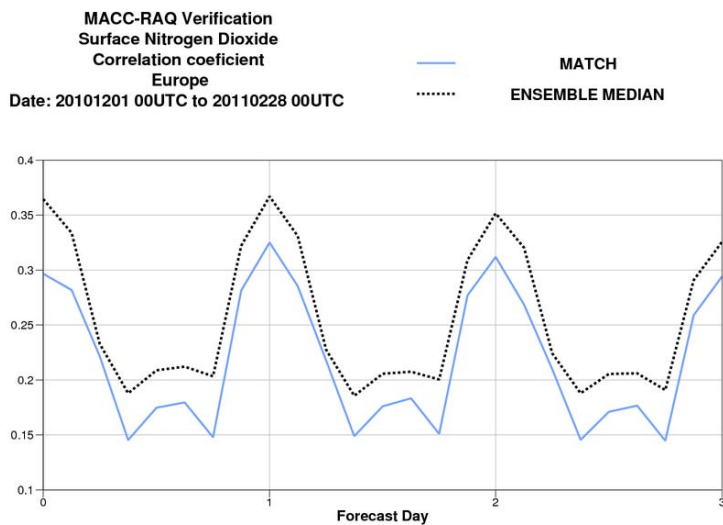
## MATCH: NO<sub>2</sub> skill scores, period #7 (December 2010, January, February 2011)



The MATCH model appears to be within the main stream of the ensemble for NO<sub>2</sub>.

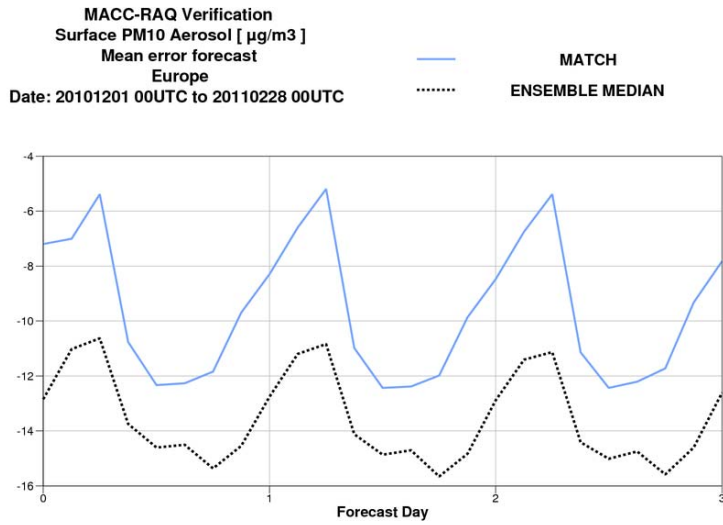


Same as for NO<sub>2</sub> bias.

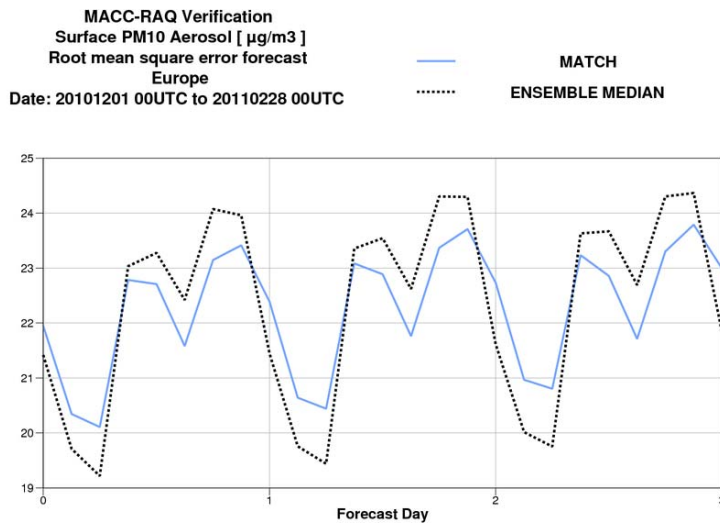


Almost the same as for NO<sub>2</sub> bias.

## MATCH: PM10 skill scores, period #7 (December 2010, January, February 2011)

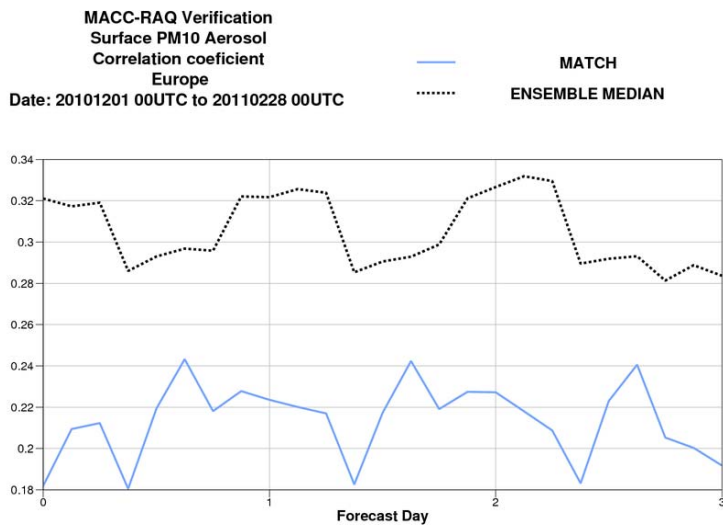


For this period both MATCH and the ensemble have an increased negative bias. The trend is though more pronounced for the ensemble with a shift of  $-5 \mu\text{g}/\text{m}^3$  while MATCH has a shift of  $-2 \mu\text{g}/\text{m}^3$ . There is no obvious explanation for this increased negative bias. Missing modelling of biomass burning would however not be part of this explanation.



MATCH PM10 values are within the same range as the ensemble.

Th



There is a general pattern that MATCH displays a lower correlation for all compounds than the ensemble.