Reviewer Responses and Manuscript Changes:

All editions to the document which quote a page number and line number are based on the updated manuscript page/line numbering.

Reviewer 1 Response:

We thank the reviewer for his/her comments. These comments are repeated below in *italics*. Our responses follow in normal text.

1-1: *Biogenic emission: Biogenic species (i.e., isoprene, mono-terpene, etc) play an important role even in NOx chemistry by controlling OH radical concentrations. For example, if isoprene emissions are overestimated, the estimated levels of NO2 can be higher than the actual levels of NO2 (because of slow NOx losses). The tropospheric chemistry can, sometimes significantly, influence the NOx analyses for summer episode. Thus, authors should clarify which biogenic inventory (e.g., MEGAN, GEIA, etc) was used in your study and explain how the biogenic inventory is uncertain (or reliable).*

As stated in Savage et al (2013) "Biogenic emissions of isoprene are from the monthly climatological data of Poupkou et al. (2010) at 0.125◦ × 0.0625◦ resolution. The use of climatological emissions for biogenic isoprene sources will diminish the ability of the model to respond to increased biogenic ozone precursor emissions during episodes, but this is not expected to be a major factor in the cases analysed in this paper. An interactive biogenic isoprene emission scheme is under development, but is not yet available for use in AQUM."

The following text will be added at the end of Section 3.1: "Poupkou et al. (2010) provide the monthly climatology of biogenic emissions on a 0.125◦ × 0.0625◦ resolution. The use of climatological biogenic isoprene emissions will partially diminish AQUM's representation of ozone from biogenic precursors. A new interactive biogenic isoprene scheme is under development by was not available for this study. However, this is a secondary issue in this paper as we focus on primary emissions of NO_x, which affect ozone concentrations".

1-2: *Biomass burning emission: I wonder whether biomass burning emission was considered in the CTM simulations. If it was considered, authors need to mention/describe it.*

Again to quote from Savage et al "Biomass burning emissions of aerosols are taken from year 2000 values from the Global Fire Emissions Database (GFED) version 1 (Randerson et al., 2005). The choice of 2000 emissions is somewhat arbitrary, but these emissions have relatively little impact on our domain."

We will add the following text at the end of Section 3.1 also: "Biomass burning emissions of aerosols come from the Global Fire Emissions Database (GFED) version 1 (Randerson et al., 2005) for 2000. The use of biomass burning emissions from 2000 is somewhat arbitrary, but within AQUM's domain these emissions have relatively little impact.

1-3: *Variations of NO_x emissions: This issue can be an important factor in the analysis of seasonal trends of columnar NO2. Authors should provide the seasonal (or monthly)* *information of NOx emission. Also, this information can be useful to analyse the seasonal trends of columnar NO2 over the London and northern England regions in Fig.3.*

We will add a new figure in line with the reviewer comments at the end of Section 3.1 in the discussion of AQUM NO_x emissions. The figure shows the seasonal cycle (fraction) applied to the AQUM annual NO_x (point and area sources) emissions total from the NAEI, ENTEC and EMEP. The seasonal cycle comes from Visschedijk et al., (2007).

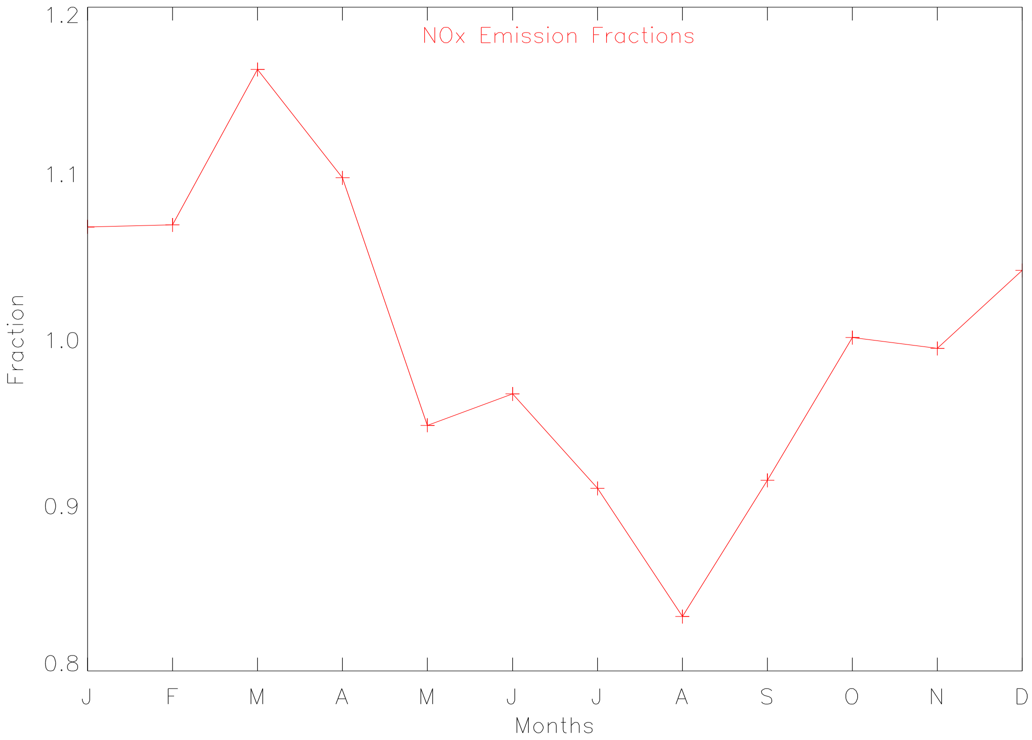


Figure: NO_x emissions seasonal cycle, based on Visschedijk et al., (2007), which is applied to AQUM's NO_x emission annual totals.

Higher columnar NO_2 over London and northern England from OMI observations are well captured by 3D-CTM simulation. However, usually, both NO_2 columns from the CTM and satellite observations during winter are higher than those during summer (e.g., van Noije et al., 2006; Huijnen et al., 2010). Authors should explain the unusual trends of the NO_2 columns (higher NO_2 columns during summer) over the regions shown in Fig. 3.

This is true over continental Europe, however, as discussed by Pope et al., (2014) and van der A et al., (2008), the meteorological variability over the UK leads to different column NO_2 signals over source regions. van der A et al., (2008) suggest that peak UK NO_x emissions occur in July. Pope et al., (2014) suggest that the transport of column NO_2 away from source regions due to strong winter dynamics outweighs the loss of UK source region column NO_2 from enhanced summer photochemistry. Therefore, UK source region column NO_2 is larger in summer than winter.

In Section 4.1 on page 17, line 9, the following text will be added: "van der A et al., (2008) suggest that peak UK NO_x emissions occur in July, while Pope et al., (2014) suggest that the transport of column NO_2 away from source regions due to strong winter dynamics outweighs the loss of UK source region column NO_2 from enhanced summer photochemistry.”

2-1: *Chemical LBCs: In the manuscript, authors mentioned that using chemical GEMS gives better result and it is consistent with the findings of Savage et al. (2013). However, authors do not mention what makes it better. What are the main differences between the GEMS and MACC LBCs? If the MACC is an improved version of the GEMS, what has been improved? Give some more detailed information on both the GEMS and MACC.*

The MACC reanalysis used a more recent version of the ECMWF model (Integrated Forecast System), assimilated more satellite products and was run at a resolution of 80 km instead of 125 km. Savage et al found that the ozone bias from January to May when using GEMS to provide LBCs data had a larger negative bias than those made with the MACC LBCs. From May to the end of the year runs using the GEMS LBCs generally perform better, with a smaller positive bias.

The following sentence will be added in Section 3.2 on page 14, line 8: “The MACC reanalysis uses a more recent version of the ECMWF model (Integrated Forecast System), assimilated more satellite products and was run at a resolution of 80 km instead of 125 km.”

2-2: *E2 (idealized point source tracer): In the E2 sensitivity test, authors tried to examine the spatial patterns of the tracer. The experiment remains unclear in the manuscript. i) How to idealize tracer from the point sources? ii) Which species are used as a tracer? It appears to be a “reactive species” having one-day lifetime. iii) What are the reasons to determine a tracer having one-day lifetime? The lifetime of NOx is less than 1 day (say, several hours) during summer (Schaub et al., 2007; Lamsal et al., 2010). iv) In Fig. 7-c, the columnar NO2 over other regions seems to be “zero”. Did you consider only tracer’s emission for this sensitivity simulation? If yes, what is the point of considering one trace species specially having one day lifetime? If no, was the tracer tagged for identification and how to tag the tracer in your model?*

We will improve the explanation of the implementation of this tracer in the manuscript to address these points. To answer the reviewer’s specific comments:

i) The tracer is emitted at the location of UK NO_x point source emissions with emissions which are the same as those of NO_x from the power station sources in the model inventory.

ii) The idealised e-folding tracer is a separate transported tracer, using the same tracer transport scheme as the chemical tracers but not having any interactions with the chemistry scheme. It has a fixed lifetime, with a first-order loss applied to the tracer at every model time-step.

iii) The choice of a lifetime of 1 day is somewhat arbitrary and is indeed longer than the lifetime of NO_x in summer at this latitude. We did not have the resources available to run tracers with a range of lifetimes for this case in order to estimate the most appropriate lifetime. However, we do not believe that this will have a significant impact on the results.

iv) Figure 7c only represents the idealised tracer field from UK NO_x point source emission sources. Therefore, the concentrations over Europe and the sea are near zero. Yes, we only consider the tracer for Run E2 to look at the correlation of the point source emissions and mean biases in Figure 6a. This is not a "tagged" NO_x tracer, as stated above – it is an idealised tracer with a simple first-order loss rate. The aim of this part of the study was to test whether there was a link between the emissions from the power stations and the observed areas of bias.

In line with comments i, ii, iii, the sentence on page 14, line 20 "Run E2 uses an idealised passive tracer from the point sources with a lifetime of one day to examine if the tracer columns correlated with summer AQUM-OMI positive biases (see Sect. 4.3)." will be modified to "Run E2 introduces a new idealised passive tracer emitted from the UK point sources with the same emissions to that of the model NO_x inventory. The idealised tracer can be transported like any chemical tracer, but cannot be lost through chemical reactions. Instead it is lost through its e-folding lifetime of one day. The point source tracer columns can then be examined to see if they correlated with summer AQUM-OMI positive biases (see Sect. 4.3). For comment iv, a new line has been added in Section 4.3 on page 19, line 21: "The minimum tracer values of 0 $\times 10^{15}$ molecules/cm^2 are over the sea and continental Europe as there is no emission of the tracer there.

2-4: *Reaction probability of N_2O_5 (γN_2O_5): As mentioned in the manuscript, usually, the reaction probability of N_2O_5 is known to range from 0.02 to 0.001 (Riemer et al., 2003; Brown et al., 2006). It may exist somewhere between the two values. What are the reasons to choose only two extreme cases in the sensitivity simulations?*

We believe the range of γ between 0.001-0.02 is not extreme in the case of air quality modelling. Multiple air quality models use γ values between 0.0-1.0. For instance, Foley et al., (2010) used γ ranges between 0-0.06 in CMAQ and Menut et al., (2013) used γ values between 0.01-1.0 in CHIMERE. Macintyre and Evans (2010) explore the sensitivity of N_2O_5 uptake on aerosol for multiple γ values between 0.0-1.0. They state that intermediate γ values between 0.001-0.02 are the most sensitive values to this process. Therefore, we base the γ range in our study on Macintyre and Evans (2010). We did not explore other values of γ as this was not the main focus of this study.

3: *When the black polygonal regions are determined in Fig. 4, authors used the two values of MB and satellite error. Here, the satellite error is the criteria (i.e., the magnitude of the MB greater than the satellite error). In that case, how did you consider the satellite error (i.e., averaged value over domain or each pixel value)? For better understanding this analysis, authors should provide some ranges of satellite errors over specific regions (e.g. northern England + London, Benelux, Po valley, North Sea, etc).*

The satellite error is the average retrieval error with the random error component reduced, using the methodology in Section 2.2, for each pixel. A second table will be added to the revised manuscript to give information on some of the locations and example average retrieval errors. Please see the response to Reviewer 2 for more information.

4: *Where is Dartmoor located in England? For the sake of reader's convenience, you would better provide some geographical information (e.g., Dartmoor, Irish Sea, North Sea, Po Valley, and many regions mentioned in the manuscript), possibly in Fig. 2.*

Please see the new table discussed in **3:** and the response to Reviewer 2.

Specific Comment 1: *Merging two figures 3 and 4 in a 2 x 3 panel*

We did consider this before submission, but feel that if the figures were merged the images would become cramped and unclear. Therefore, we prefer to leave the figures as they are.

Specific Comment 2: *P. 21763, line 23 and P. 21764, line 4 (i.e., "around 0 – 3x10^{15}" and "between 0 -6 x10^{15}"). The minimum background columnar NO_2 over background could not be "zero". Authors should provide approximate values.*

These values will be changed to the minimum background values for the discussion of Figure 3 to "O(10^{13}) -3 x10^{15} molecules/cm^2, where O(10^{13}) represents values in size of the order of 10^{13}" and "O(10^{13}) - 6 x10^{15}" in Section 4.1 on page 16, line 21, and on page 17, line 2, respectively.

Specific Comment 3: *It is a quite interesting that by introducing N_2O_5 heterogeneous chemistry, these positive biases were "significantly" reduced even during the summer episode, indicating that columnar NO_2 are significantly decreased. N_2O_5 is thermally unstable and decomposes to NO_3 and NO_2 at high temperatures. In other words, during summer, the decomposition of N_2O_5 is more active than the formation of N_2O_5. It would not have a significant impact on the columnar NO_2 during summer. Thus, you would better provide other reasons in this part.*

We agree with this comment over the rest of the domain in summer. However, over northern England, there is the largest loading of NO_x and aerosols coming from the point sources. Therefore, even with decreased levels of N_2O_5, the model shows that this process is a significant sink of NO_x. This process also occurs above the surface and in the free troposphere due to vertical transport and emissions entered in the model at 80, 180 and 320 metres (representing the emission from power stations with tall stacks). The temperatures are cooler here, so the thermal decomposition of N_2O_5 is reduced providing conditions suitable for heterogeneous chemistry to take place with the higher loading of NO_x and aerosols. Therefore, we feel that the original conclusions stand and suggest no changes to the paper based on this comment.

References:

Foley, K., Roselle, S., Appel, K., Bhave, P., Pleim, J., Otte, T., Mathur, R., Sarwar, G., Young, J., Gilliam, R., et al.: Incremental testing of the Community Multiscale Air Quality (CMAQ) modelling system version 4.7, Geoscientific Model Development, 3, 205–226, 2010.

Macintyre, H. and Evans, M.: Sensitivity of a global model to the uptake of N_2O_5 by tropospheric aerosol, Atmospheric Chemistry and Physics, 10, 7409–7414, 2010.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R. and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geoscientific Model Development, 6, 981–1028, doi:10.5194/gmd-6-981-2013, http://www.geosci-model-dev.net/6/981/2013/, 2013.

Pope, R., Savage, N., Chipperfield, M., Arnold, S. and Osborn, T.: The influence of synoptic weather regimes on UK air quality: analysis of satellite column NO_2, Atmos. Sci. Lett., 15, 211–217, doi:10.1002/asl2.492, 2014.

Poupkou, A., Giannaros, T., Markakis, K. Kioutsioukis, I., Curci, G., Melas, D. and Zerefos, C.: A model for European biogenic volatile organic compound emissions: Software development and first validation, Environ. Modell. Softw., 25, 1845–1856, 2010.

Randerson, J. T., Kasibhatla, P. S., Kasischke, E. S., Hyer, E. J., Giglio, L., Collatz, G. J. and van der Werf G. R.: Global fire emissions database (GFED), version 1, available at: http://daac.ornl.gov (last access: 2 October 2012), 2005.

van der A, R., J., Eskes, H., J., Boersma, K., F., van Noije, T., P., C., Van Roozendael, M., De Smedt, I., Peters, D., H., M., U. and Meijer, E., W.: Trends, seasonal variability and dominant NO_x sources derived from a ten year record of NO_2 measured from space, Journal of Geophysical Research, 113, 12 pp, DOI:10.1029/2007JD009021, 2008.

Visschedijk, A., Zanveld, P. and van der Gon, H.: A high resolution gridded European emission database for the EU integrated project GEMS, TNO report 2007-A-R0233/B., 2007.

Reviewer 2 Response:

We thank the reviewer for his/her comments. These comments are repeated below in italics. Our responses follow in normal text.

1. *"In equations 1-3, y is a scalar, and A and x are vectors and the notation should be adapted accordingly (e.g. in boldface or with a vector sign)."*

The A and x vectors will be changed to bold font.

2. *"A discussion of why the kernel values over Dartmoor are lower than over London is missing. This should be explained. Kernels depend on a range of aspects: albedo, viewing geometry, clouds, and also location. So what do we see when kernels over Dartmoor and London are different? Are albedo, geometry and clouds all similar between those two locations, and do we see the effect of the different* NO_2 *loading between Dartmoor and London?"*

The retrievals have been filtered out for cloud cover under 20% over both London and Dartmoor. Therefore, cloud cover is unlikely to be causing the AK differences between the two sites. The viewing angle is unlikely to have a large affect on Dartmoor and London AKs as NO_2 will be retrieved from the different locations at various viewing angles depending where OMI is in its orbit. We have looked at the surface albedo data from the OMI satellite files for 2006 and the patterns were noisy with no clear differences between the locations. Eskes and Boersma (2003) state that AKs are independent of trace gas distributions for optically thin absorbers (Eqn 15). However, for stronger absorbers they suggest that the AK depends on the true distribution of the tracer. Therefore, as London column NO_2 is greater than that of Dartmoor, i.e. regions of optically strong and thin absorbers, and the AKs are more and less sensitive, respectively.

In the revised text in Section 2.1 (page 8, lines 7-18) we will add a description of the properties influencing NO_2 AKs and which we believe are dominating factors in our case study (i.e. NO_2 loading and not cloud cover, viewing angle or surface albedo).

3. *"More generally, it is unclear what the key message is for section 2.1 and Figure 1. The authors should reconsider this section and their discussion – what is the take-home message?"*

The general purpose of section 2.1 is to discuss what an averaging kernel is, how it behaves with altitude and how it will modify the model NO_2 profiles. In the revised paper the section will start with the following sentence to inform the reader what the section aims to communicate: "Model Transfer Functions (MTF), known as "averaging kernels", allow for direct comparison between model column NO_2 and satellite retrievals. This section introduces how these MTF modify model vertical profiles and how they vary in season and location."

4. *"On page 21757 it should be clarified what causes the systematic error in the slant column. Table 1 in Boersma et al. [2004] provides this information."*

On page 9, line 13-14, we will add the sentence "The sources of systematic error in the total slant column include the NO_2 cross-section, spectral calibration, solar diffuser and temperature (Boersma et al., 2004).

5. *"Related to this, the systematic error in the slant column should not be included in Eq. (6) as it does not propagate into a tropospheric column error. Any systematic error in the NO_2 slant column density will be largely absorbed by the stratospheric assimilation procedure. Subtracting from the (biased) slant column a similarly biased stratospheric slant column results in a small tropospheric residual slant column. See e.g. Belmonte-Rivas et al. [2014]."*

Having read Belmonte-Rivas et al. (2014), we have removed the σ_{strat} term from Eqns 7-10. We will discuss the σ_{strat} term in Eqn 6 and then explain why it needs to be removed stating the reasons described by Belmonte-Rivas et al. (2014) on page 9, line 23. The σ_{strat} term has been removed from the algorithm and this had a limited impact on the number of significant pixels in Figures 4-8 as it is a small term in the error budget. New versions of Figures 4-8 have been created to account for this limited effect on the number of significant pixels in the bias plots.

6. *"However, one problem in the way they present their results, is that what they define now as the systematic error (caused by errors in the AMF), is not a true systematic error in the sense that it is a persistent and always works in the same way. The authors use the notion that the retrieval error due to AMF errors is 40% of the tropospheric column and systematic. But the AMF error changes in time (on day 1 albedo knowledge may be accurate and on day 2 albedo may be off by +0.02), and in space (for pixel 1 the cloud fraction may be underestimated, for pixel 2 it may be overestimated), and in practice at least part of these individual systematic errors may also cancel in a temporal or spatial average. The formulation in Eq. (10) leaves no room for this now, i.e. all AMF error is 40% and always proceeds systematically. The difficulty obviously is in estimating to what extent the systematic error cancels when averaging over time, but validation studies have provided some hints here (individual comparisons may be off by up to 40% but averaged over an ensemble the systematic differences are more on the order of 20%, e.g. Irie, Hains, Schaub-papers)."*

We agree with the reviewer that the uncertainty in the AMF is not totally systematic and that an average ensemble of AMF uncertainties will reduce the overall error. However, we struggled to find the value in the "*order of 20%*" in the references suggested by the reviewer. We found how the uncertainties in multiple factors such as the NO_2 profiles and surface albedo, that are used to calculate the AMF, lead to uncertainty in the vertical tropospheric column. However, these studies did not give any direct insight to the error associated in the AMF. The Irie et al., (2012) paper does discuss the 10-40% error quoted by Boersma et al., (2004), but we have already accounted for this here. Therefore, we feel that we should leave the AMF uncertainty as the conservative estimate of 40% in our study.

7. *"In the revised manuscript the above concern should be taken into account, and the authors in their discussion of Figure 2 should also indicate the magnitude of the original errors (without subtracting the pure random contribution). There should also be a discussion of why the errors appear to reduce more over continental Europe than over Sea. And why is the reduction stronger in summer than in winter? Related to more samples in summer?"*

In line with Reviewer 1's comments (3. Satellite error) on providing examples of average column NO_2 errors we have added a second table, which describes the location (lat, lon), mean column NO_2, mean error and mean error with the random error comment reduced for London, the Benelux region, northern England, the Po Valley, North Sea and Scandinavia in summer and winter. This table is reproduced below.

Place	Column NO_2		Column NO_2 Error		This Study	
	Summer	Winter	Summer	Winter	Summer	Winter
London 1°W-1°E, 51-51.5°N	9.86	10.7	9.68	9.13	4.24	4.52
Benelux 3-7°E, 50.5-52.5°N	9.57	11.4	7.09	9.24	3.98	4.82
Po Valley 7-9°E, 44.25-45.5°N	3.35	11.9	2.44	9.88	1.42	4.66
Northern England 3-0°W, 52.5-54°N	8.11	8.06	7.13	6.56	3.47	3.42
North Sea 0-8°E, 54-60°N	1.48	2.22	1.94	2.12	0.86	1.01
Scandinavia 6-16°E, 54-63°N	1.48	2.10	1.49	2.12	0.74	1.16

Table: The average column NO_2, column NO_2 error and column NO_2 error calculated by this study for multiple locations across Europe in summer and winter (x10^{15} molecules/cm^2).

The average summer satellite retrieval error reduces more than winter because the samples are larger. As summer is more prone to clear sky conditions than winter, less retrievals are filtered out in summer, the sample is bigger, and there is a larger reduction in the random error component ($N_{summer} > N_{winter}$).

When looking at the frequency of pixels in grid box averages, the frequency is higher over the sea than land. Therefore, following the same argument for summer and winter, the larger samples over the sea result in lower average retrieval errors.

The following sentence will be added on page 11, line 17-21, "The error in summer, compared with winter, and the error over sea in comparison to land, are smaller. We suggest that the larger sample size in summer and over the sea, when compared to winter and over the land, respectively, reduces the random error component further has N is larger."

8. *"P21752, L1-2: how can models 'detect' shipping lanes?"*

This line will be replaced with "Several of the regional models successfully **simulated** the shipping lanes seen by OMI."

9. *"P21752, L18: 5-6% overall or per year?"*

This is 5-6% per year and will be change to "5-6% per year" in the revised manuscript.

10. *"P21754, L2-3: please clarify the statement here as there was not much of an OMI row anomaly in 2006."*

We filter for the OMI row anomaly indirectly. All retrievals that are flagged as “-1” are of poor quality and not used in our analysis. This includes the influence of the OMI row anomalies in some pixels and are automatically set to “-1” using the Braak (2010) algorithm during the quality control process. Therefore, we are not sure of the OMI row anomaly effect on the number of pixels filtered out of our analysis, but believe it to be small in this time period. We will add this to our discussion, page 5 on line 26-27, to highlight to readers that this is accounted for when we filter out the poor quality pixels.

11. *“P21754, L7-11: the Irie et al. [2008] citation is irrelevant here, since this concerns validation of another product than used here. The paper by Irie et al. [2012] is the appropriate paper that discussed the validation DOMINO v2.0 data.”*

In line with the reviewer’s comments, the Irie et al., (2008) reference, stating *“Irie et al. (2008) compared OMI and with ground based MAX-DOAS retrievals in the Mount Tai Experiment (2006). They found the standard OMI product (version 3) overestimated the MAX-DOAS measurements by approximately* 1.6×10^{15} *molecules* cm^{-2} *(20 %), but within the OMI uncertainty limits.”* will be replaced with the Irie et al., (2012) reference. The new text will read “*Irie et al., (2012) compare SCIAMACHY, OMI and GOME-2 tropospheric column* NO_2 *with surface MAX-DOAS column* NO_2 *observations between 2006 and 2011. They found the instruments are biased by −5 ± 14%, −10 ± 14%, and +1 ± 14%, respectively, which the authors suggest are all small and insignificant.*”, on page 6, line 3-7.

12. *“P21754, L17-19: suggest to clarify ‘the true vertical distribution’.”*

This refers to the real vertical profile of a tracer (e.g. NO_2) in the atmosphere.

13. *“P21756, L4: ‘with trap’?”*

This will be corrected to “will trap”.

14. *“P21758, L8-10: how can the left-hand side of Eq. (8) become negative? The authors should explain in what sort of situations this happens, and why.”*

In Eqn 8, if the error in the tropospheric column NO_2 is small and either the N_{trop} or X_{total} terms are large, then the left hand side can be negative. The σ_{strat} term has been removed. As X_{total} is scaled by 0.03, this term is likely to be smaller than the σ_{tropak} term in most cases. However, there will be cases where the N_{trop}, scaled by the AMF uncertainty value of 0.4, will to be sufficiently large, with low uncertainty, and the left hand side of Eqn 8 will become negative. From our analysis though, this is rare and the error is set to 50% as discussed with H. Eskes (2012) via personal communication.

The following text “(e.g. when N_{trop} is large, but has small uncertainty - X_{total} is scaled by 0.03 so will be small compared to N_{trop})” will be added to page 10, line 16, after “left hand side is negative”.

Manuscript prepared for Atmos. Chem. Phys. Discuss.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 23 January 2015

Evaluation of a regional air quality model using satellite column NO_2: treatment of observation errors and model boundary conditions and emissions

R. J. Pope[1], M. P. Chipperfield[1], N. H. Savage[2], C. Ordóñez[2], L. S. Neal[2], L. A. Lee[1], S. S. Dhomse[1], and N. A. D. Richards[1]

[1]School of Earth and Environment, University of Leeds LS2 9JT, Leeds, UK
[2]Met Office, Exeter, UK

Correspondence to: R. Pope (eerjp@leeds.ac.uk)

Abstract

We compare tropospheric column NO_2 between the UK Met Office operational Air Quality in the Unified Model (AQUM) and satellite observations from the Ozone Monitoring Instrument (OMI) for 2006. Column NO_2 retrievals from satellite instruments are prone to large uncertainty from random, systematic and smoothing errors. We present an algorithm to reduce the random error of time-averaged observations, once smoothing errors have been removed with application of satellite averaging kernels to the model data. This reduces the total error in seasonal mean columns by ~~30–90~~30–70 %, which allows critical evaluation of the model. The standard AQUM configuration evaluated here uses chemical lateral boundary conditions (LBCs) from the GEMS (Global and regional Earth-system Monitoring using Satellite and in-situ data) reanalysis. In summer the standard AQUM overestimates column NO_2 in northern England and Scotland, but underestimates it over continental Europe. In winter, the model overestimates column NO_2 across the domain. We show that missing heterogeneous hydrolysis of N_2O_5 in AQUM is a significant sink of column NO_2 and that the introduction of this process corrects some of the winter biases. The sensitivity of AQUM summer column NO_2 to different chemical LBCs and NO_x emissions datasets are investigated. Using Monitoring Atmospheric Composition and Climate (MACC) LBCs increases AQUM O_3 concentrations compared with the default GEMS LBCs. This enhances the NO_x-O_3 coupling leading to increased AQUM column NO_2 in both summer and winter degrading the comparisons with OMI. Sensitivity experiments suggest that the cause of the remaining northern England and Scotland summer column NO_2 overestimation is the representation of point source (power station) emissions in the model.

1 Introduction

Air quality has a major influence on the UK both socially and economically. It can result in approximately 50 000 premature deaths per year and an average reduction in life expectancy of 7–8 months (HoC, 2010). Air pollution health effects include lung disease and cancer, cardiovascular problems, asthma and eye irritation (WHO, 2011). In 2005, poor UK air quality cost £ (€) 8.5 (10.7)–20.2 (25.5) billion and between 2007–2008 there were 74 000 asthma-related hospital admissions. Overall, these air quality-asthma incidents cost society £ (€) 2.3 (2.9) billion (HoC, 2010). Poor air quality associated with ozone concentrations over 40 ppbv can also significantly reduce crop yields e.g. Hollaway et al. (2012).

Therefore, regional models have been developed to predict hazardous levels of air pollution to help inform the public and to allow local authorities to take action to reduce/accommodate the respective health risks/effects. Air quality models have mainly been evaluated against surface observations, e.g. Savage et al. (2013). ~~However, recently~~ Recently such models have also been compared with satellite observations, taking advantage of the better spatial coverage despite the potentially large error of individual observations. In the past NO_2 satellite data has been compared mainly with global atmospheric chemistry models (e.g. Velders et al., 2001; Lauer et al., 2002; Van Noije et al., 2006). ~~However, more~~ More recently, other studies have used satellite data to evaluate models on a regional scale. Savage et al. (2008) investigated European tropospheric column NO_2 interannual variability (IAV) ~~,~~ during 1996–2000 ~~,~~ by comparing GOME with the TOMCAT chemical transport model (CTM) (Monks et al., 2012). The best comparisons were found in the JFM and AMJ seasons, especially over western Europe. They also found that synoptic meteorology had more influence on NO_2 IAV than NO_x emissions did.

Huijnen et al. (2010) compared Ozone Monitoring Instrument (OMI) tropospheric column NO_2 against a European global-regional air quality model ensemble median for 2008–2009. The ensemble compared better with the OMI data than any ~~one~~ individual model, with good agreement over the urban hotspots. Overall, the spread in the models was greatest in the summer (with deviations from the mean OMI tropospheric column in the range 40–62 %), due to the more active NO_x chemistry in this season and the differences in chemistry schemes among the contributing models, when compared to winter (20–34 %). Several of the regional models successfully ~~detected~~ simulated the shipping lanes seen by OMI.

Han et al. (2011) investigated tropospheric column NO_2 over the Korean Peninsula through comparisons between OMI data and the Community Multi-scale Air Quality Model

(CMAQ) (Foley et al., 2010). In summer, North and South Korea had similar column NO_2 from both the model and observations. ~~However, in winter~~In winter, South Korea, a more developed nation with greater infrastructure, had significantly greater NO_2 concentrations than North Korea. Overall, CMAQ overestimated OMI NO_2 concentrations by factors of 1.38–1.87 and 1.55–7.46 over South and North Korea, respectively.

Other studies investigating regional tropospheric column NO_2 through model simulations and satellite observations include Blond et al. (2007), Boersma et al. (2009) and Curier et al. (2014). Blond et al. (2007) compared CHIMERE 3-D CTM and SCIAMACHY column NO_2 over western Europe~~with good~~; they found reasonable agreement with winter and summer correlations of 0.79 and 0.82, respectively. Boersma et al. (2009) used the GEOS-Chem 3-D CTM to explain the seasonal cycle in SCIAMACHY and OMI column NO_2 over Israeli cities, with larger photochemical loss of NO_2 in summer than winter. ~~And~~ Curier et al. (2014) used a ~~synergistic~~ combination of OMI and the LOTOS-EUROS 3-D CTM to evaluate NO_x trends finding negative trends of 5–6 % per year over western Europe.

The UK Met Office's Air Quality in the Unified Model (AQUM) is used for short operational chemical weather forecasts of UK air quality. Savage et al. (2013) performed the first evaluation of the AQUM operational forecast for the period May 2010–April 2011 by using surface O_3, NO_2 and particulate matter observations from the UK Automated Urban and Rural Network (AURN) (DEFRA, 2012). Among other model-observation metrics they used the mean bias (MB), root mean square error (RMSE), modified normalised mean bias (MNMB) and the Fractional Gross Error (FGE) (Seigneur et al., 2000). See the Appendix for the definition of these metrics.

Savage et al. (2013) found that AQUM overestimated O_3 by $8.38\,\mu g\,m^{-3}$ (MNMB = 0.12), with a positive bias at urban sites but no systematic bias at rural sites. The model-observation correlation was reasonably high at 0.68. For NO_2, there was a bias of $-6.10\,\mu g\,m^{-3}$, correlation of 0.57 and MNMB of -0.26. At urban sites there was a large negative bias while rural sites had marginal positive biases. The coarse resolution of AQUM (12 km) led to an underestimation at urban sites because the model NO_x emissions are instantaneously spread over the entire grid box. The particulate matter (PM_{10}) prediction skill was lower with a correlation and bias of 0.52 and $-9.17\,\mu g\,m^{-3}$, respectively.

The aim of this paper is to evaluate AQUM using satellite atmospheric trace gas observations. The Met Office has previously compared the skill of AQUM only against AURN surface measurements, which in the case of NO_2 are not specific and include contributions from other oxidised nitrogen compounds (see Savage et al. (2013), and references therein). Therefore, for better spatial model-observation comparisons and to minimise the effect of measurement interferences, we use satellite observations over the UK. We focus on tropospheric column NO_2 data from OMI for the summer (April–September) and winter (January–March, October–December) periods of 2006. Section 2 describes the OMI satellite data used and gives a detailed account of our error analysis which determines how we can use satellite data to test AQUM. Section 3 describes AQUM and the model experiments performed. Results from the model-observations comparisons are given in Sect. 4. Section 5 presents our conclusions.

2 Satellite data

OMI is aboard NASA's EOS-Aura satellite and has an approximate London daytime overpass at 13:00 LT. It is a nadir-viewing instrument with a pixel size of $312\,km^2$ and $3240\,km^2$ along track and across track, respectively (Boersma et al., 2008). We have taken the DOMINO tropospheric column NO_2 product, version 2.0, from the TEMIS (Tropospheric Emissions Monitoring Internet Service) website, http://www.temis.nl/airpollution/no2.html (Boersma et al., 2011b, a). We have binned NO_2 swath data from 1 January to 31 December 2006 onto a daily 13:00 LT 0.25° × 0.25° grid between 43–63° N and 20° W–20° E. All satellite retrievals have been quality controlled, and retrievals/pixels with geometric cloud cover greater than 20 % and poor quality data flags (flag $= -1$) were removed. The product uses the algorithm of Braak (2010) to identify OMI pixels affected by row anomalies and sets the data flags to -1. ~~We also filter these out in this study. Even though~~ Therefore, indirectly, we are filtering out the OMI row anomalies as well, even though this effect was limited in 2006. OMI has an approximate 13:00 LT London overpass, but we used all OMI retrievals in the domain between 11:00 and 15:00 LT to get more extensive spatial coverage. Several studies have validated OMI column NO_2 against surface and aircraft measurements of tropospheric column NO_2. ~~? compared OMI and with ground based~~ Irie et al. (2012) compared SCIAMACHY, OMI and GOME-2 tropospheric column NO_2 with surface MAX-DOAS ~~retrievals in the Mount Tai Experiment (~~column NO_2 observations between 2006 ~~).~~ and 2011. They found the ~~standard OMI product (version 3) overestimated the MAX-DOAS measurements by approximately 1.6×10^{15}(20), but within the OMI uncertainty limits~~instruments are biased by -5 ± 14%, -10 ± 14%, and +1 ± 14%, respectively, which the authors suggest are all small and insignificant. Boersma et al. (2008) compared the near real time (NRT) OMI product (version 0.8) with aircraft measurements in the INTEX-B campaign. Overall, they found a good correlation (0.69) between OMI and the aircraft column NO_2, with no significant biases. Therefore, we have confidence in the OMI column NO_2 and use it for evaluation of our model.

2.1 Satellite averaging kernels

~~Eskes and Boersma (2003) define the averaging kernel~~ Model transfer functions (MTF), known as "averaging kernels (AK)", allow for direct comparison between model column NO_2 and satellite retrievals. This section introduces how these MTF (AK) modify model vertical profiles and how they vary in season and location. Eskes and Boersma (2003) define the AK to be a relationship between the retrieved quantities and the true distribution of the tracer ~~.~~ (i.e. the vertical profile of a chemical species). In other words, the satellite instrument's capability to retrieve a quantity is a function of altitude. For instance, the instrument may be more or less sensitive retrieving a chemical species near the boundary layer than in the stratosphere. Therefore, since satellite retrievals and model vertical profiles are not directly comparable, the AK (or weighting function) is applied to the model data, so the sensitivity of the satellite is accounted for in the comparisons. The AK comes in different forms for different retrieval methods. For the Differential Optical Absorption Spectroscopy (DOAS) method, the AK is in the form of a column vector, while in Optimal Estimation, the

AK is a matrix ~~,~~ whose dimensions depend on the number of pressure levels in the retrieval process.

The OMI retrievals use the DOAS technique and the AK is a column vector. Following Huijnen et al. (2010) and the OMI documentation (Boersma et al., 2011a), the AKs are applied to the model as:

$$y = \mathbf{A} \cdot \mathbf{x} \tag{1}$$

where y is the total column, ~~A~~ $\mathbf{A}$ is the AK and ~~x~~ $\mathbf{x}$ is the vertical model profile. However, here the tropospheric column is needed:

$$y_{\text{trop}} = \mathbf{A}_{\text{trop}} \cdot \mathbf{x}_{\text{trop}} \tag{2}$$

where ~~A_{trop} is:~~ $\mathbf{A}_{\text{trop}}$ is:

$$\mathbf{A}_{\text{trop}} = \mathbf{A} \cdot \frac{\text{AMF}}{\text{AMF}_{\text{trop}}} \tag{3}$$

AMF is the atmospheric air mass factor and AMF_{trop} is the tropospheric air mass factor. For the OMI product, Huijnen et al. (2010) state the AK tends to be lower than 1 in the lower troposphere (e.g. 0.2–0.7 up to 800 hPa) and greater than 1 in the mid-upper troposphere. Therefore, the OMI AKs reduce model NO_2 subcolumns in the lower troposphere and increase them in the mid-upper troposphere (Huijnen et al., 2010). Figure 1 shows example tropospheric AKs for summer and winter profiles over London (urban – higher column NO_2) and Dartmoor (rural area in southwest England – lower column NO_2), which have been coloured by their respective tropospheric ~~AMF~~AMFs. In the lower troposphere for both seasons and locations the tropospheric AKs range around 0–1. However, in the mid-upper troposphere, the London tropospheric AKs tend to be greater than Dartmoor in both seasons. London tropospheric AKs are most pronounced in winter, with ~~many~~ some tropospheric AKs over 8, while in the summer they range around 1–8. In both seasons, the tropospheric ~~AMF is~~ AMFs are biggest, 5–6, in the lower range tropospheric AKs, 0–1, and smaller, 0–1.5 as the tropospheric AKs range increases, over 2. If the tropospheric AMFs are small (i.e. near 0 suggesting the majority of the NO_2 is within the lower layers of the London boundary layer; also small tropospheric AKs there), from Eq. (3), as the full atmospheric AKs naturally increase with altitude, the tropospheric AMFs will return larger tropospheric AKs. Also, in winter over London, the shallower boundary layer ~~with~~ will trap larger winter emissions of NO_2 closer to the surface. Therefore, the tropospheric AMF will be smaller and the winter mid-upper tropospheric AKs will be larger as seen in Fig. 1. Over Dartmoor, the AKs show less seasonal variation and the majority range around 1–6 for both summer and winter. This is also seen in the tropospheric AMF, which ranges ~~between~~ around approximately 0–6, but has no clear pattern in the Dartmoor tropospheric AKs, in both seasons.

The Dartmoor AKs tend to be lower than those of London, which could be a result of multiple factors: surface albedo, viewing geometry, cloud cover, etc. As data with cloud cover higher than 20% is filtered out and the viewing geometry of London and Dartmoor will vary depending on where OMI is in its orbit (both locations are at similar latitudes), we suggest that neither is the dominant cause of the AK differences. The surface albedo data in the satellite files is noisy and shows no clear pictures between London and Dartmoor. We suggest that the different NO_2 loading between the locations is the primary factor in the AK differences. Eskes and Boersma (2003) state that AKs are independent of trace gas distributions for optically thin absorbers. However, for stronger absorbers they suggest that the AK depends on the true distribution of the tracer. Therefore, as London column NO_2 is greater than that of Dartmoor, i.e. regions of optically strong and thin absorbers, the AKs are more and less sensitive, respectively.

2.2 Differential optical absorption spectroscopy NO_2 retrieval error

The DOAS retrievals are subject to random, systematic and smoothing errors in the retrieval process. Random (quasi-systematic) errors include fitting errors, cloud errors, instrument noise and signal corruption. Systematic errors include absorption cross-sections, surface albedo and stratospheric correction uncertainties. Finally, smoothing errors include biases in the a priori profiles and sensitivity of the satellite when recording the slant column through the atmosphere. If multiple retrievals are averaged together, as in this study, the random errors will partially cancel leading to the random error being reduced by a factor of $\frac{1}{\sqrt{N}}$ (where N is the number of retrievals).

In contrast, systematic errors are unaffected by cancelling through averaging. In the following section we investigate the different error components of the satellite retrievals and derive an expression for the error in the averaged retrievals. This methodology should give smaller errors which are more representative of the time-averaged retrieval error and so allow a stricter test of the model. Boersma et al. (2004) describe the error in the DOAS $\mathrm{NO_2}$ retrievals as:

$$\sigma_{\mathrm{trop}}^2 = \left(\frac{\sigma_{\mathrm{total}}}{\mathrm{AMF_{trop}}}\right)^2 + \left(\frac{\sigma_{\mathrm{strat}}}{\mathrm{AMF_{trop}}}\right)^2 + \left(\frac{(X_{\mathrm{total}} - X_{\mathrm{strat}})\,\sigma_{\mathrm{AMF_{trop}}}}{\mathrm{AMF_{trop}^2}}\right)^2 \tag{4}$$

where σ_{trop}, σ_{strat} and σ_{total} are the uncertainties in the tropospheric ~~, stratospheric~~ vertical, stratospheric slant and total slant columns, respectively. $\mathrm{AMF_{trop}}$ is the tropospheric air mass factor, $\sigma_{\mathrm{AMF_{trop}}}$ is the error in the tropospheric air mass factor, X_{total} is the total slant column and X_{strat} is the stratospheric slant column.

σ_{total} is made up of both random and systematic error, where the random error component can be reduced by $\frac{1}{\sqrt{N}}$. The sources of systematic error in the total slant column include the $\mathrm{NO_2}$ cross-section, spectral calibration, solar diffuser and temperature (Boersma et al., 2004) . We assume that the systematic and random errors can be combined in quadrature. In Eq. (6) there are two terms for σ_{total}; $\sigma_{\mathrm{total_{ran}}}$ and $\sigma_{\mathrm{total_{sys}}}$, which are the random and systematic error components of the total slant column, respectively. Boersma et al. (2004) state that $\sigma_{\mathrm{total_{sys}}}$ can be expressed as $\sigma_{\mathrm{total_{sys}}} = 0.03X_{\mathrm{total}}$. We treat σ_{strat} here as systematic as both the OMI standard and DOMINO products estimate the stratospheric slant column using TM4 chemistry-transport model simulations and data assimilation ~~(Dirksen et al. , 2011).~~ (Dirksen et al., 2011) . According to the DOMINO OMI product documentation (which references Boersma et al., 2004, 2007; ~~Dirksen et al., 2011~~Dirksen et al. (2011)), the error in the stratospheric slant column is estimated to be 0.25×10^{15} molecules cm^{-2} in all cases. However, according to Belmonte Rivas et al. (2014) , the stratospheric slant column error does not propagate into the tropospheric column error as it is absorbed by the stratospheric assimilation procedure. Therefore, the σ_strat term is removed from Eq. (4).

Boersma et al. (2004) state that the tropospheric column is calculated as:

$$N_\mathrm{trop} = \frac{X_\mathrm{total} - X_\mathrm{strat}}{\mathrm{AMF}_\mathrm{trop}} \tag{5}$$

where N_trop is the vertical tropospheric column and can be substituted, including the σ_total ~~and σ_strat~~ estimates, into Eq. (4). This leads to:

$$\sigma^2_\mathrm{trop} = \left(\frac{\sigma_{\mathrm{total}_\mathrm{ran}}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 + \left(\frac{0.03 X_\mathrm{total}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 + \left(\underline{\frac{0.25 \times 10^{15}}{\mathrm{AMF}_\mathrm{trop}}}\underline{^2+} \frac{N_\mathrm{trop}\sigma_{\mathrm{AMF}_\mathrm{trop}}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 \tag{6}$$

σ_trop is reduced in the model-satellite comparisons when the AK is applied to the model data. Therefore, the error product, $\sigma_{\mathrm{trop}_\mathrm{ak}}$, from the OMI retrieval files with the smoothing error removed is used instead of σ_trop in Eqs. (4) and (6).

Boersma et al. (2007) suggest that the uncertainty in the tropospheric AMF is ~~between~~ around 10–40 %. Therefore, we take the conservative estimate of $\sigma_{\mathrm{AMF}_\mathrm{trop}} = 0.4 \cdot \mathrm{AMF}_\mathrm{trop}$. This leads to the new retrieval error approximation of:

$$\sigma^2_{\mathrm{trop}_\mathrm{ak}} = \left(\frac{\sigma_{\mathrm{total}_\mathrm{ran}}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 + \left(\frac{0.03 X_\mathrm{total}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 + \left(\underline{\frac{0.25 \times 10^{15}}{\mathrm{AMF}_\mathrm{trop}}}\underline{^2+} 0.4 N_\mathrm{trop}\right)^2 \tag{7}$$

All of these terms are known apart from $\sigma_{\mathrm{total}_\mathrm{ran}}$. We can rearrange to calculate this based on other variables provided in the OMI product files. This leads to:

$$\left(\frac{\sigma_{\mathrm{total}_\mathrm{ran}}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 = \sigma^2_{\mathrm{trop}_\mathrm{ak}} - (0.4 N_\mathrm{trop})^2 - \left(\frac{0.03 X_\mathrm{total}}{\mathrm{AMF}_\mathrm{trop}}\right)^2 \underline{-} \underline{\frac{0.25 \times 10^{15}}{\mathrm{AMF}_\mathrm{trop}}}\underline{^2} \tag{8}$$

In the rare case that the ~~left~~ right hand side is negative ~~,~~ (e.g. when N_{trop} is large, but has small uncertainty; X_{total} is scaled by 0.03 so will be small compared to N_{trop}), the random error component cannot be found as it would be complex, so the random error component is then set to 50 % (H. Eskes, personal communication, 2012). Now, rearranging for $\sigma_{\mathrm{total}_{\mathrm{ran}}}$, and assuming the ~~left~~ right hand side is positive, Eq. (8) becomes:

$$\sigma_{\mathrm{total}_{\mathrm{ran}}} = \mathrm{AMF}_{\mathrm{trop}} \sqrt{\left(\sigma^2_{\mathrm{trop}_{\mathrm{ak}}}\right) - (0.4N_{\mathrm{trop}})^2 - \left(\frac{0.03X_{\mathrm{total}}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2 - \left(\frac{0.25\times10^{15}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2}$$

$$\sqrt{\left(\sigma^2_{\mathrm{trop}_{\mathrm{ak}}}\right) - (0.4N_{\mathrm{trop}})^2 - \left(\frac{0.03X_{\mathrm{total}}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2} \qquad (9)$$

This quantity was calculated for each retrieval in each grid square and then the new seasonal retrieval error was calculated taking the reduced random component into account:

$$\overline{\sigma_{\mathrm{trop}_{\mathrm{ak}}}} = \sqrt{\left(\frac{\overline{\sigma_{\mathrm{total}_{\mathrm{ran}}}}}{\sqrt{N}\mathrm{AMF}_{\mathrm{trop}}}\right)^2 + \left(\frac{0.03\overline{X_{\mathrm{total}}}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2 + \left(\frac{0.25\times10^{15}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2 + \left(0.4\overline{N_{\mathrm{trop}}}\right)^2}$$

$$\sqrt{\left(\frac{\overline{\sigma_{\mathrm{total}_{\mathrm{ran}}}}}{\sqrt{N}\mathrm{AMF}_{\mathrm{trop}}}\right)^2 + \left(\frac{0.03\overline{X_{\mathrm{total}}}}{\mathrm{AMF}_{\mathrm{trop}}}\right)^2 + \left(0.4\overline{N_{\mathrm{trop}}}\right)^2}$$

where a bar superscript represents the seasonal time average.

Figure 2 shows how averaging, by decreasing the random error component, reduces the seasonal satellite tropospheric column error as calculated by our algorithm. The figure compares the simple mean of the total satellite column NO_2 error (calculated for each pixel) with our new method which reduces the estimated random error component by one over the square root of the number of observations. The reduction in the satellite column error is then presented as a percentage of the original satellite column seasonal mean error. In both summer and winter, the seasonal mean column error is reduced to 30–90 % across the domain, therefore making the OMI data much more useful for model evaluation. Table 1 gives examples of the seasonal tropospheric column NO_2 error and the reduced tropospheric column NO_2 error using our algorithm for multiple locations across Europe. The error in summer, compared with winter, and the error over sea in comparison to land, are smaller. We suggest that the larger sample size in summer and over the sea, when compared to winter and over the land, respectively, reduces the random error component further as N is larger. Only for a few retrievals over Scandinavia, does this methodology of reducing the random error component increase the overall column error (not shown here).

3 Air quality in the unified model (AQUM)

3.1 Model setup

The AQUM domain covers the UK and part of continental Europe on a rotated grid between approximately 45–60° N and 12° W–12° E. The model has a horizontal resolution of $0.11^\circ \times 0.11^\circ$ with 38 vertical levels between the surface and 39 km. The model has a coupled, online tropospheric chemistry scheme using the ~~UK~~ United Kingdom Chemistry and Aerosols (UKCA) subroutines. The chemistry scheme (Regional Air Quality, RAQ) includes 40 tracers, 23 photolysis reactions and 115 gas-phase reactions (Savage et al., 2013) including the reaction of the nitrate radical with formaldehyde, ethene, ethane, ~~propane~~propene, n-butane, acetaldehyde~~and isoprene~~, isoprene, organic nitrates and the hydroperoxyl radical. The standard model setup does not include any heterogeneous chemistry. A complete chemical mechanism is included in the online supplement to Savage et al. (2013).

The model uses the Coupled Large-scale Aerosol Simulator for Studies In Climate (CLASSIC) aerosol scheme. This is a bulk aerosol scheme with the aerosols treated as an external mixture. It contains six prognostic tropospheric aerosol types: ammonium sulphate, mineral dust, fossil fuel black carbon (FFBC), fossil fuel organic carbon (FFOC), biomass burning aerosols and ammonium nitrate. In addition, there is a diagnostic aerosol scheme for sea salt and a fixed climatology of biogenic secondary organic aerosols (BSOA). For more details of the aerosol scheme see Bellouin et al. (2011). In common with most regional ~~AQ~~ air quality forecast models in Europe, AQUM shows a small negative bias for $PM_{2.5}$ and a larger negative bias for PM_{10}. For full details of the performance of the model for aerosols, NO_2 and ozone see Savage et al. (2013).

Meteorological initial conditions and lateral boundary conditions (LBCs) come from the Met Office's operational global Unified Model (25 km × 25 km) ~~data~~forecast. Initial chemical conditions come from the previous day's AQUM forecast and aerosol and chemistry LBCs come from the ECMWF GEMS (Global and regional Earth-system Monitoring using Satellite and in-situ data) ~~reanalyses~~ reanalysis (Hollingsworth et al., 2008). The GEMS fields, available at http://www.gmes-atmosphere.eu/, provide boundary fluxes for regional air quality models such as AQUM.

This configuration of AQUM uses emission datasets from the National Atmospheric Emissions Inventory (NAEI) (1 km × 1 km) for the UK, ENTEC (5 km × 5 km) for the shipping lanes and European Monitoring and Evaluation Programme (EMEP) (50 km × 50 km) for the rest of the model domain. Over the UK the NAEI NO_x emissions datasets are made up of two source types: area and point. Area sources include traffic, light industry and urban emissions, while point sources are power stations, landfill, incinerators and refineries. Typically, the point source emissions are $100\,\mathrm{g\,s^{-1}}$ in magnitude, while the area sources tend to be $10\,\mathrm{g\,s^{-1}}$. For most of the experiments we use 2007 instead of 2006 NO_x sources because the ENTEC shipping emissions (5 km × 5 km resolution) are available for this year, while only the coarse EMEP shipping emissions are available for the earlier years (Savage et al., 2013). The difference between 2006 and 2007 point source emissions are negligible in altering the AQUM column NO_2 (not shown). Therefore, we use the 2007 emissions datasets throughout this study. The fractional seasonal cycle, which comes from Visschedijk et al. (2007), applied to AQUM's annual NO_x emissions can be seen in Figure 3.

The lightning emissions are based on a parameterisation linked to the model's convection scheme. For details see O'Connor et al. (2014). We do not have a separate parameterisation for soil NO_x emissions but given the large emissions from transport and industry, the soil NO_x emissions are unlikely to be important in this region.

Poupkou et al. (2010) provide the monthly climatology of biogenic emissions on a 0.125° × 0.0625° resolution. The use of climatological biogenic isoprene emissions will partially diminish AQUM's representation of ozone from biogenic precursors. A new interactive biogenic isoprene scheme is under development but was not available for this study. However, this is a secondary issue in this paper as we focus on primary emissions of NO_x. Biomass burning emissions of aerosols come from the Global Fire Emissions Database (GFED), version 1, (Randerson et al., 2005) for 2000. The use of biomass burning emissions from 2000 is somewhat arbitrary, but within the AQUM's domain these emissions have relatively little impact.

3.2 Sensitivity experiments

We performed one control and five sensitivity experiments to investigate the AQUM's simulation of column NO_2. Two experiments used different LBCs, two experiments used modified point source emissions and two included heterogeneous chemistry. These are summarised in Table ~~1.~~ 2.

Simulation MACC investigates the sensitivity of AQUM column NO_2 to different chemical LBCs from the global Monitoring Atmospheric Composition and Climate (MACC) ~~reanalyses~~reanalysis, which is the follow-on project of GEMS (Inness et al., 2013). The MACC reanalysis uses a more recent version of the ECMWF model (Integrated Forecast System), assimilates more satellite products and was run at a resolution of 80 km instead of 125 km.Savage et al. (2013) have undertaken a similar analysis of the MACC LBCs in AQUM. They showed that when compared with the AURN observations of O_3, AQUM-MACC performs well during the first quarter of 2006 and overestimates observations afterwards, while AQUM-GEMS has a negative bias during the first quarter of the year but compares well with observations afterwards.

We have performed additional runs to examine the impact of the point sources over the UK on NO_2 columns. The motivation behind Run E1 was to determine the impact of the NO_x point sources on the simulated column NO_2 budget, as we hypothesised that the AQUM's representation of them was the cause of some of the AQUM–OMI column NO_2 positive biases (see Sect. 4.1). Run E2 ~~uses an~~ introduces a new idealised passive tracer ~~from the~~ emitted from the UK point sources with ~~a~~ the same emissions to that of the model NO_x inventory. The idealised tracer is transported like any chemical tracer, but is not lost through chemical reactions. Instead it is lost through its e-folding lifetime of one day~~to examine if the tracer columns correlated~~. The point source tracer columns can then be examined to see if they correlate with summer AQUM-OMI positive biases (see Sect. 4.3).

Runs N_2O_5High and N_2O_5Low investigate the impact of heterogeneous chemistry on NO_2 columns. Tropospheric NO_x ($NO + NO_2$) sources are dominated by anthropogenic emissions and the loss of NO_2 to HNO_3 is through two pathways:

$$\mathrm{NO_2} + \mathrm{OH} + M \rightarrow \mathrm{HNO_3} + M \tag{R1}$$

$$\mathrm{NO_2} + \mathrm{O_3} \rightarrow \mathrm{NO_3} + \mathrm{O_2} \tag{R2}$$

$$\mathrm{NO_3} + \mathrm{NO_2} + M \rightleftharpoons \mathrm{N_2O_5} + M \tag{R3}$$

$$\mathrm{N_2O_5} + \mathrm{H_2O} \xrightarrow{\text{aerosol}} 2\mathrm{HNO_{3(aq)}} \tag{R4}$$

The standard configuration of AQUM does not include any heterogeneous reactions such as the hydrolysis of N_2O_5 on aerosol surfaces (see details of the chemistry scheme in the Supplement of Savage et al., 2013). Previous global modelling studies have shown that this process can be a significant NO_x sink at mid-latitudes in winter (e.g. Tie et al., 2003; Macintyre and Evans, 2010). Following those analyses, we have implemented this reaction, with rate k (s^{-1}) calculated as:

$$k = \frac{A\gamma\omega}{4} \tag{10}$$

where A is the aerosol surface area ($cm^2\,cm^{-3}$), γ is the uptake coefficient of N_2O_5 on aerosols (non-dimensional) and $\omega = 100\,[8RT/(\pi M)]^{\frac{1}{2}}$ ($cm\,s^{-1}$) is the root-mean-square molecular speed of N_2O_5 at temperature T (K), M is the molecular mass of N_2O_5 ($kg\,mol^{-1}$) and $R = 8.3145\,J\,mol^{-1}\,K^{-1}$.

Macintyre and Evans (2010) investigated the sensitivity of N_2O_5 loss on aerosol by using a range of uptake values (0.0, 10^{-6}, 10^{-4}, 10^{-3}, 5×10^{-3}, 10^{-2}, 2×10^{-2}, 0.1, 0.2, 0.5 and

1.0). They found that limited sensitivity occurs at low and high values of γ. At low values, the uptake pathway is an insignificant route for NO_x loss. At high values, the loss of NO_x through heterogeneous removal of N_2O_5 is limited by the rate of production of NO_3, rather than the rate of heterogeneous uptake. However, in the northern extra-tropics (including the AQUM domain), their model shows significant sensitivity to intermediate values of γ (0.001–0.02) ~~show~~ with a significant loss of NO_x. Therefore, we experiment with $\gamma = 0.001$ and 0.02 to investigate the sensitivity of AQUM column NO_2 to heterogeneous chemistry. The aerosol surface area, A, includes the contribution of seven aerosol types present in CLASSIC: sea salt aerosol, ammonium nitrate, ammonium sulphate, biomass burning aerosol, black carbon, ~~fossil fuel organic carbon (FFOC) and biogenic secondary organic aerosol (BSOA)~~FFOC and BSOA. To account for hydroscopic growth of the aerosols, the formulation of Fitzgerald (1975) is used for growth above the deliquescence point for ammonium sulphate (RH = 81 %), sea salt (RH = 75 %) and ammonium nitrate (RH = 61 %) up to 99.5 % RH. We apply a linear fit between the efflorescence (RH = 30 % for sulphate, 42 % for sea-salt and 30 % for nitrate) and deliquescence points. There is no hydroscopic growth below the efflorescence point. Look-up tables are used for the other aerosol types. Biomass burning and FFOC aerosol growth rates are taken from Magi and Hobbs (2003), BSOA growth rates come from Varutbangkul et al. (2006) and black carbon is considered to be hydrophobic (no growth).

3.3 Statistical comparisons

For the AQUM-satellite comparisons the following model-observation statistics were used: Mean Bias (MB), Root Mean Square Error (RMSE) and the Fractional Gross Error (FGE, bounded by the values 0 to 2). These statistics are described by Han et al. (2011) and Savage et al. (2013). Further details are given in the Appendix.

4 Results

4.1 Control run

Figure ~~3~~ 4 compares observed column NO_2 with the AQUM control Run C (with AKs applied). The AQUM and OMI averages have similar spatial patterns, with maximum and minimum column NO_2 over the urban and rural/ocean regions, respectively. In summer, AQUM and OMI background concentrations are around ~~0–3~~$O(10^{13})$–3×10^{15} molecules cm^{-2}~~.~~, where $O(10^{13})$ represents values in size of the order of 10^{13}. The OMI peak column NO_2 of 16–20 $\times 10^{15}$ molecules cm^{-2} is over London. AQUM simulates similar London column NO_2, but the model peak concentrations are over northern England at over 20×10^{15} molecules cm^{-2}.

In winter, the background column NO_2 is elevated with a larger spatial extent ranging ~~between 0–6~~around $O(10^{13})$–6×10^{15} molecules cm^{-2} in both the AQUM and OMI fields. However, the elevated AQUM background state has a larger coverage than that of OMI. Over the source regions, OMI column NO_2 peaks over London at 12–13 $\times 10^{15}$ molecules cm^{-2}, with similar concentrations seen in AQUM. However, AQUM peak column NO_2 are over northern England at 12–16 $\times 10^{15}$ molecules cm^{-2}. Therefore, ~~independent~~ independently of season, AQUM overestimates northern England column NO_2. Interestingly, the background column NO_2 is larger in winter for both AQUM and OMI, but column NO_2 is lower over the source regions in winter than in summer (Pope et al., 2014). van der A et al. (2008) suggest that peak UK NO_x emissions occur in July, while Pope et al. (2014) suggest that the transport of column NO_2 away from source regions due to stronger winter dynamics outweighs the loss of UK source region column NO_2 from enhanced summer photochemistry.

Figure ~~4~~ 5 shows the MB between AQUM Run C and OMI. The black polygoned regions show significant differences, i.e. where the magnitude of the MB is greater than the satellite error. In summer, there are significant positive, 5–10 $\times 10^{15}$ molecules cm^{-2}, and negative, -10 to -1×10^{15} molecules cm^{-2}, biases in northern England and the Benelux region, respectively. The negative biases are potentially linked to the coarser resolution

EMEP NO_x emissions datasets (50 km × 50 km) which average emissions over a larger grid square causing AQUM to simulate lower column NO_2 than seen by OMI. We hypothesise that the northern England biases are linked to the point source (power station) NO_x emissions from NAEI. This is further discussed in Sect. 4.3. In winter, AQUM overestimates OMI by $1–3 \times 10^{15}$ molecules cm^{-2} over the North Sea and Scotland, as the modelled winter background column NO_2 is larger; this is further investigated in Sect. 4.4 by including an additional NO_x sink in the chemistry scheme of the model. The northern England positive biases seen in summer also extend to winter, $3–5 \times 10^{15}$ molecules cm^{-2}, suggesting ~~the northern England biases are annual instead of seasonal~~ that this is not only a seasonal feature. Finally, the large bias dipole in the Po Valley appears to be related to the LBCs or the winter emissions, as summer biases are small.

We also compared AQUM against surface observations of NO_2 from AURN, found at http://uk-air.defra.gov.uk/networks/network-info?view=aurn, and maintained by DEFRA. This was to see if there was a consistent pattern in the biases in the model column and surface NO_2. However, we find similar problems to Savage et al. (2013) where surface AQUM – observation comparisons ~~have~~ show systematic negative biases at urban sites. The coarse model resolution, compared to the observation point measurements (even with roadside and traffic sites removed), results in significant model underestimation of NO_2 ~~at all~~ in urban regions. Therefore, it is difficult to draw any conclusions on the AQUM skill as the model grid-point data will struggle to reproduce the point measurement observations. Also the spatial coverage of the AURN data is very sparse over the UK and AURN NO_2 measurement interferences from molybdenum converters (Steinbacher et al., 2007) overestimate surface concentrations~~in~~, in particular at rural sites. Therefore, satellite (pixel area) data are the primary observations used to evaluate AQUM in this paper.

4.2 Impact of lateral boundary conditions

Figure ~~5a~~ 6a and b shows results of the sensitivity run with the MACC boundary conditions (Run MACC) and can be compared with Fig. ~~3a~~ 4a and b. The MACC LBCs have a limited impact on summer column NO_2 with peak concentrations over London and Northern England between 15–20 $\times 10^{15}$ molecules cm^{-2} for both runs MACC and C. However, in winter Run MACC increases column NO_2 from approximately 12 $\times 10^{15}$ to 16 $\times 10^{15}$ molecules cm^{-2} over the UK and Benelux region. When compared with OMI (Fig. ~~5a and b~~6c and d) the limited summer impact of the MACC LBCs results in biases which are similar to those in Fig. ~~4~~ 5 from the control run, with biases over northern England, 5–10 $\times 10^{15}$ molecules cm^{-2}, and continental Europe, -5 to -3×10^{15} molecules cm^{-2}. In winter, Run MACC has enhanced column NO_2 resulting in biases with OMI of between 2–5 $\times 10^{15}$ molecules cm^{-2} across the whole domain, unlike Run C with GEMS LBCs in Fig. ~~4.~~ 5. The peak positive biases are again over northern England (and the Po Valley), 5 $\times 10^{15}$ molecules cm^{-2}, suggesting that AQUM overestimates NO_2 in the region, at the OMI overpass time, ~~independent~~ independently of season or LBCs. Therefore, the GEMs LBCs appear to give better AQUM column NO_2 forecast skill than MACC does, ~~which is consistent with the findings of~~ similarly as found by Savage et al. (2013) for the comparisons with surface ozone.

4.3 AQUM NO_x emissions sensitivity experiments

We hypothesise that significant summer Run C–OMI positive biases in northern England and Scotland (Fig. ~~4~~5) are caused by the AQUM's representation of point source (mainly power station) NO_x emissions. Therefore, to better understand these biases, we investigate sensitivity experiments of NO_x emissions (Table ~~1~~2) for June-July-August (JJA) 2006 (Fig. ~~6a~~ 7a shows JJA Run C–OMI positive biases). Figure ~~6b–d~~ 7b–d shows the JJA AQUM NO_x emissions for runs C and E1 (with point sources removed) and their difference. The peak Run C NO_x emissions are around 1.8×10^{-9} kg $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$. However, with point sources removed, the differences are 1.8×10^{-9} kg $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ in point source locations, showing that they make up a significant part of the emissions budget.

Figure ~~7a~~ 8a and b highlights the impact of removing point sources as column NO_2 over northern England reduces from 15–25 $\times 10^{15}$ molecules cm^{-2} to 4–5 $\times 10^{15}$ molecules cm^{-2}. The Run E1–OMI MB now ranges between -10 ~~to~~ and -6×10^{15} molecules cm^{-2}, while the Run C–OMI MB (Fig. ~~6a) is between~~ 7a) is around

6–10 × 10^{15} molecules cm^{-2}. Therefore, the switch in sign of the biases, of similar magnitude, indicates that the point source emissions play a significant role in the AQUM column NO_2 budget.

Run E2 aimed to test whether the point sources were responsible for the positive biases in Fig. ~~6a~~ 7a by using an idealised tracer of the power station emissions. Figure ~~7c~~ 8c shows the JJA tracer column with the OMI AKs applied, where peak columns range around 16–20 × 10^{15} molecules cm^{-2} over northern England. The minimum tracer values of 0 × 10^{15} molecules cm^{-2} are over the sea and continental Europe as there is no emission of the tracer there. Inspection of Figs. ~~6a and 7c~~ 7a and 8c suggest that the peak tracer columns overlap with the large Run C–OMI positive biases.

To test this more quantitatively, the spatial correlation between these peak concentrations from Run E2 were compared against a random tracer-MB (Run C) correlation distribution. The largest 100 tracer column pixels in Fig. ~~7c~~ 8c were compared against the MBs over the same locations in Fig. ~~6a~~7a, yielding a correlation of 0.45. Then, using a Monte-Carlo approach, a random 100 sample of the Fig. ~~6a~~ 7a land-based MB pixels (we use land bias pixels only as the biases in Fig. ~~6a~~ 7a are over land) were correlated against the largest 100 tracer sample. This was repeated 1000 times and then sorted from lowest to highest. The 5th and 95th percentiles were calculated at −0.162 and 0.158, respectively. Our theory is that if the point sources are responsible for the peak Run C–OMI biases, then the peak tracer concentrations, which represent the point source emissions, should be in the same location as the peak biases. By looking at the random samples correlation, we see how the tracer-MB peak value concentration compares with randomly sampled MB locations. Since 0.45 is above the 95th percentile, this shows the tracer-MB peak correlation value is significant (is actually the greatest correlation – see Fig. ~~7d~~8d) and that AQUM's representation of point source emissions ~~are~~ is linked to the AQUM overestimation of column NO_2 in northern England and Scotland.

4.4 Sensitivity to heterogeneous removal of N_2O_5

Figure ~~8~~ 9 shows the winter and summer MBs between AQUM (with LBCs from GEMS) and OMI when heterogeneous hydrolysis of N_2O_5 is implemented in the model with $\gamma = 0.001$ (Run N_2O_5Low) and $\gamma = 0.02$ (Run N_2O_5High). In the Run C summer case (see Fig. ~~4a~~5a) there are positive northern England and Scotland biases of around 5–10 $\times 10^{15}$ molecules cm^{-2}. We have shown that these positive biases are likely linked to AQUM's representation of point source emissions. However, by introducing N_2O_5 heterogeneous chemistry these positive biases are significantly reduced. In Run N_2O_5Low (Fig. ~~8a~~9a) there is some impact on the biases as RMSE (over UK domain 8° W–2° E and 50–60° N) decreases from 3.68×10^{15} to 3.39×10^{15} molecules cm^{-2} and FGE (over UK domain 8° W–2° E and 50–60° N) also reduces very slightly. In Run N_2O_5High (Fig. ~~8c~~9c) many of the positive biases over point sources are now insignificant and the RMSE decreases to 3.08×10^{15} molecules cm^{-2}. However, over parts of continental Europe the intensity and spread of negative biases has increased, thus suggesting that $\gamma = 0.02$ might be too strong an uptake here. The FGE does go up slightly to 0.67 and we suspect that this is due to the introduction of negative biases over relatively clean or moderately polluted areas (e.g. the Irish Sea and parts of the continent). Note that the correction of errors of large magnitude (e.g. over point sources) reduces RMSE because this metric penalises the large deviations between the model and the satellite-retrieved columns, while the introduction of errors of low magnitude over less polluted areas might increase the normalised errors given by FGE. The changes at the point source locations are most significant because of the large emissions of NO_x and aerosols suitable for this heterogeneous process to take place. Therefore, we suggest that while AQUM's representation of point sources ~~results in~~ may be responsible for the summer northern England/Scotland positive biases, including N_2O_5 heterogeneous chemistry with $\gamma = 0.02$ will partially account for this. In winter, the positive biases seen in Fig. ~~4b~~5b, 2–5 $\times 10^{15}$ molecules cm^{-2}, decrease as γ increases, similarly as found for summer. In Run N_2O_5Low (Fig. ~~8b~~9b) the spatial spread of significantly positive biases is only partially reduced, resulting in small decreases of RMSE (from 5.12×10^{15} to

5.05×10^{15} molecules cm^{-2}) and FGE (from 0.63 to 0.62). For Run N_2O_5High (Fig. ~~8d~~9d) the cluster of significantly positive biases has decreased spatially yielding the best comparisons, with RMSE and FGE values of 4.48×10^{15} molecules cm^{-2} and 0.60, respectively.

5 Conclusions

We have successfully used OMI satellite observations of column NO_2 over the UK to further explore the AQUM performance, extending on previous validation of the model which had only used surface data. In order to do this we have looked in detail at the satellite errors (random, systematic and smoothing) and derived an algorithm which reduces the retrieval random error component when averaging retrievals. This allows more critical AQUM-satellite comparisons as the time average random error component can be reduced by ~~30–90~~30–70 % in all seasons.

Based on the summer and winter comparisons, the standard (operational) AQUM overestimates column NO_2 over northern England/Scotland by $5–10 \times 10^{15}$ molecules cm^{-2} and over the northern domain by $2–5 \times 10^{15}$ molecules cm^{-2}. The use of a different set of lateral boundary conditions (from the MACC reanalysis), which are known to increase AQUM's surface ozone positive bias (Savage et al., 2013), also increases the error in the NO_2 columns. The AQUM column NO_2 is increased, especially in winter, by $2–5 \times 10^{15}$ molecules cm^{-2}, resulting in poorer comparisons with OMI.

From multiple sensitivity experiments on the UK NO_x point source emissions we conclude that it was AQUM's representation of these emissions which very likely caused the northern England/Scotland summer biases. By emitting an idealised tracer in the NO_x points sources we found a significant correlation of the peak tracer columns to the AQUM – OMI MBs. Finally, introducing N_2O_5 heterogeneous chemistry in AQUM improves the AQUM–OMI comparisons in both seasons. In winter, the spatial extent of positive biases, $2–5 \times 10^{15}$molecules cm^{-2}, decreases. In summer, the northern England biases decrease both spatially and in magnitude from 5–10 to $0–5 \times 10^{15}$ molecules cm^{-2}. Therefore, this suggests that in summer the AQUM's representation of NO_x point sources is inaccurate but can be partially masked by the introduction of N_2O_5 heterogeneous chemistry.

As this study has shown the potential use of satellite observations, along with the time-averaged random error algorithm, to evaluate AQUM, the data could be used in future to evaluate ~~operation~~ operational air quality forecasts. We also show that the heterogeneous loss of N_2O_5 on aerosol is an important sink of NO_2 and should be included in the operational AQUM.

Appendix A

The equations for mean bias (MB), root mean square error (RMSE), modified normalised mean bias (MNMB) and the fractional gross error (FGE) are given here, where f is the model output, o is the satellite measurements, N is the total number of elements and i is the index.

Mean Bias (MB):

$$\text{MB} = \frac{1}{N}\sum_i (f_i - o_i) \tag{A1}$$

Modified Normalised Mean Bias (MNMB):

$$\text{MNMB} = \frac{2}{N}\sum \frac{(f_i - o_i)}{f_i + o_i} \tag{A2}$$

Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_i (f_i - o_i)^2} \tag{A3}$$

Fractional Gross Error (FGE):

$$\text{FGE} = \frac{2}{N}\sum_i \left|\frac{f_i - o_i}{f_i + o_i}\right| \tag{A4}$$

Acknowledgements. We acknowledge the use of the Tropospheric Emissions Monitoring Internet Service (TEMIS) OMI dataset used in this study. This work was supported by the UK Natural Environment Research Council (NERC) and National Centre for Earth Observation (NCEO). Richard Pope thanks the NCEO for a Ph.D. studentship to undertake this work. We are grateful to Ben Johnson ~~of~~ at the Met Office who provided advice for the implementation of aerosol growth and surface area diagnostics in CLASSIC.

References

Bellouin, N., Rae, J., Jones, A., Johnson, C., Haywood, J., and Boucher, O.: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate, J. Geophys. Res.-Atmos., 116, D20206, doi:10.1029/2011JD016074, 2011.

Belmonte Rivas, M., Veefkind, P., Boersma, F., Levelt, P., Eskes, H., and Gille, J.: Intercomparison of daytime stratospheric NO_2 satellite retrievals and model simulations, Atmospheric Measurement Techniques Discussions, 7, 895–948, doi:10.5194/amtd-7-895-2014, http://www.atmos-meas-tech-discuss.net/7/895/2014/, 2014.

Blond, N., Boersma, K. F., Eskes, H. J., van der A, R. J., Van Roozendael, M., De Smedt, I., Bergametti, G., and Vautard, R.: Intercomparison of SCIAMACHY nitrogen dioxide observations, in situ measurements and air quality modeling results over Western Europe, J. Geophys. Res.-Atmos., 112, D10311, doi:10.1029/2006JD007277, 2007.

Boersma, K., Eskes, H., and Brinksma, E.: Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res.-Atmos., 109, D04311, doi:10.1029/2003JD003962, 2004.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmos. Chem. Phys., 7, 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

Boersma, K., Jacob, D., Bucsela, E., Perring, A., Dirksen, R., van der A, R., Yantosca, R., Park, R., Wenig, M., Bertram, T., and Cohen, R.: Validation of {OMI} tropospheric {NO2} observations during INTEX-B and application to constrain emissions over the eastern United States and Mexico, Atmos. Environ., 42, 4480–4497, doi:10.1016/j.atmosenv.2008.02.004, 2008.

Boersma, K. F., Jacob, D. J., Trainic, M., Rudich, Y., DeSmedt, I., Dirksen, R., and Eskes, H. J.: Validation of urban NO_2 concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in situ surface measurements in Israeli cities, Atmos. Chem. Phys., 9, 3867–3879, doi:10.5194/acp-9-3867-2009, 2009.

Boersma, K., Braak, R., and van der A, R.: Dutch OMI NO_2 (DOMINO) data product v2.0, Tropospheric Emissions Monitoring Internet Service on-line Documentation, available at: http://www.temis.nl/docs/OMI_NO2_HE5_2.0_2011.pdf (last access: June 2014), KNMI, the Netherlands, 2011a.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905–1928, doi:10.5194/amt-4-1905-2011, 2011b.

Braak, R.: Row Anomaly Flagging Rules Lookup Table, KNMI Technical Document TN-OMIE-KNMI-950, KNMI, the Netherlands, 2010.

Curier, R., Kranenburg, R., Segers, A., Timmermans, R., and Schaap, M.: Synergistic use of {OMI} {NO_2} tropospheric columns and ~~LOTOS?EUROS~~ LOTOS-EUROS to evaluate the {NO_x} emission trends across Europe, Remote Sens. Environ., 149, 58–69, doi:10.1016/j.rse.2014.03.032, 2014.

DEFRA: Automatic Urban and Rural Network (AURN), available at: http://uk-air.defra.gov.uk/networks/network-info?view=aurn (last access: 2014), DEFRA, UK, 2012.

Dirksen, R. J., Boersma, K. F., Eskes, H. J., Ionov, D. V., Bucsela, E. J., Levelt, P. F., and Kelder, H. M.: Evaluation of stratospheric NO2 retrieved from the Ozone Monitoring Instrument: Intercomparison, diurnal cycle, and trending, Journal of Geophysical Research: Atmospheres, 116, n/a–n/a, doi:10.1029/2010JD014943, http://dx.doi.org/10.1029/2010JD014943, 2011.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.

Fitzgerald, J. W.: Approximation formulas for the equilibrium size of an aerosol particle as a function of its dry size and composition and the ambient relative humidity, J. Appl. Meteorol., 14, 1044–1049, 1975.

Foley, K. M., Roselle, S. J., Appel, K. W., Bhave, P. V., Pleim, J. E., Otte, T. L., Mathur, R., Sarwar, G., Young, J. O., Gilliam, R. C., Nolte, C. G., Kelly, J. T., Gilliland, A. B., and Bash, J. O.: Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7, Geosci. Model Dev., 3, 205–226, doi:10.5194/gmd-3-205-2010, 2010.

Han, K., Lee, C., Lee, J., Kim, J., and Song, C.: A comparison study between model-predicted and OMI-retrieved tropospheric NO_2 columns over the Korean peninsula, Atmos. Environ., 45, 2962–2971, 2011.

HoC: House of Commons Environmental Audit Report (HCEA): Air Quality: Vol 1 (2009–2010), ~~availabale~~ available at: http://www.publications.parliament.uk/pa/cm200910/cmselect/cmenvaud/229/229i.pdf (last access: February 2014), House of Commons, London, UK, 2010.

Hollaway, M. J., Arnold, S. R., Challinor, A. J., and Emberson, L. D.: Intercontinental trans-boundary contributions to ozone-induced crop yield losses in the Northern Hemisphere, Biogeosciences, 9, 271–292, doi:10.5194/bg-9-271-2012, 2012.

Hollingsworth, A., Engelen, R., Benedetti, A., Dethof, A., Flemming, J., Kaiser, J., Morcrette, J., Simmons, A., Textor, C., Boucher, O., Chevallier, F., Rayner, P., Elbern, H., Eskes, H., Granier, C., Peuch, V.-H., Rouil, L., and Schultz, M. G.: Toward a monitoring and forecasting system for atmospheric composition: the GEMS project, B. Am. Meteorol. Soc., 89, 1147–1164, 2008.

Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K. F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C.: Comparison of OMI NO_2 tropospheric columns with an ensemble of global and European regional air quality models, Atmos. Chem. Phys., 10, 3273–3296, doi:10.5194/acp-10-3273-2010, 2010.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmos. Chem. Phys., 13, 4073–4109, doi:10.5194/acp-13-4073-2013, 2013.

~~Irie,~~

Irie, H., ~~Kanaya,~~ Boersma, K., Kanaya, Y., ~~Akimoto,~~ Takashima, H., ~~Tanimoto, H., Wang, Z., Gleason, J. F., and Bucsela, E. J.: Validation of OMI tropospheric column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment 2006, Atmos. Chem. Phys., 8, 6577–6586, , 2008.~~ Pan, X., and Wang, Z.: Quantitative bias estimates for tropospheric NO2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia, Atmospheric Measurement Techniques, 5, 2403–2411, 2012.

Lauer, A., Dameris, M., Richter, A., and Burrows, J. P.: Tropospheric NO_2 columns: a comparison between model and retrieved data from GOME measurements, Atmos. Chem. Phys., 2, 67–78, doi:10.5194/acp-2-67-2002, 2002.

Macintyre, H. L. and Evans, M. J.: Sensitivity of a global model to the uptake of N_2O_5 by tropospheric aerosol, Atmos. Chem. Phys., 10, 7409–7414, doi:10.5194/acp-10-7409-2010, 2010.

Magi, B. I. and Hobbs, P. V.: Effects of humidity on aerosols in southern Africa during the biomass burning season, J. Geophys. Res.-Atmos., 108, 8495, doi:10.1029/2002JD002144, 2003.

Monks, S., Arnold, S., and Chipperfield, M.: Evidence for El Niño–Southern Oscillation (ENSO) influence on Arctic CO interannual variability through biomass burning emissions, Geophys. Res. Lett., 39, L14804, doi:10.1029/2012GL052512, 2012.

O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., Folberth, G. A., Sanderson, M. G., Telford, P. J., Voulgarakis, A., Young, P. J., Zeng, G., Collins, W. J., and Pyle, J. A.: Evaluation of the new UKCA climate-composition model – Part 2: The Troposphere, Geosci. Model Dev., 7, 41–91, doi:10.5194/gmd-7-41-2014, 2014.

Pope, R., Savage, N., Chipperfield, M., Arnold, S., and Osborn, T.: The influence of synoptic weather regimes on UK air quality: analysis of satellite column NO_2, Atmos. Sci. Lett., 15, 211–217, doi:10.1002/asl2.492, 2014.

Poupkou, A., Giannaros, T., Markakis, K., Kioutsioukis, I., Curci, G., Melas, D., and Zerefos, C.: A model for European Biogenic Volatile Organic Compound emissions: Software development and first validation, Environmental Modelling & Software, 25, 1845 – 1856, doi:http://dx.doi.org/10.1016/j.envsoft.2010.05.004, http://www.sciencedirect.com/science/article/pii/S1364815210001441, 2010.

Randerson, J., Kasibhatla, P., Kasischke, E., Hyer, E., Giglio, L., Collatz, G., and van der Werf, G.: Global fire emissions database (GFED), version 1, http://daac.ornl.gov (last access: 2 October 2012), 2005.

Savage, N. H., Pyle, J. A., Braesicke, P., Wittrock, F., Richter, A., Nüß, H., Burrows, J. P., Schultz, M. G., Pulles, T., and van Het Bolscher, M.: The sensitivity of Western European NO_2 columns to interannual variability of meteorology and emissions: a model – GOME study, Atmos. Sci. Lett., 9, 182–188, 2008.

Savage, N. H., Agnew, P., Davis, L. S., Ordóñez, 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, doi:10.5194/gmd-6-353-2013, 2013.

Seigneur, C., Pun, B., Pai, P., Louis, J.-F., Solomon, P., Emery, C., Morris, R., Zahniser, M., Worsnop, D., Koutrakis, P., White, W., and Tombach, I.: Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility, J. Air Waste Manage., 50, 588–599, 2000.

Steinbacher, M., Zellweger, C., Schwarzenbach, B., Bugmann, S., Buchmann, B., Ordonez, C., Prevot, A., and Hueglin, C.: Nitrogen oxide measurements at rural sites in Switzerland: bias of conventional measurement techniques, J. Geophys. Res.-Atmos., 112, D11307, doi:10.1029/2006JD007971, 2007.

Tie, X., Emmons, L., Horowitz, L., Brasseur, G., Ridley, B., Atlas, E., Stround, C., Hess, P., Klonecki, A., Madronich, S., Talbot, R., and Dibb, J.: Effect of sulfate aerosol on tropospheric NO_x and ozone budgets: model simulations and TOPSE evidence, J. Geophys. Res.-Atmos., 108, 2156–2202, doi:10.1029/2001JD001508, 2003.

van Noije, T. P. C., Eskes, H. J., Dentener, F. J., Stevenson, D. S., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Boersma, K. F., Butler, T., Cofala, J., Drevet, J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Krol, M. C., Lamarque, J.-F., Lawrence, M. G., Martin, R. V., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Richter, A., Rodriguez, J. M., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., and van Roozendael, M.: Multi-model ensemble simulations of tropospheric NO_2 compared with GOME retrievals for the year 2000, Atmos. Chem. Phys., 6, 2943–2979, doi:10.5194/acp-6-2943-2006, 2006.

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., Van Roozendael, M., De Smedt, I., Peters, D. H. M. U., and Meijer, E. W.: Trends, seasonal variability and dominant NOx source derived from a ten year record of NO2 measured from space, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD009021, http://dx.doi.org/10.1029/2007JD009021, 2008.

Varutbangkul, V., Brechtel, F. J., Bahreini, R., Ng, N. L., Keywood, M. D., Kroll, J. H., Flagan, R. C., Seinfeld, J. H., Lee, A., and Goldstein, A. H.: Hygroscopicity of secondary organic aerosols formed by oxidation of cycloalkenes, monoterpenes, sesquiterpenes, and related compounds, Atmos. Chem. Phys., 6, 2367–2388, doi:10.5194/acp-6-2367-2006, 2006.

Velders, G. J., Granier, C., Portmann, R. W., Pfeilsticker, K., Wenig, M., Wagner, T., Platt, U., Richter, A., and Burrows, J. P.: Global tropospheric NO_2 column distributions: comparing three-dimensional model calculations with GOME measurements, J. Geophys. Res.-Atmos., 106, 12643–12660, 2001.

Visschedijk, A., Zanveld, P., and van der Gon, H.: A high resolution gridded European emission database for the EU integrated project GEMS, TNO report 2007-A-R0233/B, 2007.

WHO: Air Quality and Health, available at: http://www.who.int/mediacentre/factsheets/fs313/en/ (last access: February 2014), World Health Organisation, Geneva, Switzerland, 2011.

Table 1. The average column NO_2, column NO_2 error and column NO_2 error as calculated in Sect. 2.2 for multiple locations across Europe in summer and winter (×10^{15} molecules/cm^2).

Place	**Column NO_2**		**Column NO_2 Error**		**Error (Sect. 2.2)**	
	Summer	**Winter**	**Summer**	**Winter**	**Summer**	**Winter**
London 1°W-1°E, 51-51.5°N	9.86	10.7	9.68	9.13	4.24	4.52
Benelux 3-7°E, 50.5-52.5°N	9.57	11.4	7.09	9.24	3.98	4.82
Po Valley 7-9°E, 44.25-45.5°N	3.35	11.9	2.44	9.88	1.42	4.66
Northern England 3-0°W, 52.5-54°N	8.11	8.06	7.13	6.56	3.47	3.42
North Sea 0-8°E, 54-60°N	1.48	2.22	1.94	2.12	0.86	1.01
Scandinavia 6-16°E, 54-63°N	1.48	2.10	1.49	2.12	0.74	1.16

Table 2. List of AQUM runs and experiments.

Run ID	Run Description
C	Control run (GEMS LBCs)
MACC	MACC LBCs
E1	No point sources emissions
E2	Idealised point source tracer
N_2O_5Low	With N_2O_5 heterogeneous chemistry with $\gamma = 0.001$
N_2O_5High	As run N_2O_5Low but with $\gamma = 0.02$

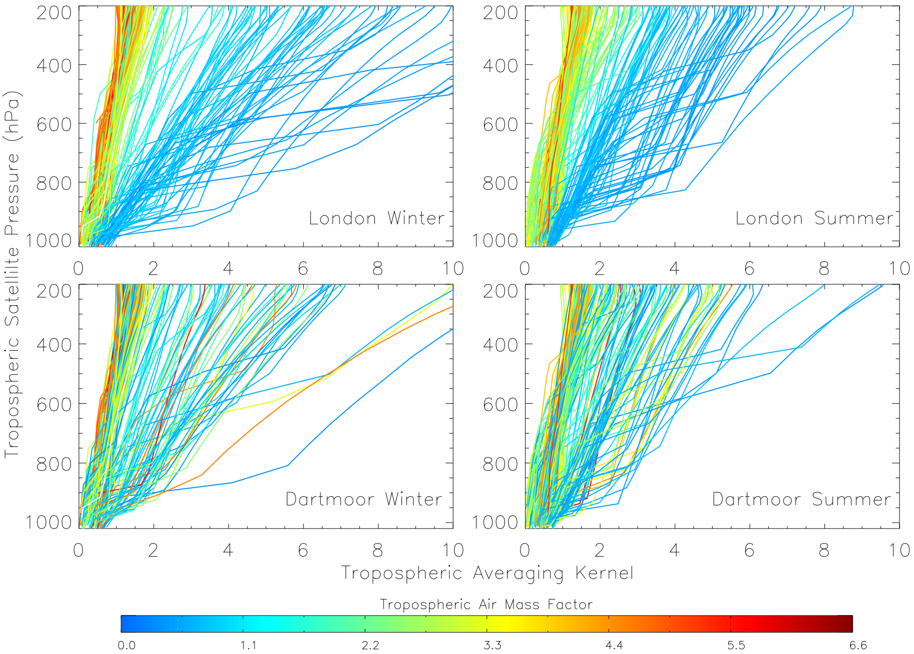


Figure 1. Example OMI averaging kernels for London (top) and Dartmoor (bottom) for summer (right) and winter (left) 2006. Averaging kernels have been coloured according to their respective tropospheric air mass factor values.

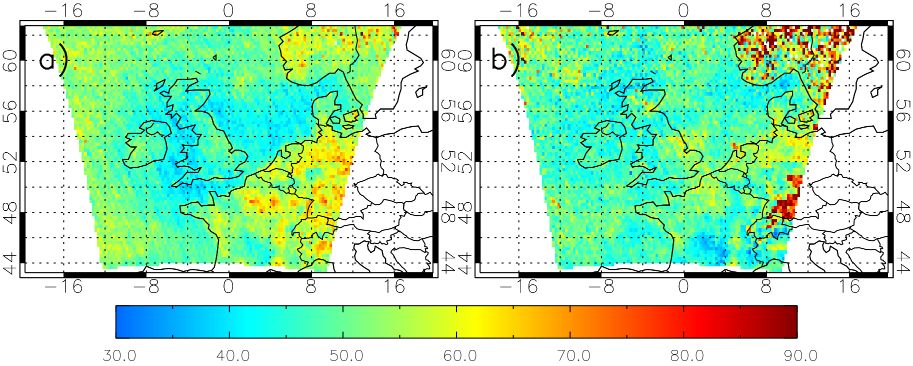


Figure 2. New seasonal satellite mean error, obtained by reduction of random error using the methodology described in Sect. 2, as a percentage of simple seasonal mean of satellite total error for 2006. Smoothing errors have been removed. **(a)** Summer and **(b)** winter.

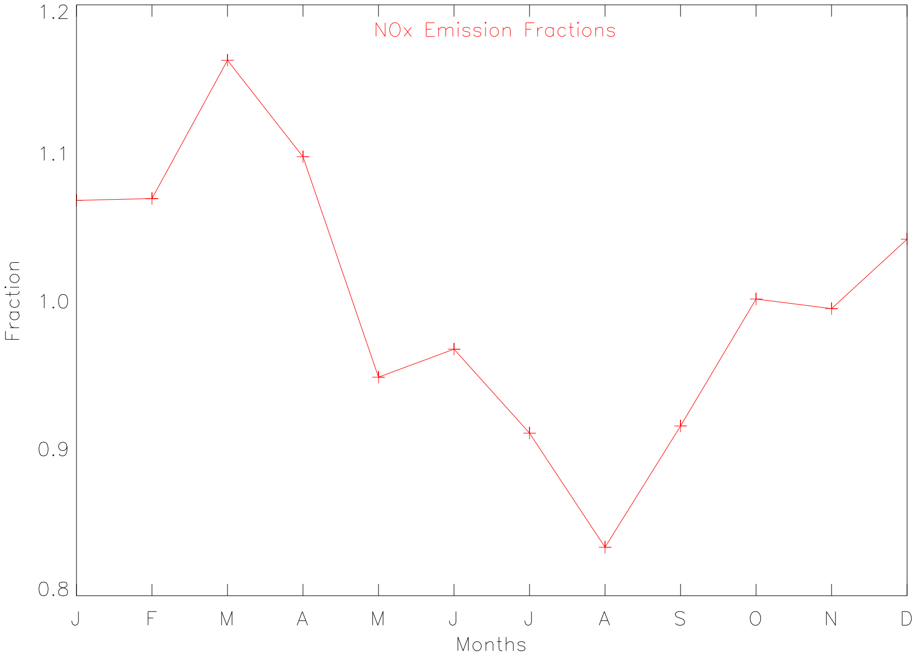


Figure 3. NO_x emissions seasonal cycle, based on Visschedijk et al. (2007) , which is applied to AQUM's NO_x emission annual totals.

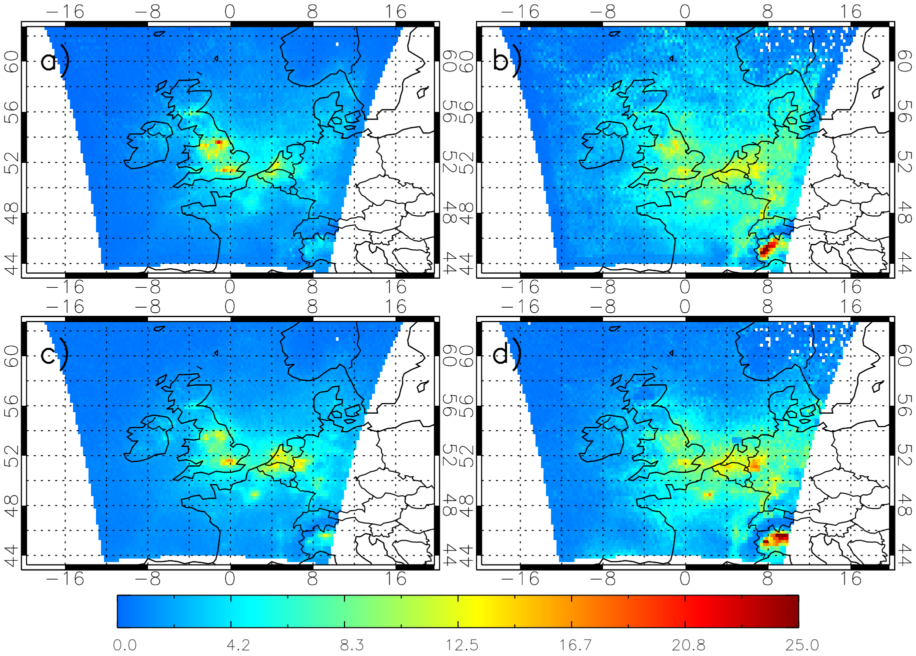


Figure 4. Tropospheric NO_2 column ($\times 10^{15}$ molecules cm^{-2}), 2006, for **(a)** AQUM Run C (with averaging kernels (AK) applied) summer, **(b)** AQUM Run C (AKs applied) winter, **(c)** OMI summer and **(d)** OMI winter.

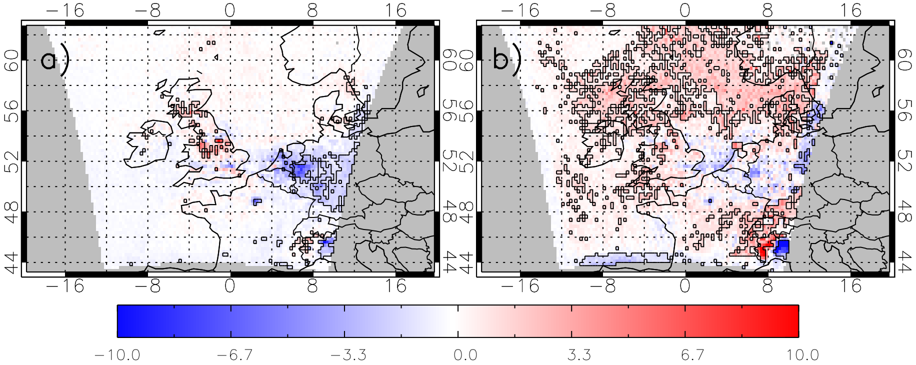


Figure 5. Mean bias in tropospheric NO_2 column ($\times 10^{15}$ molecules cm^{-2}), 2006, between AQUM Run C (AKs applied) and OMI for **(a)** summer (RMSE $= 3.68 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.65$) and **(b)** winter (RMSE $= 5.12 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.63$). The RMSE and FGE are over the UK between 8° W–2° E and 50–60° N and black polygoned regions show significant differences. Also the same for mean bias plots in Figs. ~~5–8~~6–9.

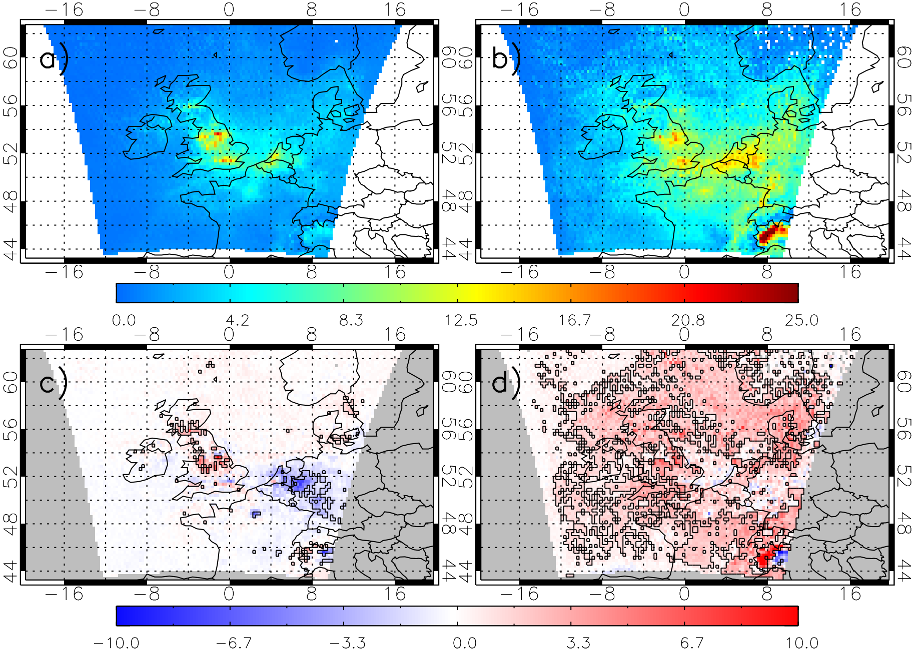


Figure 6. Tropospheric NO_2 column ($\times 10^{15}$ molecules cm^{-2}), 2006, from AQUM Run MACC (AKs applied) for **(a)** summer and **(b)** winter. AQUM Run MACC (AKs applied) and OMI mean bias for **(c)** summer (RMSE = 3.74×10^{15} molecules cm^{-2} and FGE = 0.63) and **(d)** winter (RMSE = 6.00×10^{15} molecules cm^{-2} and FGE = 0.65).

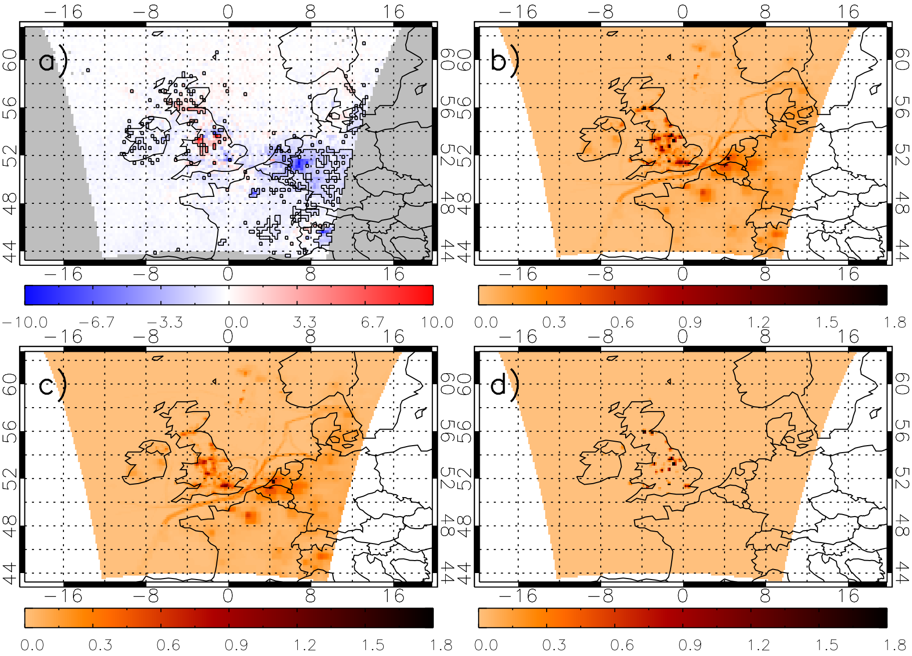


Figure 7. AQUM Run C (AKs applied)–OMI tropospheric NO_2 column ($\times 10^{15}$ molecules cm^{-2}) JJA 2006 mean bias. These are the control MBs to compare to the point source sensitivity experiments (RMSE $= 3.64 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.66$). NO_x emissions ($\times 10^{-9}$ kg m^{-2} s^{-1}), JJA 2006, used in AQUM for **(b)** Run C and **(c)** Run E1. **(d)** shows the difference between **(b)** and **(c)**.

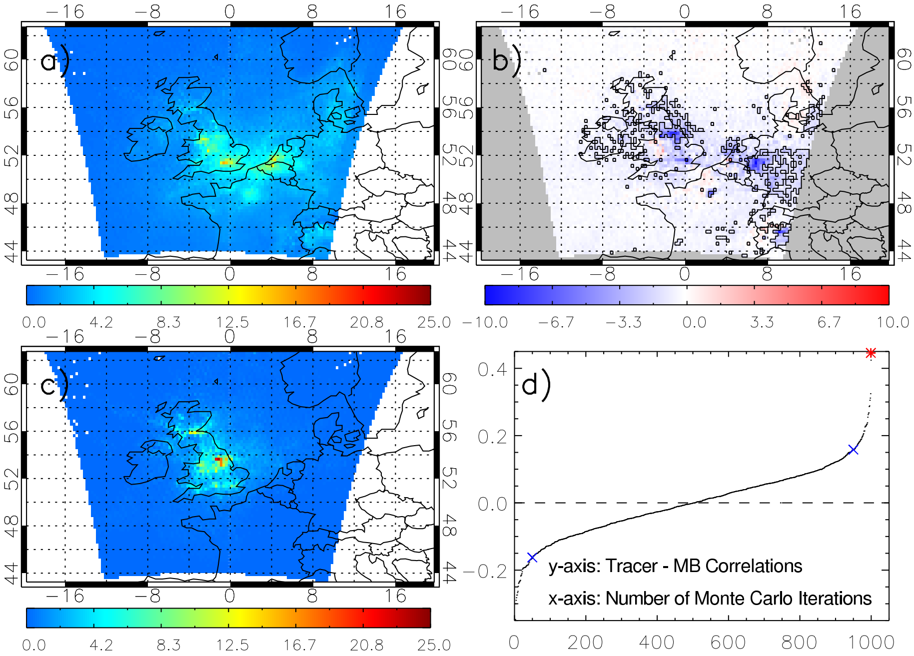
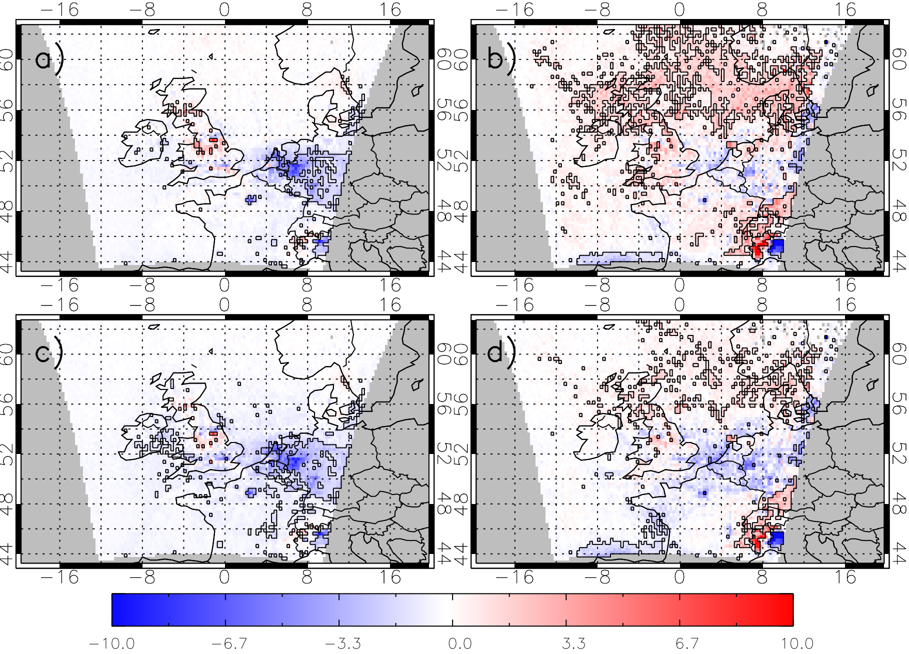


Figure 8. Tropospheric column ($\times 10^{15}$ molecules cm^{-2}), JJA 2006, for **(a)** AQUM Run E1 NO_2 (AKs applied), **(b)** AQUM Run E1 NO_2 (AKs applied)–OMI (RMSE = 3.02×10^{15} molecules cm^{-2} and FGE = 0.68) and **(c)** AQUM Run E2 Tracer (AKs applied). **(d)** Peak Run E2 and co-located Run C–OMI MB correlation (red star) significance distribution. Black dots are Run E2 and random Run C–OMI MB correlations. Blue X = 5th and 95th percentiles of the 1000 size sample.

~~MB in tropospheric column ($\times10^{15}$), 2006, between AQUM (AKs applied)–OMI for **(a)** summer γ=0.001 (RMSE=3.39$\times10^{15}$and FGE=0.65), **(b)** winter γ=0.001 (RMSE=5.05$\times10^{15}$and FGE=0.62), **(c)** summer γ=0.02 (RMSE=3.08$\times10^{15}$and FGE=0.67) and **(d)** winter γ=0.02 (RMSE=4.48$\times10^{15}$and~~



~~FGE=0.60).~~

Figure 9. MB in tropospheric NO_2 column ($\times 10^{15}$ molecules cm^{-2}), 2006, between AQUM (AKs applied)–OMI for **(a)** summer $\gamma = 0.001$ (RMSE $= 3.39 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.65$), **(b)** winter $\gamma = 0.001$ (RMSE $= 5.05 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.62$), **(c)** summer $\gamma = 0.02$ (RMSE $= 3.08 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.67$) and **(d)** winter $\gamma = 0.02$ (RMSE $= 4.48 \times 10^{15}$ molecules cm^{-2} and FGE $= 0.60$).