Atmos. Chem. Phys. Discuss., 15, C4439–C4454, 2015
www.atmos-chem-phys-discuss.net/15/C4439/2015/

Atmospheric Chemistry and Physics Discussions
Open Access

ACPD
15, C4439–C4454, 2015

Interactive Comment

Interactive comment on "Evaluation of the MACC operational forecast system – potential and challenges of global near-real-time modelling with respect to reactive gases in the troposphere" *by* A. Wagner et al.

A. Wagner et al.

Annette.Wagner@DWD.de

Received and published: 7 July 2015

Author's answers referee 1: Thank you very much for the comprehensive review of our paper! We have tidied up the numbering of the figures and tables; we have changed the introduction and re-structured the discussion and conclusion and we included the requested changes.

Concerning the use of NO2 surface observations for the validation: We absolutely agree with the referee that surface measurements of NO2 would indeed be very useful in addition to the satellite observations. However, for the global model validation this has not been implemented in MACC so far. For the validation of the global model, it was important for us to compare especially the spatial patterns of NO2 which can hardly be captured globally by the sparse amount GAW station observations. In MACC-III, the validation with MAX-DOAS is tested, however, with regional models with higher spatial resolution.

Concerning the use of the MNMB: We use the MNMB in the MACC and future CAMS evaluations because verifying chemical species concentration values significantly differs from verifying standard meteorological fields. For example, spatial or temporal variations can be much greater and the differences between model and observed values ("model errors") are frequently much larger in magnitude. Most importantly, typical concentrations can vary quite widely between different pollutant types (e.g. O3 and CO) and region (e.g. Europe vs. Antarctica), a given bias or error value can have a quite different significance. It is useful therefore to consider bias and error metrics which are normalised with respect to observed concentrations and hence can provide a consistent scale regardless of pollutant type (see e.g. Elguindi et al., 2010 or Savage et al., 2013). Moreover, the MNMB is robust to outliers, converges to the normal bias for biases approaching zero, while taking into account the representativeness issue when comparing coarse resolved global models versus site specific station observations. Though GAW stations prove regional representative in general, the experience is that local effects cannot always be ruled out reliably in long worldwide data sets, because transport, chemical processes and parameterizations are not selective for the super- to sub-grid-scale threshold. Referencing to the model/observation mean again constitutes a pragmatic workaround to avoid misleading bias tendencies, particularly in sensitive regions with sparse data coverage. Within MACC, the MNMB is used as an important standard score. It is used in the MACC quarterly evaluation reports and it appears in a lot of recent publications, e.g. Cuevas et al. (2015), Eskes et al. (2015), Sheel et al. (2014). As our paper is dedicated to the MACC special issue, we assume that most of the readers will be familiar with this metric and thus we would like to stick to the MNMB in our validations. In our paper, the MNMB is complemented with the commonly used standard metrics RMSE and R.

The specific comments of the referees have been addressed point-by-point in what follows:

Author's response point-by-point: Referee 1: Specific comments: P6279 L1: What about MACC-III? Wouldn't it be more sensible to call it something along the lines of a "series of MACC projects" or similar? -done

P6280-6281: This reads more like a textbook section on atmospheric chemistry than an introduction to a validation paper. Please be concise and focus on what is relevant for this study. It would also be useful here to discuss why we actually care about these gases and why we model them, i.e. what are some potential health effects or other impacts of these gases. See the submitted MACC validation paper by Eskes et al. (2015) in GMDD for an example on this.

-the introduction has been re-written to: The impact of reactive gases on climate, human health and environment has gained increasing public and scientific interest in the last decade (Bell et al., 2006, Cape 2008, Mohnen et al., 2013, Seinfeld and Pandis 2006, Selin et al., 2009). As air pollutants, carbon monoxide (CO), nitrogen oxides (NOx) and ozone (O3) are known to have acute and chronic effects on human health, ranging from minor upper respiratory irritation to chronic respiratory and heart disease, lung cancer, acute respiratory infections in children and chronic bronchitis in adults (Bell et al., 2006, Kampa and Castanas 2006). Tropospheric ozone, even in small concentrations, is also known to cause plant damage in reducing plant primary productivity and crop yields (e.g. Ashmore 2005). It is also contributing to global warming by direct and indirect radiative forcing (Forster et al., 2007, Sitch et al., 2007). Pollution events can be caused by local sources and processes but are also influenced by continental and intercontinental transport of air masses. Global models can provide the transport patterns of air masses and deliver the boundary conditions for regional models, facilitating the forecast and investigation of air pollutants. The EU-funded research project MACC - Monitoring Atmospheric Composition and Climate, (consisting of a series of European projects, MACC to MACC-III), provides the preparatory work that will form the basis of the Copernicus Atmosphere Monitoring Service (CAMS). This service is established by the EU to provide a range of products of societal and environmental value with the aim to help European governments respond to climate change and air quality problems. MACC provides reanalysis, monitoring products of atmospheric key constituents (e.g. Inness et al., 2013), as well as operational daily forecasting of greenhouse gases, aerosols and reactive gases (Benedetti et al., 2011, Stein et al., 2012) on a global and on European-scale level, and derived products such as solar radiation. An important aim of the MACC system is to describe the occurrence, magnitude and transport pathways of disruptive events, e.g., volcanoes (Flemming and Inness, 2013), major fires (Huijnen et al., 2012, Kaiser et al., 2012) and dust storms (Cuevas et al., 2015). The product catalogue can be found on the MACC website, http://copernicus-atmosphere.eu. For the generation of atmospheric products, state-of-the-art atmospheric modelling is combined with assimilated satellite data (Hollingsworth et al., 2008, Inness et al., 2013, 2015, more general information about data assimilation can be found in e.g. Ballabrera-Poy et al., 2009 or Kalnay 2003). Within the MACC project there is a dedicated validation activity to provide up-to-date information on the quality of the reanalysis, daily analyses and forecasts. Validation reports are updated regularly and are available on the MACC websites. The MACC global near-real-time (NRT) production model for reactive gases and aerosol has operated with data assimilation from September 2009 onwards, providing boundary conditions for the MACC regional air quality products (RAQ), and other downstream users. The model simulations also provide input for the stratospheric ozone analyses delivered in near-real-time by the MACC stratospheric ozone system (Lefever et al., 2014). In this paper we describe the investigation of the potential and challenges of near-real-time modelling with the MACC analysis system between 2009 and 2012. We concentrate on this period because of the availability of validated independent obser-

ACPD
15, C4439–C4454, 2015

Interactive Comment

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

vations (namely surface observations from the Global Atmosphere Watch Programme GAW, the European Monitoring and Evaluation Programme EMEP, as well as total column satellite data from the MOPITT, SCIAMACHY and GOME-2 sensors) that are used for comparison. In particular, we study the model's ability to reproduce the seasonality and absolute values of CO and NO2 in the troposphere as well as O3 and CO at the surface. The impact of changes in model version, data assimilation and emission inventories on the model performance is examined and discussed. The paper is structured in the following way: Section 2 contains a description of the model and the validation data sets as well as the applied validation metrics. Section 3 presents the validation results for CO, NO2 and O3. Section 4 provides the discussion and section 5 the conclusions of the paper.

P6281-6282 etc: Sometimes you talk about MACC/MACC-II, sometimes about MACCII and sometimes about MACC. Please be consistent. I recommend introducing the series of MACC projects (including MACC-III) once in the beginning and then referring to it simply as MACC in the remainder of the manuscript. Again, take a look at the submitted MACC validation paper by Eskes et al. (2015) in GMDD for finding out how to do this in a better way. -done

P6281 L19-20: This is worded a bit strangely. It is not the series of MACC projects that form the basis of CAMS, but rather the work that has been carried as part of MACC represents the preparatory activities that in the end are supposed to result in the operational CAMS. -done

P6281 L26: Are there more recent references on how data assimilation is being carried out within MACC/CAMS? If yes, cite them here. Maybe Inness et al. 2013 or similar? -done

P6282 L17-21: It is not clear how the availability of independent observations limits the period of this study to 2009-2012. For sure all the satellite datasets (MOPITT, SCIAMACHY, GOME-2) were available many years before 2009 and with exception of

ACPD

15, C4439–C4454, 2015

Interactive Comment

SCIAMACHY also continued after 2012. Surely GAW and EMEP data were available outside this period as well? Be precise about what is the limiting factor here. -The data availability is only limiting the end of the validation period. We chose 2009 as the beginning because the MACC_osuite model run with data assimilation was introduced in 09/2009.

P6282 L25: are -> is -done

P6282 L28: "encloses"? Better write something like "provides" or "contains" -done

P6283 L19: "MACC_osuite". Can you provide an explanation for this rather odd technical acronym? -done P6283 L24: Be specific about the spatial resolution of the model. Is it 100 km x 100 km or irregular (and/or give it in degrees lat/lon)? -done P6284 L9: What do you mean by "go back"? Do you mean the emissions are taken from or based on the RETRO-REAS inventory? Also, how exactly were the emissions merged? -done P6284: Give more information about the spatial resolution of the various emission Inventories -done P6284 L26: "lists up" -> "lists" -done P6285 L20: Has this been studied (if yes, provide results) or is this just an assumption? -This is the experience (unpublished, however) of our validation work within MACC. P6286 L5: WMO 2010 is not included in the list of references -done P6286 L6: Why specify "tropospheric" here? These are surface observations, right? -done P6286 L24: Why didn't you use vertical interpolation between the two closest model levels. Discuss why the resulting error is negligible (or why not).

The selection of the vertical model level is indeed a challenge. Within MACC, we initially did extensive sensitivity tests for level selection, but had to conclude that there is no clear optimal approach for all stations, terrains and species. At the lowest levels, the narrow spacing of the model levels is often not significantly resolved by model processes and parameterizations, at large model/real surface differences the missing surface influence (e.g. deposition) could introduce more problematic inconsistencies (e.g. in diurnal cycle) than a precisely chosen model altitude, be it w.r.t. altitude,

ACPD
15, C4439–C4454, 2015

pressure or temperature (which all has been applied in published studies but each has clear pros and cons).

P6289 L1: The labels "Fires-Alaska" and "Fires-Siberia" look awkward compared to the other regions. Clarify why these specifically refer to fires and that they are only used for CO validation with MOPITT. Also in some of the Figures these labels are not used consistently. Please fix. -done

P6289 L11: "UV-VIS". Also I would recommend either writing "UV-VIS and NIR" or "ultraviolet-visible and near-infrared" and not mixed. -done

P6289: This section requires a discussion about the expected uncertainty of the satellite-based NO2 retrievals. Also, what is a reasonable minimum threshold of detection for the tropospheric NO2 column derived from SCIAMACHY and GOME-2? We agree that a short section on uncertainties is needed and have added the following paragraph: Satellite observations of tropospheric NO2 columns have relatively large uncertainties, mainly linked to incomplete stratospheric correction (important over clean regions and at high latitudes in winter and spring) and to uncertainties in air mass factors (mainly over polluted regions) (e.g. Boersma et al., 2004 and Richter et al., 2005). The uncertainty varies with geolocation and time but in first approximation can be separated into an absolute error of 5x1014 molec cm-2 and a relative error of about 30%, whichever is larger. As some of the contributions to this uncertainty are systematic, averaging over longer time periods does not reduce the errors as much as one would expect for random errors. Over polluted regions, the uncertainty from random noise in the spectra is small in comparison to other error sources, in particular for monthly averages.

The question of a detection limit for satellite NO2 observations is an interesting one. Averaging of large amounts of data will lower the random noise in the data significantly as has been demonstrated for many trace gases in studies looking at multi-annual averages. While the number of available measurements is limited in real world obser-

ACPD
15, C4439–C4454, 2015

vations, it is not clear to us whether or not a detection limit in the sense of an absolute threshold exists for this type of absorption spectroscopy measurements. We therefore preferred not to give a "detection limit" but rather an absolute uncertainty which for practical applications has the same meaning.

P6290 L7: "linearly in time" -done

P6291 L8: Why does the MNMB used here range from -2 to 2 rather than -1 to 1? Why is this metric multiplied by 2? When using this metric in percent, as the authors do in this study you get a bounded range of -200% to 200%. How should this be interpreted? Please provide additional detail about the statistical properties of this non-standard evaluation metric. The MNMB is a normalization based on the mean of the observed and forecast value. It is used as a standard score within MACC and been adopted in order to avoid asymmetry, which occurs in bias assessment when the mean observation is used as a reference, see also Elguindi et al. (2010). A detailed discussion on this has been added in the general part of the author's answers above.

P6291 L20: Keep the section headers consistent. Either spell out the species or not, but do not mix. -done

P6291 L23: It shows not one but two maps -done P6291 L23: Figure 11? Figures 2-10 have not even been discussed yet. This also applies throughout the rest of the paper. Renumber Figures and Tables based on when they are introduced in the manuscript -done

P6292 L4: "far north" -> Better write "high latitudes in the northern hemisphere" or something similar to be specific -done

P6293 L6: better write "norther hemisphere winter months" -done

P6293 L24: This is not clear from Figure 14. It seems to show negative values of around -30% for Dec 2010? -done, mistake, is supposed to mean Dec 2012

P6293 L25-27: Can you provide an explanation for why Dec 2012 behaves so differ-

ACPD
15, C4439–C4454, 2015

ently? -We suppose because of the limited data availability towards the end of the validation period.

P6293 L27: "diurnal O3 cycle". This is misleading - Figure 15 does not really analyse the diurnal cycle but rather simply differentiates the result by day and night. Consider rewording this. -done

P6294 L21-23: Why do you need to refer to RMSEs and correlation coefficients in this sentence, when you are just talking about MNMBs? Please revise. -done

P6294 L24 "northern hemisphere" -done

P6295 L1: These correlation coefficients are indeed extremely low. A Pearson correlation coefficient of what is on average about 0.3 (Fig 2) translates to an RËĘ2 of 0.09! And this is even for monthly averages and not hourly/daily observations - so the random error should already be reduced to a large extent. If a model can explain less than 10% of the variability in monthly averages, I think quite a bit of explanation about possible reasons for the poor performance is necessary. Please add a discussion on this here. -the low values of 0.1 during the period January 2011 to October 2011 result from the reading error in the fire emissions. The generally only moderate correlation coefficient is related to mismatches in the strong short-term variability seen in both the model and the measurements. Data assimilation also presumably impacts this.

P6295 L3: How was the subset of stations in Figure 3 selected? Were only those stations selected at which the model performed well, or was some other selection process used? Please add information about this in the text. - done in the text, we chose representative examples for every region to underline the results in the text but certainly not only stations were the model performed well, as can be seen by the deviations between model and observations in the plots.

P6295 L25 to P6296 L10: This section discusses solely differences between MOPITT and IASI but not the relevance of these differences with respect to the model. Please

ACPD
15, C4439–C4454, 2015

Interactive Comment

revise to better indicate how these differences affect the model performances? Is it due to assimilation of IASI CO products in the model? -done in the text

P6296 L24: Be careful about interpreting too much into satellite-based NO2 columns over the open oceans. The NO2 levels there tend to be below the detection limits of the instruments and the patterns observed there often represent no true geophysical signal. We agree that NO2 columns, which are close to the uncertainty, should be interpreted with care. We do not think that this is limited to columns over the open oceans although there might be small problems with the spectral signature of vibrational Raman scattering and possibly also liquid water absorption. The patterns seen in the satellite data, which are referred to in the text are clearly linked to outflow from the continents and we do not think they are artefacts. The enhanced values in the Southern Ocean close to Antarctica are a well-known artefact from incomplete removal of stratospheric variability. In response to the comment by the reviewer we have changed the corresponding paragraph as follows: In the northern hemisphere, background values of NO2 VCD over the ocean are lower in the simulations than in the satellite data. The same is true for the South Atlantic Ocean to the west of Africa (see Fig.15). This might suggest a model underestimation of NO2 export from continental sources or too rapid conversion of NO2 into its reservoirs. However, as the NO2 columns over the oceans are close to the uncertainties in the satellite data, care needs to be taken when interpreting these differences.

P6298: Please clearly distinguish here between CO and NO2 here. These are intermixed in the discussion making it difficult to follow. -done

P6299: This section also requires a brief discussion of the potential uncertainties introduced by transitioning from SCIAMACHY to GOME-2 in 2012 and how it affects the validation of NO2. -we have repeated the validation with GOME-2 data and the conclusions are the same as for the combined SCIAMACHY and GOME-2 validation. For the region plots, the daily mean values of SCIAMACHY show a stronger temporal variability compared to the daily mean values of GOME-2. This is due to the differences in the

ACPD

15, C4439–C4454, 2015

data sampling (GOME-2 has a better spatial coverage). Figures are attached below.

P6300 L5: Again, you are not really studying/validating the diurnal cycle. Please reword. -done

P6316: The combined label/region field is a bit confusing. Do only the GAW stations have a label whereas the EMEP stations have a region acronym? For clarity please highlight this in the caption and list the region acronyms. -done

P6319: This table has unrealistically high number of significant digits. Please modify. -done

P6322: The panels in this plot are missing labels as a) b) c), yet the caption refers to them. Also, why does the caption only refer to a) and b) instead of all three. Please be consistent. Also the panels are very small, such that the legend is not readable. -done

P6324: The legend here does not list the region names as "Fires-Alaska" and "Fires-Siberia", as they were introduced previously. Please decide on a label for these regions and then stick to it consistently in text and Figures. -done Discussion Paper P6325: Same in this Figure. -done

P6326: It would be helpful to use different symbols/colours for SCIAMACHY and GOME-2 in this Figure. For the daily values, different symbols are not an option due to the dense plotting in the long time series. We could use different colours; however, this would not be in accordance to the other Figures, then. We could, if this helps, introduce a vertical dashed line in the x-axis.

P6328: Same in this Figure. -see above

P6329: The caption says "daily" but the Figure shows monthly averages. Please correct. -done

P6331: This Figure has an unclear colour scale, making the interpretation of MNMBs close to zero challenging. Plots with divergent colour scale such as this should ideally

ACPD

15, C4439–C4454, 2015

have only one colour gradient for positive and negative values, respectively, with a neutral colour (white or grey) in between. I recommend shades of red for positive values and shades of blue for negative values with white or grey symmetrically around zero. -This is the colour bar used in the MACC validation exercises. We would like to keep it, for in this case we wanted to especially highlight large biases.

P6332: These plots are extremely busy and the legend is unreadable. Please consider ways of reducing the overplotting to increase the visual impact of the Figure. Also, once again, please consistently format and label the panels. Why does subplot a) consist of two panels and subplot b) of one panel. Why not have 3 separate subplots? -done

P6333: Describe either in the caption or in the text how this seemingly random subset of stations was selected. -done, we wanted to give some examples for every region to underline the findings in the text.

Boersma, K.F., Eskes, H.J., Brinksma, E.J.: Error analysis for tropospheric NO2 retrieval from space. J. Geophys. Res., 109, D4, doi:10.1029/2003JD003962, 2004.
Cuevas, E., Camino, C., Benedetti, A., Basart, S., Terradellas, E., Baldasano, J.M., Morcrette, J.-J., Marticorena, B., Goloub, P., Mortier, A., Berjón, A., Hernández, Y., Gil-Ojeda, M., Schulz, M.: The MACC-II 2007-2008 Reanalysis: Atmospheric Dust Evaluation and Characterization over Northern Africa and Middle East, Atmos. Chem. Phys. 15, 3991–4024, doi:10.5194/acp-15-3991-2015, 2015.

Elguindi, N., Clark, H., Ordóñez, C., Thouret, V., Flemming, J., Stein, O., Huijnen, V., Moinat, P., Inness, A., Peuch, V.-H., Stohl, A., Turquety, S., Athier, G., Cammas, J.-P., and Schultz, M.: Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC, Geosci. Model Dev., 3, 501-518, doi:10.5194/gmd-3-501-2010, 2010.

Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrillat, S., Cuevas, E., Engelen, R., Flentje, H., Gaudel, A., Griesfeller, J., Jones, L., Kapsomenakis, J.,Katragkou, E., Kinne, S., Langerock,

ACPD
15, C4439–C4454, 2015

B., Razinger, M., Richter, A., Schultz, M., Schulz, M., Sudarchikova, N., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Validation of reactive gases and aerosols in the MACC global analysis and forecast system. Geosci. Model Dev. Discuss., 8, 1117-1169, doi:10.5194/gmdd-8-1117-2015, 2015.

Richter, A., Burrows, J. P., Nüß, H., Granier, C, Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437-132,doi:10.1038/nature04092, 2005.

Savage, N. H., Agnew, P., Davis, L. S., Ordonez, C., Thorpe, R., Johnson, C. E., O'Connor, F. M., and Dalvi, M.: Air quality modelling using the Met Office Unified Model (AQUM OS24-26): model description and initial evaluation, Geosci. Model Dev., 6, 353–372, 2013, doi:10.5194/gmd-6-353-2013, 2013. Sheel, V., Sahu, L.K., Kajinu, M., Deushi, M., Stein, O., Nedelec, P.: Seasonal and interannual variability of carbon monoxide based on MOZAIC observations, MACC reanalysis, and model simulations over an urban site in India. J. Geophys. Res., 119, 14, 9123–9141, 2014.

Please also note the supplement to this comment:
http://www.atmos-chem-phys-discuss.net/15/C4439/2015/acpd-15-C4439-2015-supplement.pdf

---

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 6277, 2015.

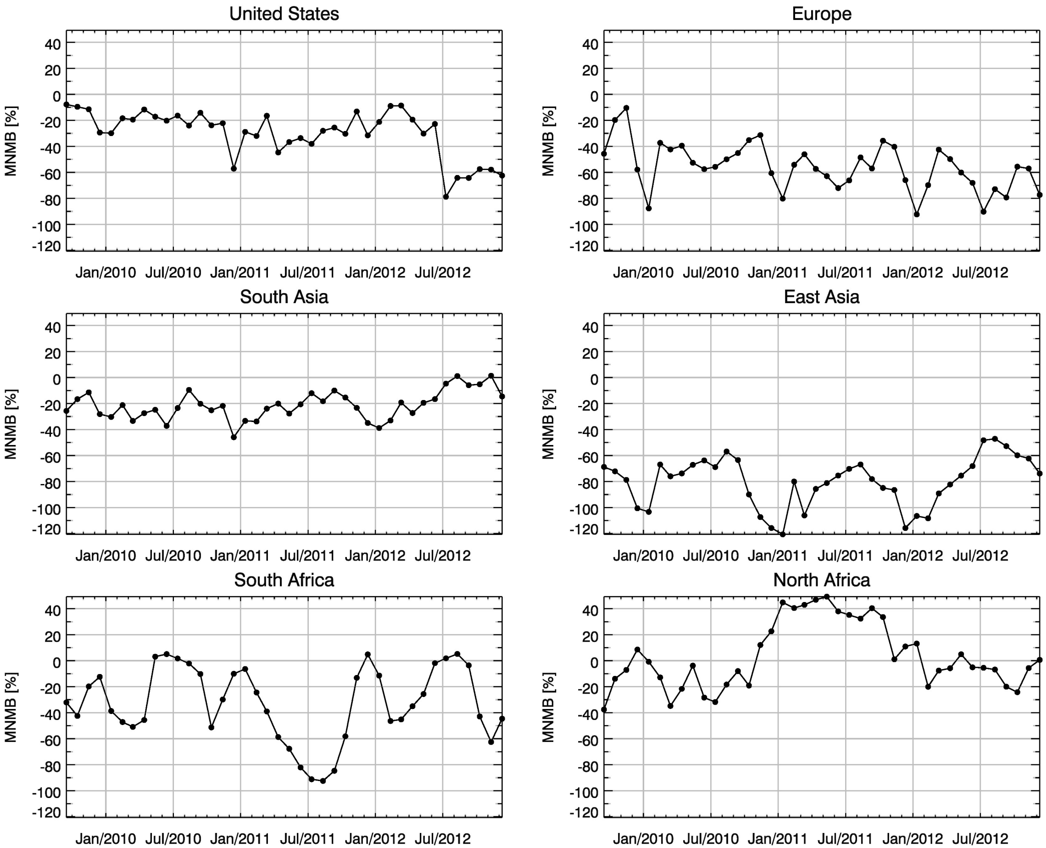

Fig. 1.

ACPD
15, C4439–C4454, 2015

Interactive Comment

ACPD
15, C4439–C4454, 2015

Interactive Comment

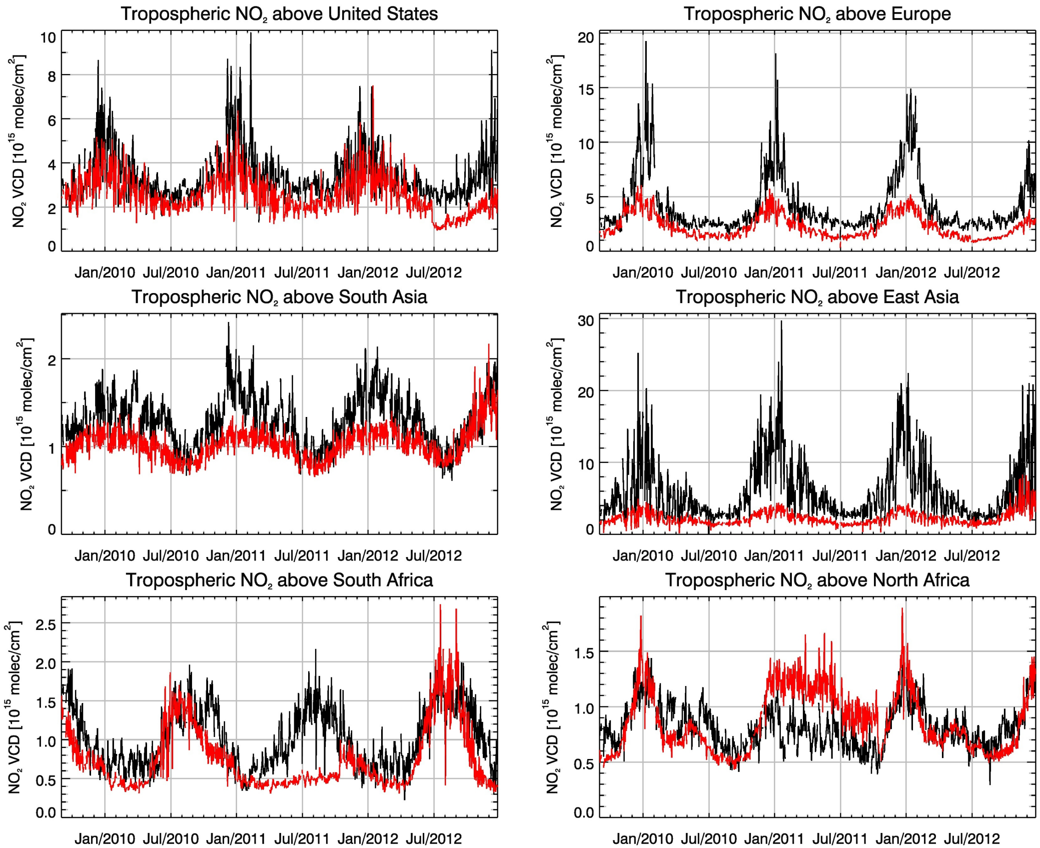

Fig. 2.

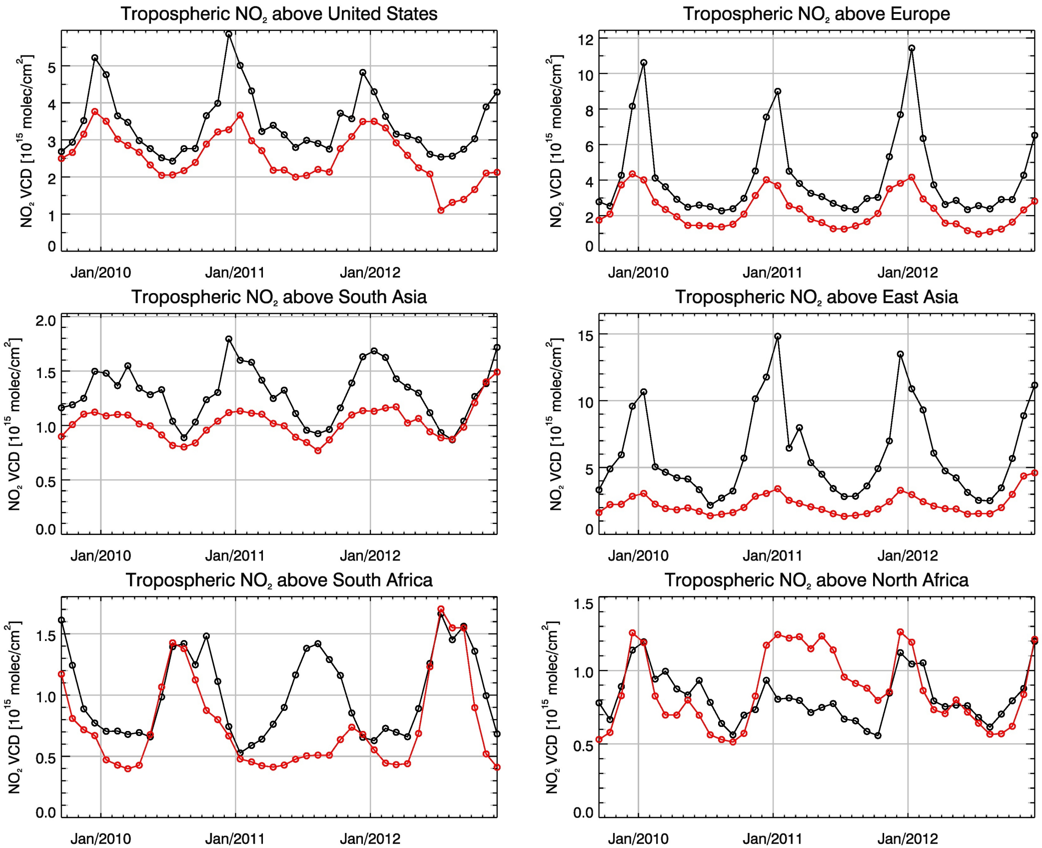

Fig. 3.

ACPD
15, C4439–C4454, 2015

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper