

Supplementary Material for “Influence of the North Atlantic Oscillation on European tropospheric composition: an observational and modelling study”

Supplementary Material (SM) 1: Automated Urban and Rural Network (AURN)

Surface ozone and NO₂ observations are from the Automated Urban and Rural Network (AURN), which comprises 198 sites measuring surface air pollutants since 1973 (DEFRA, 2015). However, not all sites have been monitoring for the full period. Here we have used midday AURN surface data from 2000 to 2010 for urban background, suburban, rural and remote AURN sites only (totalling 114 sites), which are representative of the surrounding area. Kerbside, urban traffic and industrial sites were excluded as they represent point measurements subject to large variability and levels of emissions. These monitoring sites have been interpolated to a 0.25° × 0.25° longitude–latitude grid.

Figure SM1a & b show AURN surface ozone sampled under the North Atlantic Oscillation positive (NAO+) and negative (NAO-) phases for December-January-February (DJF) 2000-2010. Surface ozone typically ranges between 35-60 µg/m³ and 25-55 60 µg/m³ in the two phases, respectively. Figure SM1c shows the difference between AURN ozone under NAO+ and NAO- with significant positive differences (95% confidence interval based on the Wilcoxon Rank Test, WRT, Pirovano et al., 2012) of 5-10 µg/m³ across the UK. For NO₂, NAO+ (Figure SM1d) and NAO- (Figure SM1e) composite concentrations typically range between 20-60 and 30-60 µg/m³, respectively. Figure SM1f highlights significant negative differences of -10 to -5 µg/m³. Though there are several significant positive differences. Overall, this suggests that the strength of westerly flow and the Atlantic storm tracks have an impact on UK surface air quality. NAO+ introduces ozone-rich Atlantic air masses into the UK, while reducing ozone sink gases (e.g. nitrogen oxides, NO_x) through transport out into the North Sea and continental Europe. Similar patterns are seen for Ozone Monitoring Instrument (OMI) tropospheric column NO₂ (Figure 2 of the main manuscript). Under NAO- conditions, westerly flow is reduced and re-orientated over continental Europe. Therefore, emitted gases such as NO_x are more likely to accumulate over source regions and destroy ozone producing NO₂. Even though NO_x in the presence of volatile organic compounds (VOCs) yield larger NO₂ concentrations, which can be photolysed to produce ozone, the intensity of winter-time incoming solar radiation is weaker leading to slower ozone production rates. Therefore, emission of NO_x results in greater ozone loss than production rates. The weakening of the westerlies also decreases the transport of ozone into the UK.

SM 2: TOMCAT Chemistry Transport Model (CTM) Evaluation

Nitrogen Dioxide (NO₂)

Previous studies have shown TOMCAT simulates tropospheric NO₂ with a suitable level of confidence. Monks et al., (2017) evaluated TOMCAT against OMI tropospheric column NO₂ for DJF and June-July-August (JJA) 2008. Over Western Europe and the North Atlantic, there were no significant ($|\text{mean bias}| > \text{satellite error}$) model biases. Over North America, there were some significant positive biases of 1-2 × 10¹⁵ molecules/cm². The largest significant biases in the region were over Scandinavia and Eastern Europe (3-5 × 10¹⁵ molecules/cm²). Savage et al., (2004) compared TOMCAT with GOME (Global Ozone Monitoring Experiment) observed tropospheric column NO₂ for 1997 and found that the model captures the seasonal cycle over Europe observed by the satellite but does underestimate its amplitude. Furthermore, Savage et al., (2004) found a strong correlation between TOMCAT and GOME monthly means over Europe (0.79) and North America (0.89). Emmons et al., (2015) found good agreement between TOMCAT and aircraft in situ NO₂ measurements as part of the POLMIP project. Comparisons with the UK AURN surface NO₂ (not presented here) show that TOMCAT has a systematic negative bias, which is sometimes outside the range of the observational variability, and this was also seen by Monks et al., (2017) at sites over central Europe. Huijnen et al., (2010) note that it is difficult for global models to reproduce surface in situ observations of short-lived species due to their coarse horizontal resolution, which might

partially explain the systematic negative bias. The monthly emissions inventories used in the model may also contribute to this bias. However, overall TOMCAT has sufficient skill to be used as a tool to investigate NAO influences on tropospheric NO₂ on the spatial and temporal scales observed by satellites.

Peroxyacetyl Nitrate (PAN)

Pope et al., (2016) used Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) measurements in combination with TOMCAT to investigate the global distribution of PAN. MIPAS PAN measurements were provided by two groups; University of Leicester (UoL) and the Karlsruhe Institute for Technology (KIT). Significant differences were identified between the two products. In the tropics at 150 hPa, TOMCAT was biased positively with respect to the UoL and negatively with respect to KIT products. However, in the mid-latitudes, TOMCAT significantly underestimated both MIPAS products at this pressure level. In the same study, TOMCAT was also compared with aircraft data over North America and Greenland. In the summer-time flights agreement was good although in the spring campaign TOMCAT was positively biased. Comparisons with the Emmons et al., (2000) aircraft climatology generally showed TOMCAT to simulate PAN well in the troposphere. Given the mixed model-satellite comparisons, but reasonable model-aircraft evaluation, we use TOMCAT PAN to investigate links with NAO circulation patterns.

Ozone (O₃)

Richards et al., (2013) compared summer-time tropospheric ozone over the Mediterranean from TOMCAT with observations by the Tropospheric Emission Spectrometer (TES, 2005-2008) and the Global Ozone Monitoring Experiment – 2 (GOME-2, 2007-2008). Agreement was found to be good in both cases, with correlations of 0.52-0.76 with respect to TES and 0.91-0.94 with respect to GOME-2. Globally, Monks et al., (2017) found that TOMCAT agreed well with GOME-2 (0-6 km sub-column) ozone in JJA, but significantly underestimated tropical ozone in DJF by -5 to -3 DU. Monks et al., (2017) also compared TOMCAT against ozonesondes in the northern hemisphere (NH, 30-90°) and between 1000-750 hPa the model generally reproduces their seasonal cycle within their uncertainty range. The layer-average correlations ranged from 0.77-0.80 depending on year and TOMCAT biases were -1.3 to -1.2%.

In the current study, TOMCAT was evaluated against ozonesondes in the UK, from Lerwick and the Automated Urban and Rural Network (AURN) sites, and sub-column (0-6 km) ozone retrieved from the Ozone Monitoring Instrument (OMI) by the Rutherford Appleton Laboratory. Figure SM2a shows the midday seasonal cycle for TOMCAT and AURN +/- their uncertainty ranges (standard deviations). Overall, TOMCAT has a significant positive bias (13.2 µg/m³) and root mean square deviation (17.8 µg/m³) with respect to the AURN data, which are outside their estimated observational error (standard error with the influence of autocorrelation removed (Bence, 1995) of 9.3 µg/m³). The overestimation of TOMCAT ozone is most pronounced between May and September. However, TOMCAT performs well in winter and is within the range of observational variability. This is backed up by the ozonesonde profile comparisons in Figure SM2b & c, where TOMCAT is seen to reproduce well the structure of the mean ozone profile for April to Sept and in winter (October - March) the ranges of model and ozonesonde variability are shown to overlap.

The OMI 0-6km sub-column ozone retrieved by RAL using an algorithm derived from that reported by Miles et al., (2015a) was compared with TOMCAT for the winter period (October – March) between 2005-2015. Initially, the OMI sub-column ozone was sampled under the different NAO phases, but it was difficult to identify a robust signal. Therefore, the OMI product was instead used to evaluate TOMCAT over the North Atlantic and Europe for an extended winter period. Figure SM3a & b show TOMCAT and OMI ozone sub-columns. The model has been co-located in time and space with the satellite observations and the OMI averaging kernels (AKs) have been applied to account for the satellite's vertical sensitivity, as discussed by Miles et al., (2015b). In the North

Atlantic region, winter-time TOMCAT and OMI ozone ranges between 20 and 25 DU, but OMI is larger by 2-3 DU over the continents. Figure SM3c indicates biases of -3 to 3 DU over the North Atlantic, which are not significant. Biases are classed as significant if $|\text{mean bias}| > \text{satellite error}$ (Figure SM3d). However, in the tropics and sub-tropics, the comparisons between TOMCAT and OMI degrade. The ozone hotspots (i.e. northern India, Arabian Sea, Red Sea and Gulf of Guinea) seen by OMI are simulated by TOMCAT but typically 5 DU lower. TOMCAT appears to miss enhanced ozone over the South Atlantic and sub-tropical North Atlantic (25-27 DU). South of 30°N, TOMCAT significantly underestimates OMI sub-column ozone by -15 to -5 DU across the majority of the domain. This was also seen by Monks et al. (2017) when comparing TOMCAT with GOME-2 sub-column ozone.

Overall, TOMCAT reasonably reproduces winter-time ozone over north-western Europe when compared with surface sites and ozonesondes. Comparisons with satellite data also show reasonable winter-time agreement over the North Atlantic and Europe when/where the NAO has the largest influence on the meteorology. Therefore, we are confident to investigate the impacts of NAO circulation pattern on tropospheric ozone using TOMCAT.

SM 3: Satellite Derived Tropopause Height

Figures SM 4 & 5 show the winter-time (DJF), averaged between 2002 and 2012, tropopause height (TH) derived from the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) sampled under the NAO phases. For each MIPAS scan the thermal tropopause was calculated from the temperature profile retrieved for this geolocation in an earlier step of the processing chain. The calculation of the thermal tropopause was performed according to the definition of the World Meteorological Organisation (WMO). Corresponding to this definition the tropopause is the lowest level, at which the lapse rate dT/dz is less than 2K/km and does not exceed this value in an overlying layer of at least 2 km.

During NAO+, the minimum TH (8-9 km) is located over northern Canada and maximum TH (13 km) is over the sub-tropical North Atlantic. Under NAO- conditions, peak TH of 12-13 km is again over the sub-tropical North Atlantic, but this does not propagate as significantly into the North Atlantic as in NAO+. Over northern Canada and Iceland/southern Greenland, the TH has increased to 9-10 km and 10-11 km, respectively, when compared with NAO+. However, the TH over the Arctic is much lower (8-10 km) in NAO-, while nearly reaching 11 km in NAO-. The larger TH in NAO+ over the southern North Atlantic is linked to enhanced mid-latitude depressions and storm tracks, while the enhancement over Iceland/Southern Greenland in NAO- highlights the UTLS anticyclone above the climatological Icelandic low pressure system in the troposphere.

SM 4: Satellite Uncertainties and Errors

Satellite anomalies of the different chemical species sampled under the two NAO phases, relative to their winter-time average, will see a large cancelling of the systematic errors. Therefore, we only consider the random errors in this study. The random errors for tropospheric column NO₂, MIPAS PAN at 150 hPa and TES ozone profiles, sampled under the two NAO phases, are shown in Figure SM6. In Figures SM6a & b, the random errors are approximately 10-40%, represented as a percentage of the retrieved value, over north-western Europe. However, the random errors, which cancel by a factor of $1.0/\sqrt{n}$ where n is the sample size, are heavily dependent on the sample size. In winter, retrieval of NO₂ in the high latitudes is less frequent resulting in a reduced sample size and larger random errors (i.e. approximately 100%). MIPAS PAN errors typically range between 0-5% (Figures SM6c & d) over the North Atlantic, but peak at 15-20% over the Canadian Arctic region where the sample size is lower. TES lower tropospheric ozone random errors (standard deviations) in all four zones, see main manuscript, typically ranges between 10-20% in the region of peak TES sensitivity (900-650 hPa).

References

- Bence, J. R.: Analysis of Short Time Series: Correcting for Autocorrelation, *Ecology*, 76, 628–639, <http://www.jstor.org/stable/1941218>, 1995.
- DEFRA 2015 Automated Urban and Rural Network (AURN) (uk-air.defra.gov.uk/networks/network-info?view=aurn).
- Emmons, L. K., Hauglustaine, D. A., Muller, J., Carroll, M. A., Brasseur, G. P., Brunner, D., Staehelin, J., Thouret, V., and Marenco, A.: Data composites of airborne observations of tropospheric ozone and its precursors, *J. Geophys. Res.: Atmos.*, 105, 20 497–20 538, doi:10.1029/2000JD900232, 2000.
- Emmons, L. K., Arnold, S. R., Monks, S. A., Huijnen, V., Tilmes, S., Law, K. S., Thomas, J. L., Raut, J.-C., Bouarar, I., Turquety, S., Long, Y., Duncan, B., Steenrod, S., Strode, S., Flemming, J., Mao, J., Langner, J., Thompson, A. M., Tarasick, D., Apel, E. C., Blake, D. R., Cohen, R. C., Dibb, J., Diskin, G. S., Fried, A., Hall, S. R., Huey, L. G., Weinheimer, A. J., Wisthaler, A., Mikoviny, T., Nowak, J., Peischl, J., Roberts, J. M., Ryer-son, T., Warneke, C., and Helmig, D.: The POLARCAT Model Intercomparison Project (POLMIP): overview and evaluation with observations, *Atmos. Chem. Phys.*, 15, 6721–6744, doi:10.5194/acp-15-6721-2015, 2015.
- Huijnen, V., Eskes, H., Poupkou, A., Elbern, H., Boersma, K., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., et al.: Comparison of OMI NO₂ tropospheric columns with an ensemble of global and European regional air quality models, *Atmospheric Chemistry and Physics*, 10, 3273–3296, 2010.
- Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, *Atmospheric Measurement Techniques*, 8, 385–398, doi:10.5194/amt-8-385-2015, <http://www.atmos-meas-tech.net/8/385/2015/>, 2015a.
- Miles, G. M., Siddans, R., and Kerridge, B. J.: RAL Ozone Profile Algorithm Product User Guide Version 1.1, 2015b.
- Monks, S. A., Arnold, S. R., Hollaway, M. J., Pope, R. J., Wilson, C., Feng, W., Emmerson, K. M., Kerridge, B. J., Latter, B. L., Miles, G. M., Siddans, R., and Chipperfield, M. P.: The TOMCAT global chemical transport model v1.6: description of chemical mechanism and model evaluation, *Geoscientific Model Development*, 10, 3025–3057, doi:10.5194/gmd-10-3025-2017, <https://www.geosci-model-dev.net/10/3025/2017/>, 2017.
- Pirovano, G., Balzarini, A., Bessagnet, B., Emery, C., Kallos, G., Meleux, F., Mitsakou, C., Nopmongkol, U., Riva, G., and Yarwood, G.: Investigating impacts of chemistry and transport model formulation on model performance at European scale, *Atmospheric Environment*, 53, 93 – 109, doi:<https://doi.org/10.1016/j.atmosenv.2011.12.052>, <http://www.sciencedirect.com/science/article/pii/S1352231011013422>, aQMEII: An International Initiative for the Evaluation of Regional-Scale Air Quality Models - Phase 1, 2012.
- Pope, R. J., Richards, N. A. D., Chipperfield, M. P., Moore, D. P., Monks, S. A., Arnold, S. R., Glatthor, N., Kiefer, M., Breider, T. J., Harrison, J. J., Remedios, J. J., Warneke, C., Roberts,

J. M., Diskin, G. S., Huey, L. G., Wisthaler, A., Apel, E. C., Bernath, P. F., and Feng, W.: Intercomparison and evaluation of satellite peroxyacetyl nitrate observations in the upper troposphere–lower stratosphere, *Atmospheric Chemistry and Physics*, 16, 13 541–13 559, doi:10.5194/acp-16-13541-2016, <http://www.atmos-chem-phys.net/16/13541/2016/>, 2016.

Richards, N. A. D., Arnold, S. R., Chipperfield, M. P., Miles, G., Rap, A., Siddans, R., Monks, S. A., and Hollaway, M. J.: The Mediterranean summertime ozone maximum: global emission sensitivities and radiative impacts, *Atmos. Chem. Phys.*, 13, 2331–2345, doi:10.5194/acp-13-2331-2013, 2013.

Savage, N. H., Law, K. S., Pyle, J. A., Richter, A., Nüß, H., and Burrows, J. P.: Using GOME NO₂ satellite data to examine regional differences in TOMCAT model performance, *Atmospheric Chemistry and Physics*, 4, 1895–1912, doi:10.5194/acp-4-1895-2004, <http://www.atmos-chem-phys.net/4/1895/2004/>, 2004.

Figures:

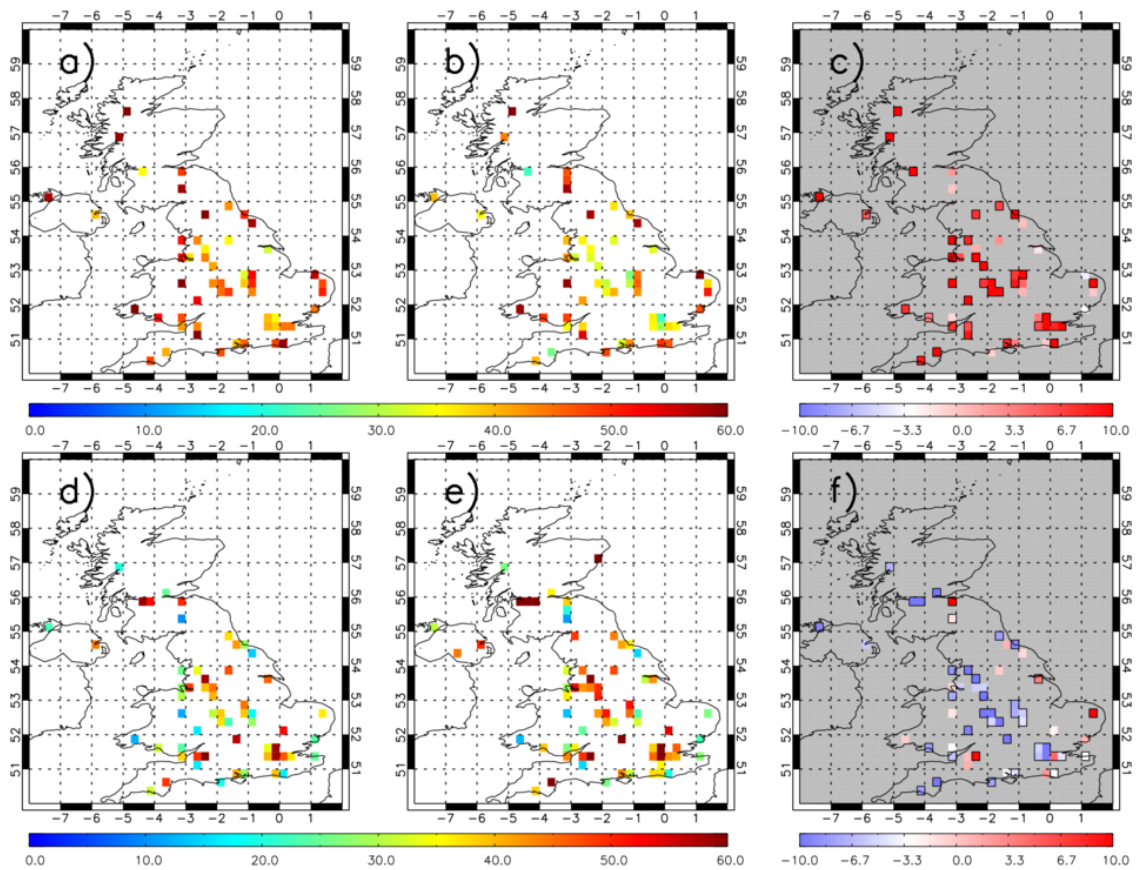


Figure SM1: Automated Urban and Rural Network (AURN) midday surface ozone and NO₂ ($\mu\text{g}/\text{m}^3$) sampled under the winter-time (DJF) NAO for 2000-2010. Panels a) and b) show AURN ozone sampled under the positive and negative NAO phases, while d) and e) show the same but for NO₂. Panels c) and f) represent a)-b) and d)-e) where the black polygoned regions show significant differences (95%) based on the WRT.

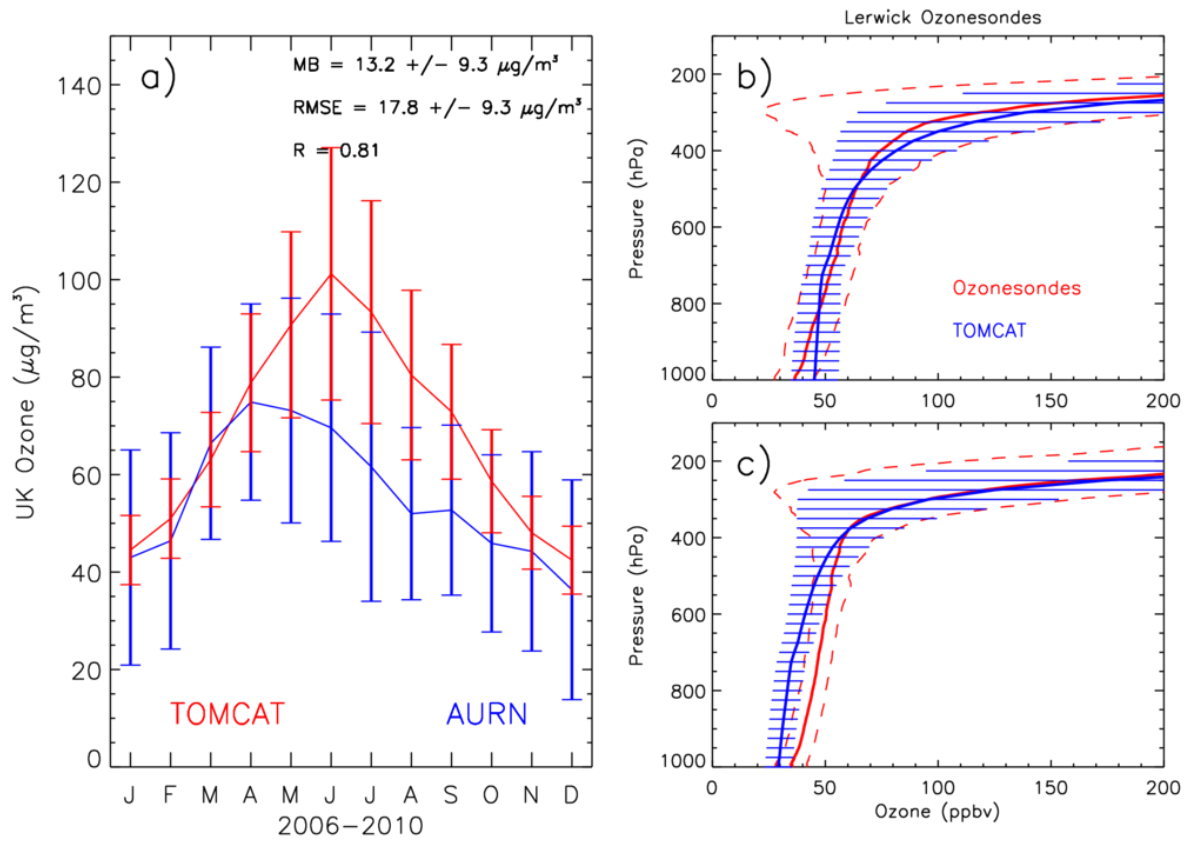


Figure SM2: a) TOMCAT (co-located to AURN sites) and AURN midday seasonal cycles for 2006 to 2010. Vertical lines show TOMCAT and AURN seasonal standard deviations. b) and c) show TOMCAT (red)-ozonesonde (blue) comparisons for Lerwick, Shetland Islands, UK, for summer (April–September) and winter (October–March), 2006–2015, respectively. Blue horizontal and red dashed lines show ozonesonde and model standard deviations.

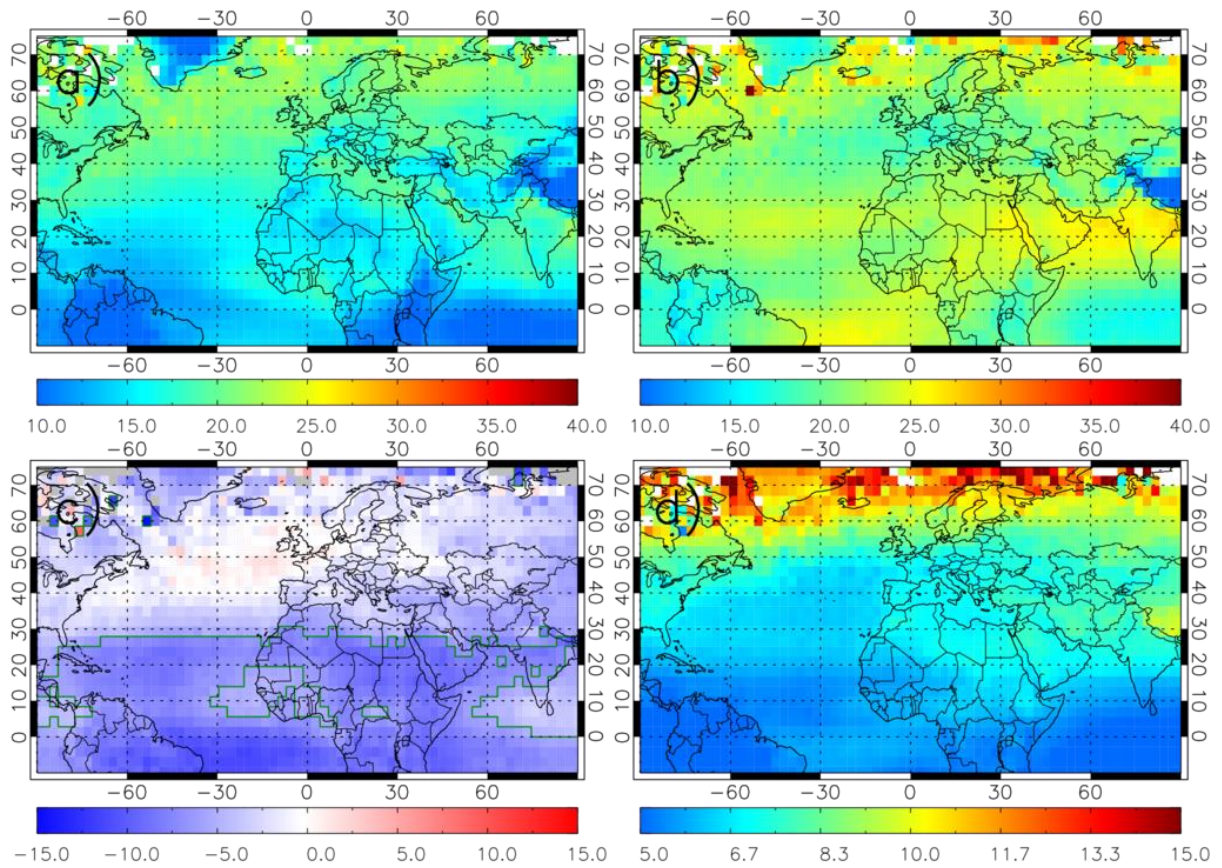


Figure SM3: Winter-time (November-March) sub-column (0-6 km) ozone (DU), 2006-2015, for a) TOMCAT, b) OMI, c) TOMCAT-OMI mean bias and d) OMI error. Green polygoned regions show significant differences where the magnitude of the mean bias is greater than the satellite error.

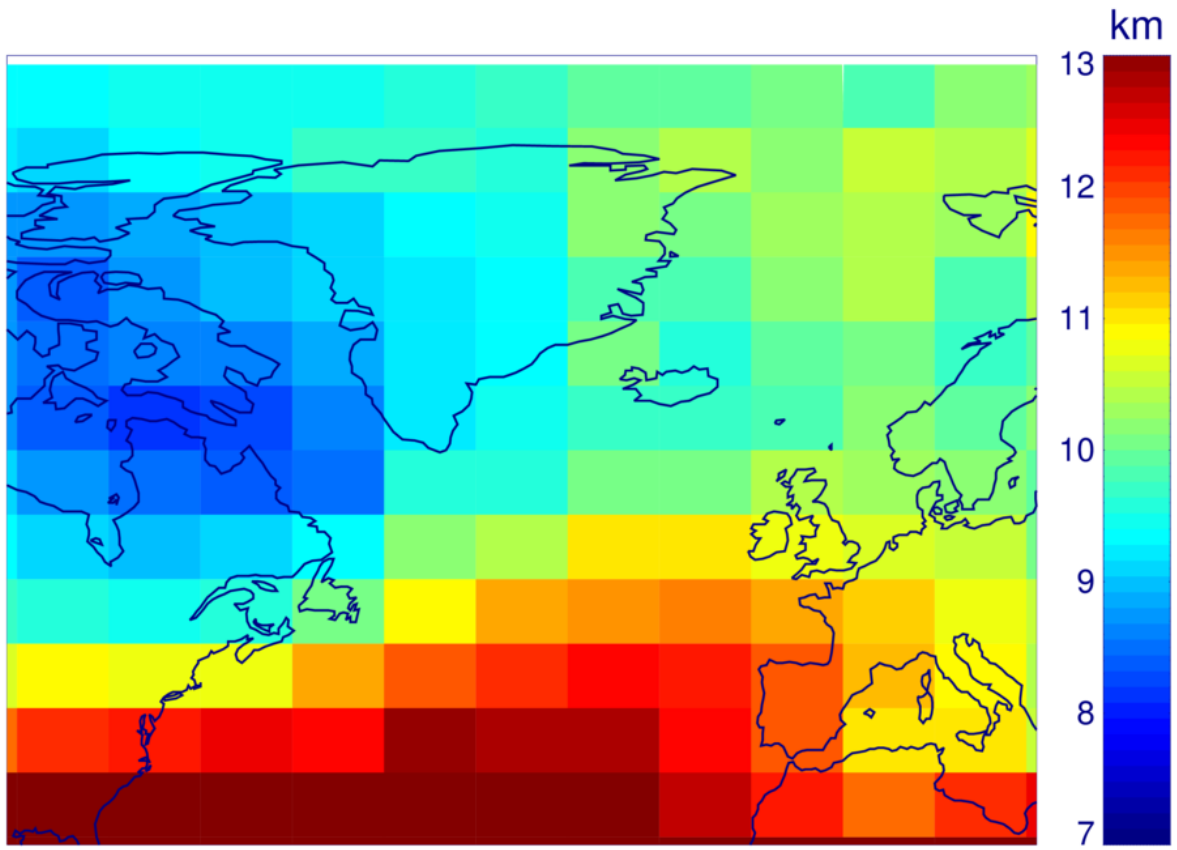


Figure SM4: Winter-time (DJF) MIPAS derived tropopause height averaged between 2002-2012 for NAO+.

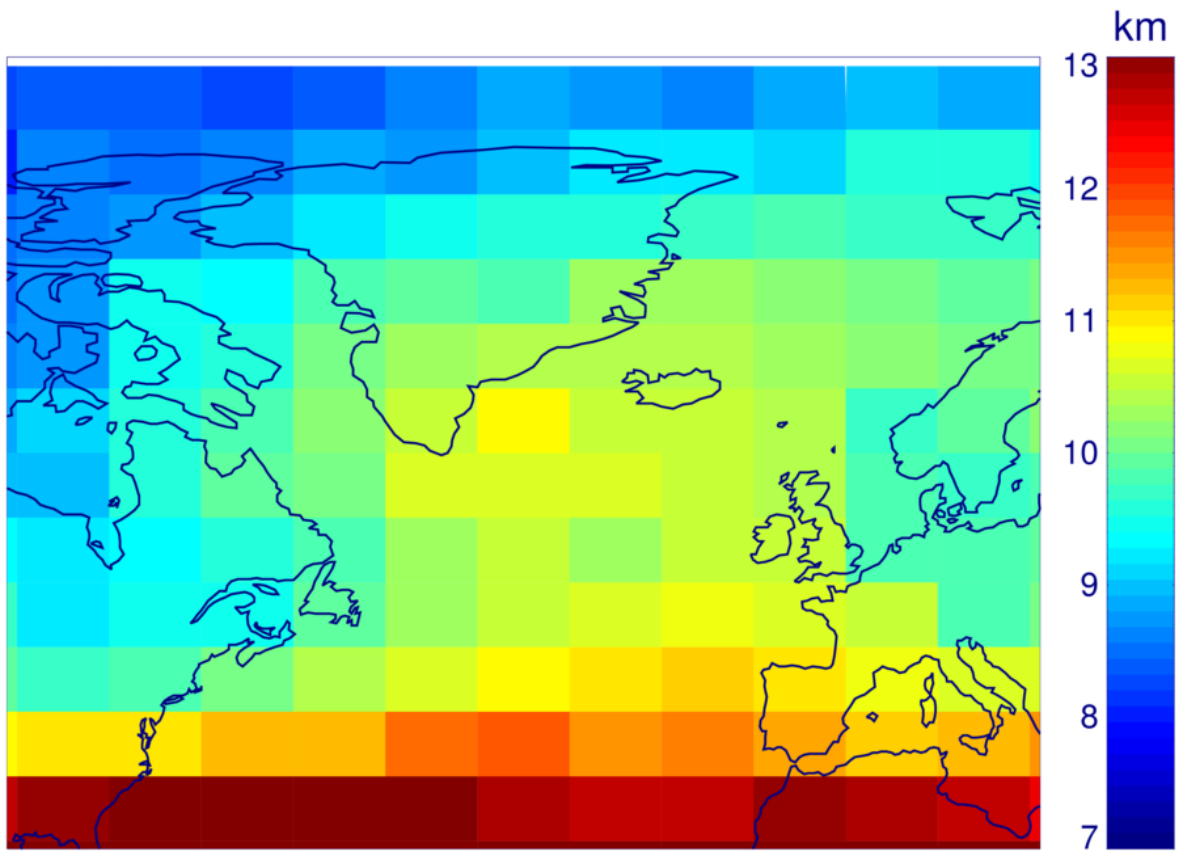


Figure SM5: Winter-time (DJF) MIPAS derived tropopause height averaged between 2002-2012 for NAO-.

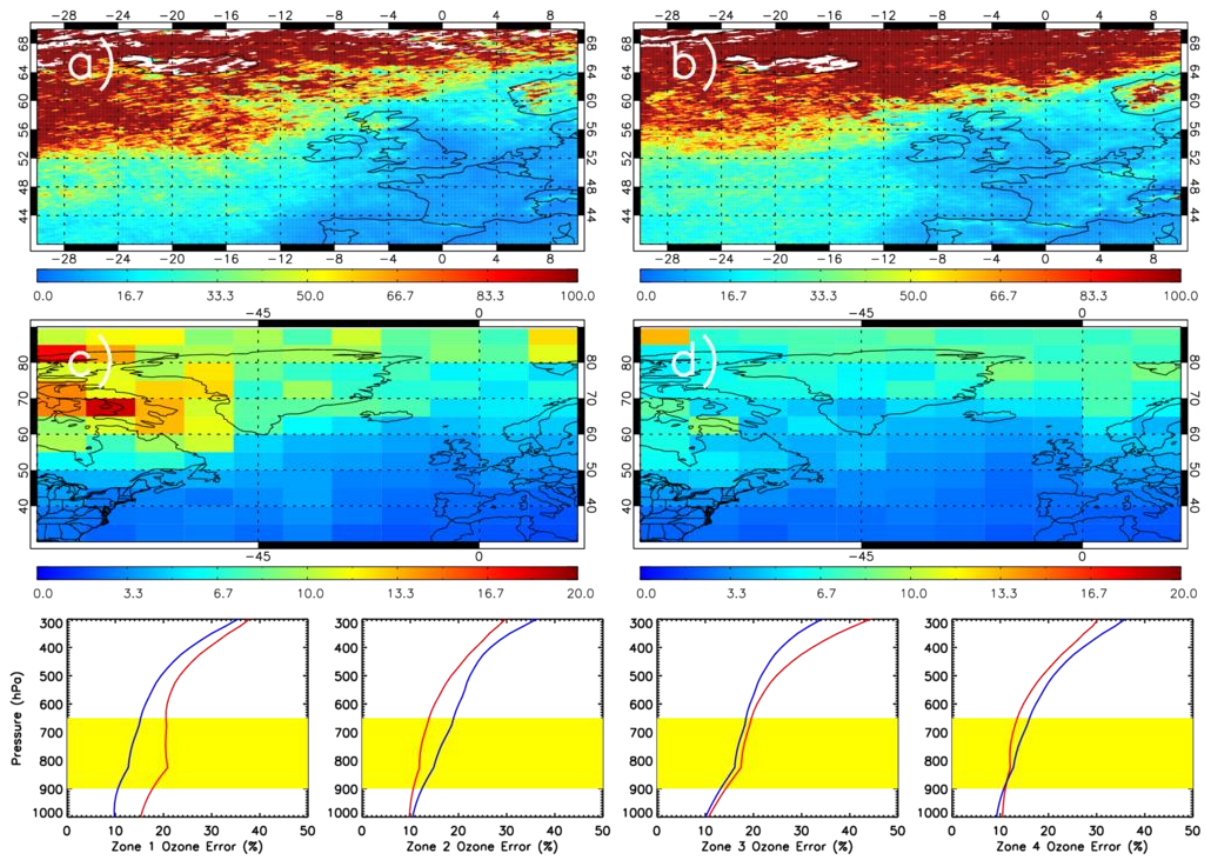


Figure SM6: Satellite random errors (%) for a) OMI tropospheric column NO_2 sampled under NAO+, b) OMI tropospheric column NO_2 sampled under NAO-, c) MIPAS PAN at 150 hPa sampled under NAO+, d) MIPAS PAN at 150 hPa sampled under NAO-. Zones 1-4 (bottom panels), from the main manuscript, show TES ozone random errors (standard deviations, %) under NAO+ (red lines) and NAO- (blue lines). The yellow box shows the region of TES peak sensitivity to retrieving lower tropospheric ozone.