Please find below detailed replies to your comments. Since we have made significant changes to the manuscript, we invite you to read our replies together with the new version of the manuscript (once it is uploaded to ACPD).

Multi model evaluation of short-lived pollutants over East Asia during summer 2008 (reviewer 1)

This paper reports on an exercise where six global and 1 regional model are performing the same model experiment (i.e. using the same ECLIPSE emissions, summer 2008) and compare model results with a collection of measurements. While the paper pulls together an interesting set observations, the paper is not well framed in terms of defining a key-question or hypothesis. For instance can the models be used to constrain emissions? Can the set of models be used to provide better estimates of air pollution, RF, or any other parameters, compared to earlier studies.

While the region of interest is highly relevant, it is not clear why the problem is tackled with six global models, and only 1 regional model. It is also not clear what the new findings are, compared to what is already known.

Authors' reply: We thank the reviewer for giving us the opportunity to explain more clearly the rationale our study. We appreciate their remarks and suggestions. The Introduction of the new version of the manuscript has been rewritten in order to clarify the purpose of this study and to better explain the objectives and rationale. It includes the following text:

"Evaluating the CLimate and air quality ImPacts of Short-livEd pollutants~(ECLIPSE) project developed new emission inventories for present-day global SLP emissions as well as future scenarios designed to benefit both air quality and climate with a focus on Asia and Europe (see Stohl et al. (2015) for discussion of the ECLIPSE rationale and summary of results)."

« An important component of ECLIPSE is the so-called reality check to evaluate model performance over pollutant source (Europe : Stohl et al. (2015), China/Asia~, the focus of this paper) and receptor (Arctic : Eckhardt et al. (2015)) regions. In these evaluations, the ECLIPSE models were run with the same present-day ECLIPSE emission inventory~(ECLIPSEv4a) for 2008 and 2009. Note that the same global models were used to estimate sector/regional emission responses and, in a sub-set of cases, to predict, using the ECLIPSE emission scenarios, future atmospheric composition and associated impacts on climate and air quality. The ECLIPSE global chemistry-climate models may not be the most suitable tools to assess air quality impacts, however they are the tools used to evaluate climate and air quality impacts together. To address this point, a regional model is also included in the evaluation, and one of the global chemical-transport models is run at relatively high horizontal resolution (50 km) compared to the other global models. »

We note that the reviewer has mainly focused their remarks on the analysis of the trace gas results. The new findings about trace gas results including a discussion about the causes for model discrepancies are now detailed in Sect. 3.6. We have also expanded the revision of the paper to include further analysis of the aerosol results (Sect. 4.3).

I could imagine that the paper would be better placed in the ACP sister journal, GMD (characterizing model performance), an option that the editor and authors should seriously consider. In summary I think the paper in its current form has to little scientific novelty to justify publication in ACP. The authors could consider however choosing a more in depth discussion of certain aspects of the evaluation. In my comments below I mention some potential deepening of the analysis. I suggest that such work doesn't necessarily have to be done with all models, but additional sensitivity studies with 1 or 2 models could make the paper more relevant and attractive.

Authors' reply: We thank the reviewer for their suggestions. Following consultation with the Editor, we decided to follow the recommendation of both anonymous referees to analyse out results in more depth and continue with ACP. We have therefore substantially rewritten the paper providing additional scientific analysis that was, as the reviewer points out, missing from the submitted version of the manuscript. In addition, this paper does not focus on describing model developments, it aims to quantify the ability of models to accurately simulate the spatial and vertical distributions of trace gases and aerosols over Asia in order assess uncertainties related to model estimates of air quality and climate impacts of SLPs.

The discussion and analysis of the results in the paper have been substantially updated. New findings are now detailed in two discussion sections on the trace gases (Sect. 3.6) and aerosols (Sect. 4.3). This includes further analysis of causes in model-observation discrepancies based on, for example, trace gas ratios and modelled diurnal cycles, in the case of ozone. We also included a sensitivity study, using one model, to quantify the impact of additional emission mitigation (vehicles, power/chemical plants) in Beijing province in the context of the Beijing Olympic and Paralympic Games (August/September 2008).

Detailed comments:

P 11051 l 5 troposphere includes surface=>better tropospheric columns

Authors' reply: "troposphere" has been replaced by "tropospheric columns" where necessary in the text.

l 6 I don't think a mismatch of NO_2 between global models and a few stations can be conclusive on emissions, given coarse resolution of the models. The regional model should be more representative but even then it will depend much on the vicinity of local sources.

Authors' reply: We agree with reviewer #1 that this conclusion was too general and not backed up by the results. The abstract and text in the main paper (e.g. Sections 3.3, 3.6) have been revised to better reflect the finding of this study with regard to NO_2 distributions and NOx emissions.

l. 9 I think to some extent this provides the argument that coarse resolution models can not be used for such detailed comparison in polluted areas. This is not new knowledge.
Authors' reply: We agree that it is difficult to compare global models with surface data collected in urban locations, for example. On the other hand, it is important to assess these models over a variety of scales and locations ranging from urban to rural and on continental scales. Such models need to be able to capture average patterns in terms of gradients from polluted to background environments as well as vertical distributions of pollutants. As explained above, these models have been used in the ECLIPSE project to assess air quality and climate impacts together. With regard to the representativeness of the model/observations comparison we now explain, in several places in the text, which observations we have used, for which purpose and which spatial scales they represent.

For example, in the Introduction:
"In order to assess model performance over East Asia for air quality, as well as climate, we use a variety of different datasets covering the urban, regional, and continental scales. Ozone, aerosol and precursor data at surface sites in urban and rural locations are used, together with CAREBEIJING aircraft data collected in the lower troposphere south of Beijing, to evaluate model performance in terms of local and regional pollution from major emission regions. Continental scale horizontal and vertical transport of ozone and aerosols, important for radiative impacts, are assessed downwind of the main emission regions using aerosol lidar data as well as satellite aerosol lidar and tropospheric ozone, CO and NO_2 column data."

l. 10 What are the important conclusions? Are the global model RFs all wrong? What did we learn?
Authors' reply: The Abstract and Summary section have been rewritten to make the our conclusions much clearer to the reader. We have also, as noted above, included 2 new sections discussing possible causes for the trace gas and aerosol discrepancies compared to the observations and the implications of these findings for assessment of air quality and climate impacts. Several points are highlighted that contribute to model uncertainty with regard to radiative forcing estimates. Here are the main conclusions :

" Models show systematic positive biases in ozone, especially at rural surface locations, and compared to satellite data downwind of major Chinese emission regions. The general underestimation of CO over and downwind of emissions is linked to this, most likely due excessive destruction by OH, suggesting that CO lifetimes are too short. Reasons for ozone differences varies between models but appears to be linked to model ability to simulate VOC and NOx regimes in polluted and less polluted environments. This may also be linked to inter-model spatial variability in compared to NO2 surface data and NO2 satellite column data. The latter, however, indicates a possible underestimation in Nox emissions over Korea and Japan as well as under (over)-

estimation of emissions to the south/east (west) of the Chinese NCP emission region. These findings point to the need to employ adequate model resolution to improve simulated responses to emissions when moving from ozone titration to ozone production regimes within large polluted conurbations, their surroundings and downwind. Overestimation of Asian ozone and its transport downwind implies that radiative forcing from this source may be overestimated. Sensitivity analyses, based on one model, suggest that emission mitigation over Beijing cannot explain these discrepancies.

Satellite-derived AOD measurements were reproduced quite well by the models over China even if evaluation of individual aerosol components over Asia overestimate ECLIPSE model-mean surface BC and sulphate aerosols in urban China in summer 2008. The effect of the short-term mitigation measures taken during the Olympic Games in summer 2008 is too weak to explain divergences between the models and observations. Our results rather point to an overestimation of emissions close to the surface in urban areas, particularly for SO2. A potential reason for this is the fact that the spatial distribution of power plant emissions has been changing dramatically in the last decades in China Liu et al. (2015), a change that has not been captured well in the ECLIPSEv4a dataset. A strong underestimation by ECLIPSE models of aerosol loadings has been identified over northern India, suggesting that the emissions of BC and precursors of other aerosols are underestimated in the ECLIPSEv4a inventory. Improvements have already been included, such as higher emissions from kerosene lamps, in the ECLIPSEv5 dataset. Model deficiencies in the representation of pollution accumulation due to the Indian monsoon may also play a role. The underestimation of the scattering aerosols in the lower troposphere, above the boundary layer, suggests too much vertical transport of pollutants towards the free troposphere and/or insufficient deposition in the boundary layer, leading to overestimated aerosol residence times in models."

p. 11055 l. 15 Although of course attractive to use, CAREBEING was taking place in a period of lower than usual emissions, while this was not considered in the models. This is of course strange, but perhaps only minor fraction of analysis is affected? Not clear. Why not use one model with 'lower' emissions in the period and region of interest?
Authors' reply: A sensitivity run has been performed using the WRF-Chem model to quantify the local and regional effects of emission mitigation in Beijing Province (June-September 2008) based on the emission mitigation plan described in Wang et al. (2010). Pollutant emissions from transport, industrial and solvent use sectors were reduced in the Beijing region from 1-15 August 2008. Prior studies are discussed briefly in the Introduction and details and findings are discussed mainly in Section 3.6 and also in Section 4.3:
The following text is included in Section 3.6:
"To assess the possible impact of emission mitigation measures in Beijing during the period analyzed in this study, the WRF-Chem model was run for 2 weeks (1 to 15 August 2008) with reduced pollutant emissions from the transport, industrial and solvent use sectors were reduced following the mitigation strategy during the Olympics described in

Wang et al. (2010). For example, emissions of all species in the transport sector were reduced by 75% in Beijing and 20% in the area 200 km from Beijing, corresponding to eight model grid cells around Beijing in this model. Emissions linked to the industrial sector or to solvents were reduced by 50% in the same region. Most pollutant concentrations are reduced resulting, for example, in lower CO, by about 30 ppbv, locally in and around Beijing in the emission reduction run compared to the base run. This results in ozone reductions of up to 6-7 ppbv in the region of Beijing. Based on these results, it appears that these reduction measures cannot explain the discrepancies between the models and the observations discussed earlier."

Fig. 1 shows the difference between surface concentrations in the base simulation and surface concentrations in the simulation where the emissions have been reduced during the Olympics for two pollutants: CO and ozone. In general, most pollutants are reduced (CO, ozone, SO_2, BC, OC, SO_4) in Beijing and the impact of emissions abatement is too low to explain discrepancies in the models compared to observations.

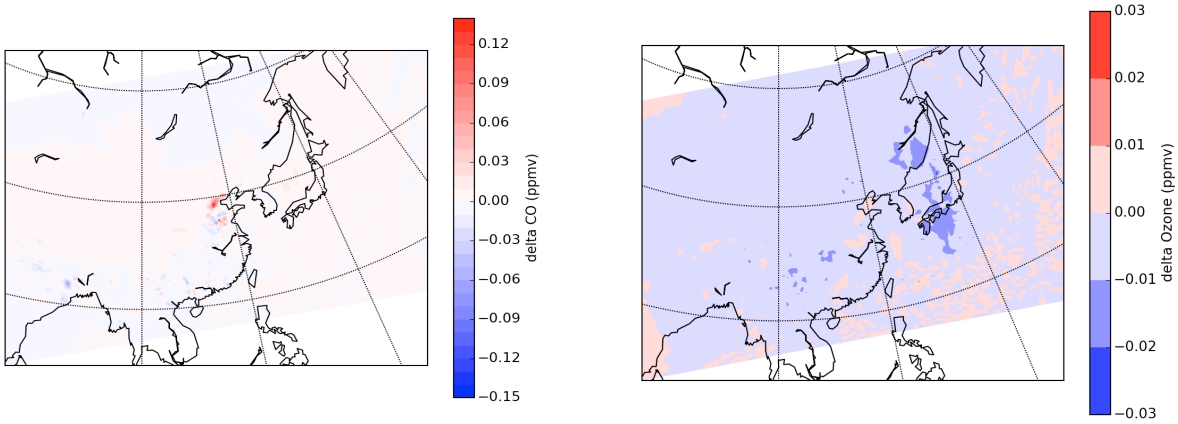

Fig. 1: Differences between surface concentrations in CO (left) and ozone (right) in the base simulation and surface concentrations in the simulation where emissions have been reduced during the Olympics using the WRF-Chem model.

p. 11056- table 1: Is it correct that NorESM was not running 2008? If so, why is it then included?

Authors' reply: NorESM is a chemistry-climate model. As such, it was not nudged to meteorology, but forced using reanalyses of sea surface temperature for 2008 (now noted in Table 1). Compared to the other models, driven by ECMWF or NCEP analyses, NorESM has a monsoon circulation that penetrates to far to the north over China. NorESM relative humidities are also higher over China, for example, compared to ECMWF and NCEP analyses which is likely to affect ozone photochemistry as well as aerosol formation. These factors, linked to model representation of the Asian monsoon, are discussed in the results sections. NorESM was included in this evaluation to provide consistency with companion papers (e.g Stohl et al., 2015) and because it was used to derive present-day and future emission impacts on air quality and climate.

p. 11056 Can you provide some statement on the effect of not including seasonality in most emissions? (Especially since the focus on the short period). Here a sensitivity experiment would have been appropriate.
Authors' reply: As mentioned in the text (P. 11056, L. 21-22), all ECLIPSE models used a seasonal cycle for the domestic heating sector, based on Streets et al. (2003). WRF-Chem uses seasonal cycles for all emission sectors, and only two other sectors (energy, industry) exhibit small variations (~28 %) relative to the mean emissions. Stein et al. (2014) highlighted the importance of including a seasonal cycle for the CO emissions to correct a low bias in the northern hemisphere found in many previous global studies, especially in winter, but it is unlikely that this is having a significant effect over Asia in summer where CO is already underestimated.

p. 11057 l. 11 There is some strange reasoning: on the on hand dust emissions in WRF-CHEM are too high, and two lines later it is argued that it doesn't matter anyway, so why leave it out?
Authors' reply: The reviewer is correct: the text in the manuscript was unclear. According to the literature, dust sources should not impact East Asia during summer. In the current version of WRF-Chem dust emissions are too high at this time year leading to unrealistic amounts of dust at high altitudes, and thus, excessive AODs and backscattering values (Saide et al., 2012). The following text has been added to the manuscript :
« WRF-Chem provides online dust and sea-salt emissions but only the latter are used in the ECLIPSE simulations due to an overestimation of dust loads, as reported by Saide et al. (2012). The main dust sources in East Asia are located in dry regions of China and Mongolia, north of the Himalayas (Taklamakan, Gobi and Gurbantunggut deserts). Most of the dust events occur in spring (Huang et al., 2013) whilst in summer, due to the Asian summer monsoon flux, rather little dust is transported to coastal areas (Kim et al., 2007). Thus, neglecting this source in WRF-Chem summertime simulations is not expected to introduce large bias in modeled aerosol loads. »

p. 11059 Was NorESM nudged, or calculating own meteorology? If the latter, why was it included?
Authors' reply: NorESM is a chemistry-climate model. It was not nudged to meteorology, but forced using reanalyses of sea surface temperature for 2008. See previous reply on this point.

p. 11059 what means FNL?
Authors' reply: As defined p. 11059, FNL means "final" analyses. This has been clarified in the text.

p. 11061 Does it make sense to compare WRFCHEM with the IASI ozone profiles?
Authors' reply: Yes, in our opinion, it makes sense to compare WRF-Chem with IASI ozone columns since the model provides results up to 20 km. The IASI sensor is most sensitive at 4 and 16 km as indicated by the averaging kernel (not shown). In addition, WRF-Chem ozone profiles are completed between 20-40 km.

p. 11061 explain the resolution of IASI and how models at different resolution where compared. One month or one year?
Authors' reply: Ozone columns are compared on a monthly basis (August 2008). IASI data have been averaged on a 1x1° grid and model results are scaled to this grid. The following sentence has been added in section 3.1:
"IASI data are averaged on a 1x1° grid and model results were scaled to this grid."

p. 11062 indeed global models have difficulties putting the correct location of the monsoon flux.
Authors' reply: We agree with this comment. As pointed out by the reviewer this may be having an influence on the NorESM modelled aerosol results. This point is now considered in section 4.3. NorESM-derived AODs is slightly underestimated in northern India. This is an indication that the effect of pollution accumulation due to the Indian monsoon is not adequately represented. NorESM has indeed lower winds associated with the Indian monsoon (Fig. 1).

p. 11063 l. 20 Which time period? interesting to see that the models are having so different bias/error with regard to NO2. However I would like to be sure that they really used the same emissions. Then some dedicated experiments could point to explain why the difference are so large (lifetime/mixing/deposition)?
Authors' reply: All the models use the same emissions, which have been specifically designed for the ECLIPSE project. Further analysis of model discrepancies in NO_2, CO and ozone is now included in the new Section 3.6. For this purpose we compared available model results to observed trace gas ratios at different locations. In the case of polluted locations differences compared to emitted and observed trace gas ratios (e.g. CO/NOx) are used to indicate a lack of chemical processing or mixing of air masses. We also use, for certain models, ratios of NO:NO_2 and ozone:NOz (NOy-NOx) to examine whether models are in VOC or NOx limited regimes with respect to photochemical ozone production.

p. 11064 If air quality is an issue, the authors should worry also about the vertical resolution of the first layer, the height of the emissions, the height of the measurements, and how the model was sampled.
Authors' reply: All the models prescribe surface emissions. The first vertical layer height varies between 25 and 192 m depending on the model and we did not find any significant correlation between the first layer height and modelled pollutant concentrations. Model results were sampled from the first layer in grid cells corresponding to station locations. This point is now addressed in Section 3.4:
"There is significant variability in modelled NO_2 compared the observations at polluted and rural sites. This could be caused by differences in model vertical resolution near the surface although no correlation was found between the height of the first model layers and pollutant concentrations."

p. 11065 Given the lifetime of CO, it would be instructive to learn what OH concentrations the models are having now, and what would be needed to reproduce the right gradients? Can the coarse resolution models be used for this?
Author's reply: We now include a more detailed discussion about the differences between the model results and CO observations in Section 3.6. We were able to examine OH concentrations in certain models. It appears that the underestimation in CO is linked to the overestimation in ozone, particularly downwind of China. As explained previously, the use of complementary datasets covering different environments and spatial scales allows us to draw more firm conclusions.
The following text is included in Section 3.6:
"The ECLIPSE models also systematically underestimate CO downwind compared to surface data over Korea, Japan, and compared to IASI CO data over Japan and downwind over the north-western Pacific Ocean. Whilst inclusion of additional seasonality in the ECLIPSE emissions (already included for domestic combustion), might improve agreement in winter and spring (Stein et al., 2014), this is unlikely to explain these summertime differences. Low model CO appears to be linked to the clear overestimation in modeled ozone at rural sites and compared to IASI 0-6~km column data. Excessive ozone resulting in too much destruction of CO by OH and may suggest that modeled CO lifetimes are too short. This hypothesis is consistent with the findings of Monks et al. (2015) who concluded that, in models run with the same emissions, differences in OH~(chemical schemes) are a more likely cause of the systematic CO underestimation in the Northern Hemisphere and the Arctic than differences in vertical transport. Indeed, we find that surface August mean modeled OH (not shown) is higher in the NorESM model (due to the penetration of the monsoon flux) compared to, for example, TM4-EPCL and WRF-Chem over the main Chinese emission regions. In contrast, excessive modeled CO over the central Pacific, where concentrations are low, may be due to the position of the Pacific anticyclone in the meteorological analyses used by the majority of models. A shift in the position of the anticyclone to the south could result in this pattern of negative (positive) biases over the north (south) Pacific potentially as a result of transport that is too zonal. This may also explain low modeled

CO in the Arctic noted by Monks et al. (2015).”

p. 11066 l 25 I understand that properly taking into account aerosol scattering is a big thing for NO2 retrievals. Was this done at these locations. What could the error?
Authors' reply: The observed NO2 columns were retrieved from GOME-2 satellite measurements using an algorithm described by Boersma et al. (2004). The NO2 columns are retrieved with a precision of 35-60 % and the main uncertainties are dominated by uncertainty in the estimate of the tropospheric air mass factors (AMF) (Boersma et al., 2004). Cloud fraction, surface albedo and profile shape are the most important factors influencing the AMF estimation uncertainties. Boersma et al. (2004) concluded that the aerosol correction is indirectly incorporated into the cloud correction and that an independent aerosol correction procedure would lead to differences of less than 10% in the correction factor. The following sentence has been added in Section 3.3 (Tropospheric NO2 columns):
“Column retrievals do not include corrections for aerosol scattering, which are estimated to less than 10 % by Boersma et al. (2004)”

p. 11068 l. 25 What was assumed for size distribution for particulate emissions?
Authors' reply: Particulate emissions are distributed differently depending on each model: three models only considered total mass (EMEP, NorESM and OsloCTM2), TM4-ECPL includes all the emitted mass in the fine mode, HadGEM3 distributed biofuel and fossil-fuel combustion in modes centred on 150 and 60 nm, respectively, both with a sigma of 1.59. ECHAM equally distributes ship, industrial and power-plant emissions in the accumulation and coarse modes, respectively centred on 75 and 750 nm with sigma of 1.59 and 2.0 and other emissions are equally separated between the Aitken mode, centred at 30 nm and the accumulation mode whereas biomass and fossil fuels are emitted in a single mode, centred at 15 nm with a sigma of 1.8. WRF-Chem distributes the emissions in its 8 size bins centred on 60, 117, 234, 468, 937.5, 1875, 3750, 7500 nm and with the following percentages: 25, 25, 15, 15, 7.5, 7.5, 2.5, 2% respectively.

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

Liu, F., Zhang, Q., Tong, D., Zheng, B., Li, M., Huo, H., and He, K. B.: High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010, Atmospheric Chemistry and Physics, 15, 13 299–13 317, doi:10.5194/acp-15-13299-2015, http://www.atmos-chem-phys.net/15/13299/2015/, 2015.

Saide, P. E., Spak, S. N., Carmichael, G. R., Mena-Carrasco, M. A., Yang, Q., Howell,

S., Leon, D. C., Snider, J. R., Bandy, A. R., Collett, J. L., Benedict, K. B., de Szoeke, S. P., Hawkins, L. N., Allen, G., Crawford, I., Crosier, J., and Springston, S. R.: Evaluating WRF-Chem aerosol indirect effects in Southeast Pacific marine stratocumulus during VOCALS-REx, Atmospheric Chemistry and Physics, 12, 3045–3064, doi:10.5194/acp-12-3045-2012, http://www.atmos-chem-phys.net/12/3045/2012/, 2012.

Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., and Clerbaux, C.: On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, Atmos. Chem. Phys., 14, 9295-9316, doi:10.5194/acp-14-9295-2014, 2014.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res.-Atmos., 108, 8809, doi:10.1029/2002JD003093, 2003.

Wang, B., Shao, M., Lu, S. H., Yuan, B., Zhao, Y., Wang, M., Zhang, S. Q., and Wu, D.: Variation of ambient non-methane hydrocarbons in Beijing city in summer 2008, Atmos. Chem. Phys., 10, 5911-5923, doi:10.5194/acp-10-5911-2010, 2010.

Worden, H. M., Cheng, Y., Pfister, G., Carmichael, G. R., Zhang, Q., Streets, D. G., Deeter, M., Edwards, D. P., Gille, J. C., and Worden, J. R.: Satellite-based estimates of reduced CO and CO_2 emissions due to traffic restrictions during the 2008 Beijing Olympics, Geophys. Res. Lett., 39, L14802, doi: 10.1029/2012GL052395, 2012.

Please find below detailed replies to your comments. Since we have made significant changes to the manuscript, we invite you to read our replies together with the new version of the manuscript (once it is uploaded to ACPD).

This paper describes the evaluation of a number of global and regional atmospheric composition models over the East Asian region during 2008. It focuses on the summer period, coinciding with the Beijing Olympics, and addresses comparisons with both gas-phase and aerosol observations from a range of different measurement platforms and locations.

The paper is interesting, and is potentially very useful in characterizing the behavior of the models involved. However, the major weakness is that it does not provide sufficient interpretation and attribution of why models differ from observations or each other. Reproducing observations over this highly polluted region requires a good simulation of emissions, chemical processing, local dynamics, and regional meteorology. The analysis presented includes suggestions of the causes of particular differences (e.g. that emissions of NOx are too high), but these are not firmly supported with clear evidence or convincing argument to back them up. A more thorough attribution of model differences is needed, so that developers and users of the models involved can make some progress towards model improvement, and so that others attempting a similar comparison are aware of the critical factors involved. This would provide a stronger message and make the paper much more useful to a wider audience.
Authors' reply: We thank the reviewer for their insightful comments and suggestions. We agree that the submitted manuscript needed to be improved in terms of a more in-depth analysis of the model evaluation results and the attribution of causes to explain modelled discrepancies compared to observations. We have substantially revised the paper to include, in particular, two new sections discussing the reasons for model differences in ozone and other trace gases (Section 3.6) and aerosols (Section 4.3). The Abstract and Conclusions have also been substantially revised to more accurately represent our findings. In addition, the Introduction has been updated in order to explain more clearly the motivation and context of this study. We describe the rationale employed to assess the performance of models used for air quality and climate assessments using a range of different observations. Since we have made substantial modifications to the manuscript, as requested by the reviewers, we highlight the main changes to the text that are relevant to the reviewer's specific points in the following:

" Models show systematic positive biases in ozone, especially at rural surface locations, and compared to satellite data downwind of major Chinese emission regions. The general underestimation of CO over and downwind of emissions is linked to this, most likely due excessive destruction by OH, suggesting that CO lifetimes are too short. Reasons for ozone differences varies between models but appears to be linked to model ability to simulate VOC and NOx regimes in polluted and less polluted environments. This may also be linked to inter-model spatial variability in compared to NO2 surface data and NO2 satellite column data. The latter, however, indicates a possible underestimation in NOx emissions over Korea and Japan as well as under (over)-estimation of emissions to the south/east (west) of the Chinese NCP emission region. These findings point to the need to employ adequate model resolution to improve

simulated responses to emissions when moving from ozone titration to ozone production regimes within large polluted conurbations, their surroundings and downwind. Overestimation of Asian ozone and its transport downwind implies that radiative forcing from this source may be overestimated. Sensitivity analyses, based on one model, suggest that emission mitigation over Beijing cannot explain these discrepancies.

Satellite-derived AOD measurements were reproduced quite well by the models over China even if evaluation of individual aerosol components over Asia overestimate ECLIPSE model-mean surface BC and sulphate aerosols in urban China in summer 2008. The effect of the short-term mitigation measures taken during the Olympic Games in summer 2008 is too weak to explain divergences between the models and observations. Our results rather point to an overestimation of emissions close to the surface in urban areas, particularly for SO2. A potential reason for this is the fact that the spatial distribution of power plant emissions has been changing dramatically in the last decades in China Liu et al. (2015), a change that has not been captured well in the ECLIPSEv4a dataset. A strong underestimation by ECLIPSE models of aerosol loadings has been identified over northern India, suggesting that the emissions of BC and precursors of other aerosols are underestimated in the ECLIPSEv4a inventory. Improvements have already been included, such as higher emissions from kerosene lamps, in the ECLIPSEv5 dataset. Model deficiencies in the representation of pollution accumulation due to the Indian monsoon may also play a role. The underestimation of the scattering aerosols in the lower troposphere, above the boundary layer, suggests too much vertical transport of pollutants towards the free troposphere and/or insufficient deposition in the boundary layer, leading to overestimated aerosol residence times in models."

General Comments

The purpose of the model evaluation is briefly outlined at the end of the introduction, but includes several very different issues (air quality, climate change, long-range transport) which have different evaluation requirements. However, these issues are lumped together in this study. Please provide a clear indication of which parts of the evaluation are relevant to which issue, so that the wider implications are immediately clear.

Authors' reply: The Introduction has been substantially revised to better explain the rationale for this study and to provide a clearer description about the different observations that were used to evaluate model performance on a range of scales and at different locations. The ECLIPSE models have been used, in related studies (summarized in Stohl et al., 2015), to estimate present-day and future chemical composition and their impacts on air quality and climate. The models used in these assessments range from chemistry-climate models to chemical-transport models, also including aerosol treatments. One goal of ECLIPSE was also to evaluate these models over emission source regions (Europe, Asia). As such the models need to be evaluated both in urban and rural locations for air quality as well as over source regions and downwind for climate change impacts. Comparison of model results with a range of different datasets including surface data, vertical profiles and satellite data allows assessment of model

performance over a range of scales relevant for air quality and climate.

East Asian emissions are rising rapidly and are less well quantified than in many other parts of the world, and they were temporally reduced by a poorly-known amount during the Olympic period. Given these large uncertainties, it is unclear why this region and period were chosen for the study. Please provide a clear justification for the choice. The impact of the Olympic emission reduction is one aspect of interest in many previous studies, but is not exploited here.
Authors' reply: Here, we present results from the evaluation of the ECLIPSE models over East Asia. As noted above, this region was targeted due to its still high pollution levels, climate impacts and as a region where SLP mitigation options are being actively considered. It is also a region where significant uncertainties surround model estimates of radiative forcing. For example, Kinne et al. (2006) showed important underestimation of observed AOD by multiple models over East Asia in summer and pointed out that uncertainties in the direct radiative forcing could be larger than inter-model differences in AOD suggest. Even in the recent AeroCom model comparison, inter-model variation in radiative forcing is largest in this region (Myhre et al., 2013a). Samset et al. (2014) pointed to overestimation of modeled BC compared to aircraft observations in the upper troposphere downwind of Asia over the Pacific suggesting that BC lifetimes are too long in current models and that BC direct radiative forcing is overestimated by about 25%.

The ECLIPSE model evaluation over East Asia focuses on the summer period (August and September 2008). This was motivated by the availability of intensive observations from the CAREBEIJING 2008 measurement campaign (the only campaign conducted during 2008 and 2009 for which data were available) and by the fact that severe ozone pollution episodes still occur over this region, and the wider Northern China Plain (NCP) at this time of year.

We have revised the text in the Introduction to include:
"The ECLIPSE model evaluation over East Asia focuses on the summer period (August and September 2008). This was motivated by the availability of intensive observations from the CAREBEIJING 2008 measurements campaign (Huang et al. 2010, Zhang et al., 2014) and by the fact that severe ozone pollution episodes occur over NCP at this time of year even if the maximum is generally earlier in the late spring for trace gases (Naja and Akimoto, 2004; Li et al., 2007; He et al., 2008; Safieddine et al., 2013) and aerosols (Cao et al., 2004; Sun et al., 2004; Yang et al., 2005; Huang et al., 2006). During the summer months, whilst the monsoon circulation brings cleaner air from the Pacific Ocean into southern and eastern Asia reducing pollutant concentrations (Lin et al., 2009; Kim et al., 2007). However, high pollution episodes with enhanced aerosol concentrations and decreasing visibility still occur in coastal regions, due to increases in relative humidity increasing aerosol sizes (Flowers et al., 2010). The monsoon flux also induces transport of high ozone concentrations inland (He et al., 2008)."

To assess the possible impact on our results of emission mitigation measures applied in

the context of the Beijing Olympic and Paralympic games, the WRF-Chem model was run for 2 weeks (1 to 15 August 2008) with reduced pollutant emissions from the transport, industrial and solvent use sectors, following the mitigation strategy during the Olympics described in Wang et al. (2010). According to the related literature (e.g. Zhou et al., 2010; Worden et al., 2012), the impact of such emission mitigations was mostly local and only affected the Beijing area. Thus, regarding our observation datasets, only the in-situ observations in Beijing (at the Peking University, Fig. 4 and 10) were affected by them. For example, emissions of all species in the transport sector were reduced by 75% in Beijing and 20% in the area 200 km from Beijing, corresponding to eight model grid cells around Beijing in this model. Emissions linked to the industrial sector or to solvents were reduced by 50% in the same region. Most pollutant concentrations are reduced resulting, for example, in lower CO, by about 30 ppbv, locally in and around Beijing in the emission reduction run compared to the base run. This results in ozone reductions of up to 6-7 ppbv in the region of Beijing. With regard to aerosols, results from the WRF-Chem simulations with reduced emissions due to additional mitigation measures in the Beijing area show that the measures taken for Olympic Games leads to small reductions in surface BC, OC and sulphate concentrations by 0.3, 1 and 1μg m−3 , respectively. Based on these results, it appears that these reduction measures cannot explain the discrepancies between the models and the observations discussed earlier.

Fig. 1 shows the difference between surface concentrations in the base simulation and surface concentrations in the simulation where the emissions have been reduced during the Olympics for two pollutants: CO and ozone. Concentrations of CO, ozone, SO_2, BC, OC, sulphates decrease, except for NO_2 that increases due to a competition between NO and NO_2, in Beijing. Overall, we find that the impact of emissions abatement is too low to explain discrepancies between the models and the observations. These results are described in sections 3.6 (trace gases) and section 4.3 (aerosols).

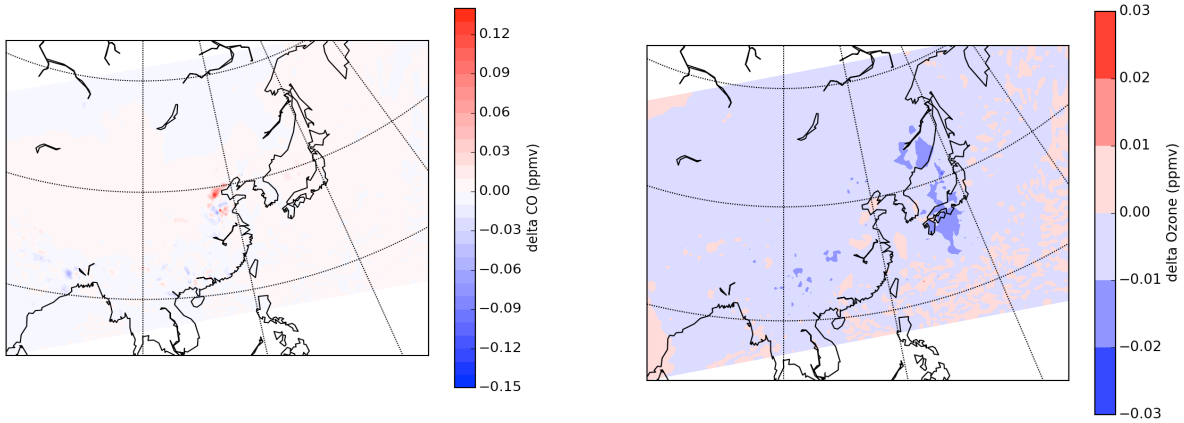

Fig. 1: Differences between modelled surface concentrations in CO (left) and O3 (right) in the base simulation and surface concentrations in the simulation where emissions have been reduced during the Olympics using the WRF-Chem model.

The main discussion sections (e.g. Section 3.5) are insubstantial, major biases are missing, and the links between biases that are described are not clearly and logically laid

out. The result is that these sections are inconclusive and provide the reader with little new information about why the models and observations differ. While it may not be possible to diagnose model biases completely with the limited information available, clearer analysis of the data will provide new and useful information. One issue is that the satellites provide estimates of NO2 (not NOx) and that emissions are of NO (not NOx). Biases in chemical processing thus strongly influence comparison with measurements, and biases are expected in this region where aerosol loadings can be very high. Another issue is that no attempt is made to assess how representative the surface or aircraft measurements are of the wider region, and thus whether it is reasonable to expect the models to match them. A clearer and more detailed discussion of the expected chemistry and transport biases would make the paper much more useful.

Authors' reply: We agree with reviewer #2 that we needed to analyse our results in more detail in order to provide more concrete and useful conclusions. For trace gases, we have now included a new Section 3.6 that examines possible reasons for model discrepancies in much more detail. We identify several rather robust and systematic biases that are common to the majority of models and which have implications for model ability to correctly assess air quality and climate impacts of Asian pollution. In order to examine the causes of the discrepancies in different models, we make use of modelled trace gas ratios (e.g. CO/NOx, ozone/NOz) together with a more detailed analysis of the model results (e.g. modelled ozone diurnal cycles). With regard to the specific point about NO_2, we agree with the reviewer and have clarified this in the text throughout the new version of the paper.

With regard to aerosols, we have included a new section 4.3 which discusses reasons for the discrepancies in the model aerosol simulations compared to the observations and their implications for assessment of climate and air quality impacts.

With regard to the representativeness of the model/observations comparison we now explain, in several places in the text, which observations we have used, for which purpose and which spatial scales they represent.

For example, in the Introduction:

"In order to assess model performance over East Asia for air quality, as well as climate, we use a variety of different datasets covering the urban, regional, and continental scales. Ozone, aerosol and precursor data at surface sites in urban and rural locations are used, together with CAREBEIJING aircraft data collected in the lower troposphere south of Beijing, to evaluate model performance in terms of local and regional pollution from major emission regions. Continental scale horizontal and vertical transport of ozone and aerosols, important for radiative impacts, are assessed downwind of the main emission regions using aerosol lidar data as well as satellite aerosol lidar and tropospheric ozone, CO and NO_2 column data."

We agree that it is difficult to compare global models with surface data collected in urban locations, for example. On the other hand, it is important to assess these models over a variety of scales and locations ranging from urban to rural and on continental scales. Such models need to be able to capture average patterns in terms of gradients from polluted to background environments as well as vertical distributions of pollutants.

These models have been used in the ECLIPSE project to assess air quality and climate impacts together. Analysis of the CAREBEIJING aircraft data is described in Sections 3.5 and contributes to the discussion in Section 3.6. We make it clearer that this data was collected to the south of the main Beijing conurbation, in air masses largely originating from the south. See also our reply to the specific comment below.

Specific Comments

Abstract, l.6: "emissions of NOx are too high": or that chemical timescales or transport are incorrect. Addressing the contributions of these different biases is the key to strengthening the paper, and simultaneous evaluation of several variables across a number of models in this study should permit this.
Authors' reply: We agree with the reviewer that this particular conclusion was too general and required further analysis. The Abstract has now been substantially revised to better reflect the conclusions of our study and to take into the account the additional analysis of the trace gas results presented in the new Section 3.6 and the aerosol results presented in Section 4.3. See also replies above.

Abstract, l.13: I agree with the statement "These results have important implications for accurate prediction of pollution episodes...", but the reader needs to know what the implications are, and how they could be resolved.
Authors' reply: As noted above, the Abstract has been substantially revised. Overall, the ECLIPSE model evaluation highlights several significant biases in the model results, based on simulations using the same emissions. The Summary (Section 5) has also been revised which we invite the reviewer to read - the final paragraph from this section now states:
"In summary, the ECLIPSE model evaluation highlights significant differences between the models and observations, even when models are run using the same emissions over East Asia. Nevertheless, an important finding is that overall the Earth System Models show a similar level of performance as the Chemistry Transport Models, which is encouraging for the further use of the Earth System Models for determining both climate and air quality impacts. Somewhat better general agreement is found for trace gases compared to aerosols, for which model simulations are very variable. For both trace gases and aerosols, models have difficulties reproducing gradients between urban and rural~(downwind) locations and vertical distributions. Improved model resolution as well as improved understanding and model treatments of processes affecting pollutant lifetimes are needed. Model evaluations using a variety of observations are required so that difference aspects of model behavior can be tested. Results from this study suggest that significant uncertainties still exist in chemistry-climate simulations which has implications for the use of such models in the assessment of radiative effects of short-lived climate forcers on climate and regional/global air quality."

Section 2.4: It would be helpful to include a statement on the purpose of the CAREBeijing flights to provide an indication of how representative they were of the spatial and temporal scales resolved by the models. Were they intended to sample urban outflow, or rural regions? What biases have previous analysis of these flight data identified?

Authors' reply: The paper analysing the CAREBEIJING aircraft observations (Zhang et al., 2014), did not provide the intended target of the flights nor did they identify clear biases. However, a back-trajectory analysis conducted by Zhang et al. (2014) highlighted 4 different origins for the air masses sampled during the flights: south, north, east and mixing. More polluted air masses that were sampled originated from the south. These airborne observations are representative of the somewhat less polluted background atmosphere in China with only a partial influence from urban areas. Section 2.4 has been updated. We also added the following text describing how the model results were extracted and their representativeness:

"Model results were extracted along the flight paths corresponding to 2 or 3 model pixels (depending on the model) using hourly (or 3-hourly) output. This allows a fairer evaluation against the observations especially since trace gases have important diurnal cycles. Whilst the model results are an average over fairly large spatial scales, such a comparison provides useful insights into the vertical distribution of pollutants simulated by the models over a region which more representative of the less polluted background."

Section 2.5: The meteorological data are highlighted as a source of model differences, but no indication is provided of how good they are (how close are they to observations?) or what differences they introduce. Fig 1 shows that the NorESM fields are very different from the others. What can we learn from pollutant comparisons if the underlying transport patterns are substantially different from those observed?

Authors' reply: The models were forced, primarily using meteorological analyses and reanalyses from ECMWF, or NCEP in the case of WRF-Chem, in order to nudge the model results closer to reality in the case of the large scale dynamics and humidity, for example. These meteorological analyses have also been corrected with assimilated observations a posteriori. At this time of the year, the flow over the southern part of East Asia is influenced by the Asian summer monsoon with dominant synoptic winds blowing from the south linked to the Anticyclonic circulation over the Pacific Ocean to the east. This results in high levels of relative humidity over the region. ECMWF and NCEP wind fields are rather similar suggesting that differences in large-scale transport patterns are not the main cause of differences in trace gases and aerosols discussed in later sections.

As pointed out by the reviewer, NorESM was not nudged to meteorology, but forced using reanalysed sea surface temperatures for 2008. Figure 1 shows that this model does have a monsoon circulation, but that it penetrates to far to the north over China. NorESM was included in this evaluation to provide consistency with companion papers

(e.g Stohl et al., 2015) and because it was used to derive present-day and future emission impacts on air quality and climate. NorESM relative humidities are also higher over China, for example, compared to ECMWF and NCEP analyses which is likely to affect ozone photochemistry as well as aerosol formation. These factors, linked to model representation of the Asian monsoon, are discussed in the results sections.

Section 3.1.1: Are these spatial or temporal statistics/correlations? This should be stated clearly here, and in the captions to the relevant tables. I assume they are spatial statistics, in which case the start of the section (p.11060, l.6) should state clearly that it is the spatial patterns which are being evaluated here.
Authors' reply: These are spatial statistics/correlations. We have revised Section 3.1 and merged the original sub-sections to make the text clearer. The point about the use of IASI data to evaluate models on continental scales is covered in the introduction to Section 3 and elsewhere in the manuscript. See also the reply to the next comment. The Table captions have also be updated.

p.11061,l.14: What does the good agreement actually tell us here? Does it just reflect the latitudinal variation in tropopause height, or is there a more interesting story?
Authors' reply: We agree with reviewer#2 that the good agreement is actually reflecting the latitudinal variation of the tropopause height since the highest concentrations are observed and simulated in the north of the domain, where the tropopause height is lower with higher amounts of stratospheric ozone included in the 0-20 km column. But, the main finding here is the fact that the overestimation of the 0-6 km column ozone by the models is not due to a general overestimation in the stratospheric ozone flux in this region, since models show good agreement with 0-20 km IASI ozone columns (high correlation coefficients (R > 0.93). The main discrepancies reported in the modelled tropospheric columns are mostly linked to photochemistry and transport rather than stratospheric intrusions (as discussed in Sect. 3.6). Section 3.1 has been re-written to make this point clearer.

p.11062,l.26: observed "spatial" patterns?
Authors' reply: Yes, the text has been updated (see above).

p.11063,l.29: Biases are substantially worse at 0-6 km than at 0-20 km, as expected, as they are not dominated by stratospheric influences. What can you say about the source of the biases here? (e.g., overestimation over East China Sea, etc.)
Authors' reply: Reasons for model biases compared to IASI, and the other datasets used in the evaluation, are now discussed in detail in the new discussion Section 3.6.

Section 3.5: As noted in the general comments above, a clearer and more detailed discussion of the expected chemistry and transport biases in this section would make the paper much more useful.

Authors' reply: The section (3.6) discussing the results and causes for model discrepancies has been totally rewritten to include an analysis of possible causes for biases in the different models. We attempt to draw out some systematic biases in the results and to attribute these to particular issues in different models. Note that we have also included a more complete discussion of aerosol discrepancies in Section 4.3.

Section 4.2: How are the aerosol composition biases identified here likely to influence the optical properties assessed earlier? The large overestimation of BC is a particular issue. Comparing the performance in different variables at the same time should allow stronger conclusions to be drawn.

Authors' reply: In general, biases in the aerosol composition mostly concern BC in urban environment. This bias is expected to increase the AOD and Rapp values, although its impact is limited because of (i) the small BC contribution to the total aerosol mass, and (ii) the smaller contribution to extinction of a given BC particle compared to similar (in mass) OC and sulphate particles. More specifically, some models show important biases in very scattering species (OC and sulphate) which are correlated with excessive values of AOD and Rapp, e.g. EMEP and NorESM (over Korea only).

We have added a point about the BC/SO4 ratios in the discussion in Section 4.3:

"The BC/SO4 ratio observed at Beijing is almost constant (0.2) with enhanced values detected episodically (02-03 August, 16-17 August, 01-02 September), where the ratio can reach 1. All models reproduce this ratio reasonably well (not shown), but two (HadGEM and EMEP) show high oscillations between 0.1 and 6 (mean value of 2). Over Gosan, the observed BC/SO4 ratio is lower (0.1) underlining that Gosan is a more remote site from local sources. Models also present a good agreement, except EMEP (0.4) and TM4-ECPL (0.2), which overestimate the ratio. Such discrepancies may affect model responses to emission perturbations and thus radiative forcing."

We also highlight evidence for :

"a general underestimation of scattering aerosols in the boundary layer associated with overestimation in the free troposphere pointing to modeled aerosol lifetimes that are too long. This is likely linked to a too strong vertical transport and/or insufficient deposition efficiency during transport or export from the boundary layer, rather than chemical processing (in the case of sulphate aerosols). The top-heavy distribution of sulfate implies substantial errors in the simulated aerosol-cloud interactions, with too weak effects if not enough of the sulfate impacts the boundary-layer clouds."

p.11075,l.21: The short lifetimes aspect is important here (see general comment about Section 3.5), but you don't explain why this might be the case.

Author's reply: We agree with the reviewer that it is important to identify biases in modelled lifetimes of ozone precursors, such as CO. Our analysis suggests that this more likely linked to excessive ozone production than to transport errors. This is explained more clearly in the new version of the manuscript where we have included further discussion in Section 3.6:

"The ECLIPSE models also systematically underestimate CO downwind compared to surface data over Korea, Japan, and compared to IASI CO data over Japan and downwind over the north-western Pacific Ocean. Whilst inclusion of additional seasonality in the ECLIPSE emissions (already included for domestic combustion), might improve agreement in winter and spring (Stein et al., 2014), this is unlikely to explain these summertime differences. Low model CO appears to be linked to the clear overestimation in modeled ozone at rural sites and compared to IASI 0-6~km column data. Excessive ozone resulting in too much destruction of CO by OH and may suggest that modeled CO lifetimes are too short. This hypothesis is consistent with the findings of Monks et al. (2015) who concluded that, in models run with the same emissions, differences in OH~(chemical schemes) are a more likely cause of the systematic CO underestimation in the Northern Hemisphere and the Arctic than differences in vertical transport. Indeed, we find that surface August mean modeled OH (not shown) is higher in the NorESM model (due to the penetration of the monsoon flux) compared to, for example, TM4-EPCL and WRF-Chem over the main Chinese emission regions. In contrast, excessive modeled CO over the central Pacific, where concentrations are low, may be due to the position of the Pacific anticyclone in the meteorological analyses used by the majority of models. A shift in the position of the anticyclone to the south could result in this pattern of negative (positive) biases over the north (south) Pacific potentially as a result of transport that is too zonal. This may also explain low modeled CO in the Arctic noted by Monks et al. (2015)."

p.11075,l.25: Coarse resolution models are not appropriate tools for representing regional air quality, so this sentence simply supports the findings of many previous studies.

Authors' reply: We agree with reviewer#2 's comment and removed this sentence. We have also added the following sentence in the Introduction : "The ECLIPSE global chemistry-climate models may not be the most suitable tools to assess air quality impacts, however they are the tools used to evaluate climate and air quality impacts together. To address this point, a regional model is also included in the evaluation and one of the global chemical-transport models is run at relatively high horizontal resolution (50 km) compared the other global models."

Based on results from this study, we have however included in the summary section :

"This evaluation has important implications for accurate assessment of air pollutants on regional air quality and global climate based on global model calculations. Overall, Earth System Models perform as well as Chemical Transport Models in terms of atmospheric composition, which is encouraging for the further use of the Earth System Models in determining the climate impacts. Ideally, models should be run at higher resolution over source regions to better simulate urban-rural pollutant gradients/chemical regimes, and also to better resolve pollutant processing and loss by wet deposition as well as vertical transport. Discrepancies in vertical distributions require further quantification since these are a key factor determining estimated radiative forcing from short-lived pollutants."

p.11076: The final paragraph of the conclusions makes a number of very vague and general statements about model weaknesses and about improvements needed, but there are no specific new conclusions derived from this study. Please sharpen up this paragraph to include some firm guidance on how weaknesses should be addressed.
Authors' reply: Conclusions have been rewritten accordingly with modifications in Sections 3.6 (trace gases) and 4.3 (aerosols). We have substantially rewritten the paper providing additional scientific analysis that was, as the reviewer points out, missing from the submitted version of the manuscript.

Table 3: The NMB of the model mean is a lot larger than that of any of the contributing models. This looks odd, please check.
Authors' reply: We thank the reviewer for pointing this. The correct value is 2.5, not 25.0. It has been replaced in Table 3.

Typos
p.11052, l.11/l.26: "Hong Kong"
p.11052, l.21: missing "the" before industrial
p.11055,l.5-6: "allow to control" would be clearer as "address"
p.11056, l.18: grammar incorrect: perhaps remove "description"?p.11067, l.6: Monks reference should be 2015.
p.11075, l.13: "climate" not needed.

Authors' reply: All the typos have been corrected.

Multi-model evaluation of short-lived pollutant distributions over East Asia during summer 2008

B. Quennehen[1,*], J.-C. Raut[1], K. S. Law[1], N. Daskalakis[2,3,**], G. Ancellet[1], C. Clerbaux[1], S.-W. Kim[4], M. T. Lund[5], G. Myhre[5], D. J. L. Olivié[6], S. Safieddine[1,***], R. B. Skeie[5], J. L. Thomas[1], S. Tsyro[6], A. Bazureau[1], N. Bellouin[7], M. Hu[8], M. Kanakidou[2], Z. Klimont[9], K. Kupiainen[9], S. Myriokefalitakis[2,3], J. Quaas[10], S. T. Rumbold[11,****], M. Schulz[6], R. Cherian[10], A. Shimizu[12], J. Wang[8], S.-C. Yoon[4], and T. Zhu[8]

[1]LATMOS/IPSL, UPMC Univ. Paris 06 Sorbonne Universités, UVSQ, CNRS, Paris, France
[2]Environmental Chemical Processes Laboratory, Department of Chemistry, University of Crete, Heraklion, Crete, Greece
[3]Institute of Chemical Engineering Sciences (ICE-HT), FORTH, Patra, Greece
[4]School of Earth and Environmental Sciences, Seoul National University, Seoul, Korea
[5]Center for International Climate and Environmental Research - Oslo (CICERO), Oslo, Norway
[6]Norwegian Meteorological Institute, Oslo, Norway
[7]Department of Meteorology, University of Reading, Reading, RG6 6BB, UK
[8]College of Environmental Sciences and Engineering, Peking University, Beijing, China
[9]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[10]Institute for Meteorology, Universitat Leipzig, Leipzig, Germany
[11]Met Office Hadley Centre, Exeter, EX1 3PB, UK
[12]National Institute for Environmental Studies, Ibaraki, Japan
[*]now at Univ. Grenoble Alpes/CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), 38041 Grenoble, France
[**]now at LATMOS/IPSL, UPMC Univ. Paris 06 Sorbonne Universités, UVSQ, CNRS, Paris, France
[***]now at Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States
[****]National Centre for Atmospheric Science, University of Reading, Reading, RG66BB, UK

Correspondence to: Boris Quennehen (quennehen@latmos.ipsl.fr)

Abstract.

The ability of ~~six global and one regional model~~ seven state of the art chemistry-aerosol models to reproduce distributions of tropospheric ozone and its precursors, as well as aerosols over eastern Asia in summer 2008 is evaluated~~using satellite and~~. The study focuses on the performance of models used to assess impacts of pollutants on climate and air quality as part of the EU ECLIPSE project. Models, run using the same ECLIPSE emissions, are compared over different spatial scales to in-situ ~~observations. Whilst ozone precursors (NO_2 and CO) are generally underestimated by the models in the troposphere, surface NO_2 concentrations are overestimated, suggesting that emissionsof NO_x are too high. Ozone integrated columns and vertical profiles are generally well modeled, but the global models face difficulties simulating the ozone gradient at the surface between urban and rural environments, pointing to the need to increase model resolution. The accuracy of simulated aerosol patterns over eastern China and northern India varies between the models,~~ surface, vertical profile and satellite data. Several rather clear biases are found between model results and observations including overestimation of ozone at rural locations downwind

of the main emission regions in China as well as downwind over the Pacific. Several models produce too much ozone over polluted regions which is then transported downwind. Analysis points to different factors related to the ability of models to simulate VOC limited regimes over polluted regions and NO_x limited regimes downwind. They may also be linked to biases compared to satellite NO_2 indicating overestimation of NO_2 over and ~~although most of the~~ to the north of the northern China Plain emission region. On the other hand, model NO_2 is too low to the south and east of this region and over Korean/Japan. Overestimation of ozone is linked to systematic underestimation of CO particularly at rural sites and downwind of the main Chinese emission regions. This is likely to be due to enhanced destruction of CO by OH. Model-observation discrepancies over Beijing do not appear to be due to emission controls linked to the Olympic Games in summer 2008.

With regard to aerosols, most models reproduce the ~~observed pollution features~~ satellite-derived AOD patterns over eastern China~~, significant biases are noted in the magnitude of optical properties (aerosol optical depth,~~. Our study nevertheless reveals an overestimation of ECLIPSE model-mean surface BC and sulphate aerosols in urban China in summer 2008. The effect of the short-term emission mitigation in Beijing is too weak to explain the divergences between the models. Our results rather point to an overestimation of SO_2 emissions, in particular, close to the surface in Chinese urban areas. However, we also identify a clear underestimation of aerosol ~~backscatter). These results have~~ concentrations over northern India, suggesting that the rapid recent growth of emissions in India, as well as their spatial extension, is underestimated in emission inventories. Model deficiencies in the representation of pollution accumulation due to the Indian monsoon may also be playing a role. Comparison with vertical aerosol lidar measurements highlights a general underestimation of scattering aerosols in the boundary layer associated with overestimation in the free troposphere pointing to modeled aerosol lifetimes that are too long. This is likely linked to a too strong vertical transport and/or insufficient deposition efficiency during transport or export from the boundary layer, rather than chemical processing (in the case of sulphate aerosols). The top-heavy distribution of sulfate implies substantial errors in the simulated aerosol-cloud interactions, with too weak effects if not enough of the sulfate impacts the boundary-layer clouds.

This evaluation has important implications for accurate ~~prediction of pollution episodes affecting air quality and the radiative effects of these~~ assessment of air pollutants on regional air quality and global climate based on global model calculations. Overall, Earth System Models perform as well as Chemical Transport Models in terms of atmospheric composition, which is encouraging for the further use of the Earth System Models in determining these impacts. Ideally, models should be run at higher resolution over source regions to better simulate urban-rural pollutant gradients/chemical regimes, and also to better resolve pollutant processing and loss by wet deposition as well as vertical transport. Discrepancies in vertical distributions require further quantification since these are a key factor determining estimated radiative forcing from short-lived ~~climate pollutantsover Asia.~~ pollutants.

1 Introduction

Short-lived pollutants (SLPs), defined here as tropospheric ozone and aerosols, including black carbon (BC)~~were recently~~, are the focus of several important efforts by the scientific community due to their potential role in emerging strategies aiming

to mitigate global climate change and improve air quality (Shindell et al., 2012; Anenberg et al., 2012). Due to their relatively short lifetimes (e.g., aerosol lifetime in the troposphere is about one week (Pruppacher and Klett, 1997)), the impact of ~~SLP~~SLPs (as well as ozone precursors) emission reductions on near-term reductions in the rate of climate warming has been examined in several recent studies (Ramanathan and Carmichael, 2008; Jackson, 2009; Penner et al., 2010; Shoemaker et al., 2013; Smith and Mizrahi, 2013; Rogelj et al., 2014).~~In particular, due to recent fast increases in SLP and ozone precursor emissions (Streets et al., 2003; Richter et al., 2005; Klimont et al., 2013; Wang et al., 2014; Klimont et al., 2016), East Asia is a key region being targeted by mitigation strategies. This region is already highly polluted with elevated near surface pollutant concentrations, especially over eastern China (Wang et al., 2011b; Yang et al., 2011), producing severe environmental impacts (Ma et al., 20 Shao et al. (2006) presented observations from 360 cities in China showing that the national ambient air quality limits were exceeded in 70% of the cases. More specifically, three areas in eastern China are shown, by several studies, to be major polluted regions (Chan and Yao, 2008; Ma et al., 2012; Boynard et al., 2014): the North China Plain (NCP, which includes Beijing and Tianjin), the Yangtze river delta (YRD, which includes Shanghai), and the Pearl river delta (PRD, which includes Honk-Kong and Guangzhou). It is therefore important to assess the source of this regional pollution and to evaluate the ability of chemistry-aerosol-climate models to simulate the distributions of SLPs over this region.~~

Ozone is a reactive species impacting both climate and air quality. In the troposphere, it is produced photochemically from the oxidation of carbon monoxide (CO), ~~methane,~~ non-methane volatile organic compounds (VOCs) by OH radicals in the presence of nitrogen oxides (~~NO_x~~). ~~It also~~ NO_x). Methane is also an important ozone precursor. Tropospheric ozone also has natural sources such as the flux from the stratosphere. Due to photochemical loss, it has a lifetime in the lower troposphere of a few weeks (Stevenson et al., 2006). It is also removed by dry deposition to the surface. Radiative forcing due to tropospheric ozone over the industrial era is estimated to be 0.40 ± 0.20 ~~W m^{-2}~~ W m^{-2} (Myhre et al., 2013b). ~~Tropospheric ozone pollution over the three most polluted Chinese regions is influenced by the Asian summer monsoon (Liu et al., 2002; Xu et al., 2008; Lin et al., 2009; He e regional photochemical pollution (Lin et al., 2008). Long-term time series of ozone measurements were reported for Hong-Kong (Wang et a Beijing (Ding et al., 2008), showing increasing background concentrations at both locations, whereas seasonal and day-to-day variations over the NCP, YRD, and PRD regions in 2008 were studied by Dufour et al. (2010) using satellite observations. A seasonal maximum is found over East Asia in late spring due to increased solar radiation whereas a minimum is found in summer for the PRD and YRD, and in winter for the NCP (Naja and Akimoto, 2004; Li et al., 2007; He et al., 2008; Safieddine et al., 2013) The main mechanism driving the summer minimum is the intrusion of low-ozone marine air masses coming from the tropical Pacific ocean, i.e., the monsoon (Lin et al., 2009).Several regional model evaluations of trace gases have been carried out with a focus on East Asia. In the frame of the MICS-Asia phase II project, Han et al. (2008) showed good agreement between eight regional models and SO_2 observations from the TRACE-P campaign (Fuelberg et al., 2003), whereas modeled ozone accuracy varied with the region and the season. Model performance regarding NO_x was relatively poor with a systematic underestimation by all the models. A previous study by Zhang et al. (2006), comparing TRACE-P airborne observations with a regional model simulation, found similar results as Han et al. (2008) but with a more accurate reproduction of NO_x concentrations. CO concentrations were underestimated by Gao et al. (2009) at two surface stations in Japan whereas ozone~~

seasonal variation and concentration simulated were in agreement with the observations.Atmospheric aerosol plays a major role in the Earth's radiative balance by scattering (McCormick and Ludwig, 1967) and absorbing solar radiation (Haywood and Shine, 1995). Aerosols also affect the formation, lifetime and albedo of clouds (Albrecht, 1989; Twomey, 1977; Ackerman et al., 2000), causing indirect effects on the radiative balance. According to recent estimates, atmospheric aerosols , emitted by anthropogenic and natural sources (e.g., heating, transportation, biomass burning and dust), have, since pre-industrial times, modified the aerosol direct effect by -0.35 $\pm$ 0.50 W m^{-2} (Myhre et al., 2013b) W m^{-2}, whereas the total (direct and indirect) effects, which include cloud adjustments due to aerosols, modified the Earth's radiative balance by -1.3 $\pm$ 0.4 W m^{-2} (Cherian et al., 2014). Radiative forcing due to absorbing species (e.g., BC) at the top-of-atmosphere is shown to be positive -0.9 (from -1.9 to -0.1) W m^{-2} (Myhre et al., 2013b) and result from a negative forcing from most aerosols and a positive contribution from BC absorption of solar radiation (Haywood and Shine, 1995). BC is the carbonaceous component of soot, resulting from incomplete combustion. In a recent extensive study, Bond et al. (2013) estimated the BC direct radiative forcing of BC from fossil fuel and biofuel emissions for the industrial era to be 0.51 W m^{-2} W m^{-2} whereas Myhre et al. (2013b) reported a positive forcing of 0.40 W m^{-2} W m^{-2} (0.05-0.80). BC emissions are almost entirely anthropogenic and 90% of BC emissions are due to diesel engines, industry, residential burning, and open burning (Bond et al., 2013). Aerosol impacts on air quality are also a serious problem since aerosols reduce visibility (sometimes dramatically, e.g., Zhao et al. (2011); Chen et al. (2012); Li et al. (2013)) and cause, together with ozone, serious health deterioration and premature deaths (Nawahda et al., 2012).

Many studies have discussed observations of aerosol properties over East Asia (e.g., Kim et al. (2007)) and, more specifically, over the three regions mentioned above. A clear seasonal cycle is observed for organic carbon (OC) and BC with maximum and minimum concentrations in winter and summer, respectively (Cao et al., 2004; Sun et al., 2004; Yang et al., 2005; Huang et al., 2006).The ability of global models to simulate monthly aerosol optical depths (AODs) was examined by Kinne et al. (2006) over East Asia in summer with comparisons showing important underestimation of observed AODs. Kinne et al. (2006) also pointed out that uncertainties in the direct radiative forcing could be larger than differences in AOD suggest. The aerosol vertical distribution was investigated by Koffi et al. (2012) who compared global model results with aerosol extinction mean profiles from satellite observations over specific regions, including East Asia. They found generally good agreement in the vertical shape of extinction profiles between the models and the observations. However, accuracy depended highly on the model, season and region. Large discrepancies in modeled surface BC concentrations compared to observations were reported by Koch et al. (2009) whereas Samset et al. (2013) and Allen and Landuyt (2014) concluded that BC vertical distributions are a major source of uncertainty in BC radiative forcing estimations.

East Asia is a key region being targeted by SLP mitigation strategies due to the recent rapid increases in precursor emissions (Streets et al., regional and global radiative forcing, severe episodes of air pollution and other environmental impacts (Ma et al., 2010). National ambient air quality exceedances occur in many cities (Shao et al., 2006) especially in eastern China (Wang et al., 2011b; Yang et al the North China Plain (NCP, including Beijing and Tianjin), the Yangtze river delta (YRD, including Shanghai) and the Pearl river delta (PRD, including Hong-Kong and Guangzhou). In this context, the European Union (EU) Evaluating the CLimate and air quality ImPacts of Short-livEd pollutants (ECLIPSE) project was launched with the objective to develop and assess

effective emission abatement strategies in order to provide sound scientific advice about measures to mitigate climate change and improve air quality . ECLIPSE models versions are more recent than those evaluated in the cited evaluation studies, thus one focus of the project is the evaluation of global and regional chemical-aerosol-climate models over two important emission regions (Europeand Asia) developed new emission inventories for present-day global SLP emissions as well as future scenarios designed to benefit both air quality and climate with a focus on Asia and Europe (see Stohl et al. (2015) for discussion of the ECLIPSE rationale and summary of results). A new ECLIPSE emission inventory was developed for methane, aerosols, ozone and their precursors, including, in particular, improvements over China. Several sources, such as brick making kilns, were updated, and previously unaccounted sources such as wick lamps, diesel generators and high-emitting vehicles, were included. These emissions were used to perform a detailed analysis of climate metrics for different emission sectors, regions (including China, India) and seasons using state of the art Earth System Models (ESMs). The results were used as a basis for refining a mitigation scenario by including additional measures with beneficial air quality and one receptor region (the Arctic). The aim of this paper is to evaluate the performance of six global and one regional model, run using the same ECLIPSE emission dataset, by comparison against observations over East Asia. Diagnoses used in this study allow to control the driving processes affecting both air quality short-term (e. g., surface concentration)20-year) climate impacts. Compared to the ECLIPSE current legislation scenario, taking into account current and planned legislation for emission reduction, the ECLIPSE mitigation scenario, taking into account these additional measures (e.g. gas flaring, diesel engines, and coal/biomass stoves) would reduce global anthropogenic methane and BC emissions by 50 and climate80% respectively by 2050. It is estimated that, in the decade 2041-2050, the mitigation scenario would result in 0.2 K less surface temperature warming globally (Stohl et al., 2015) and, at the same time extend, for example, life expectancy in China by 1.8 months in 2030.

An important component of ECLIPSE is the so-called reality check to evaluate model performance over pollutant source (Europe (Stohl et al., 2015), China/Asia , the focus of this paper) and receptor (Arctic (Eckhardt et al., 2015)) regions. In these evaluations, the ECLIPSE models were run with the same present-day ECLIPSE emission inventory (ECLIPSEv4a) for 2008 and 2009. Note that the same global models were used to estimate sector/regional emission responses and, in a sub-set of cases, to predict, using the ECLIPSE emission scenarios, future atmospheric composition and associated impacts on climate and air quality. The ECLIPSE global chemistry-climate models may not be the most suitable tools to assess air quality impacts, however they are the tools used to evaluate climate and air quality impacts together. To address this point, a regional model is also included in the evaluation, and one of the global chemical-transport models is run at relatively high horizontal resolution (50 km) compared to the other global models. The ECLIPSE evaluation over Europe showed that many models underestimate CO, and overestimate ozone, whilst modeled AOD was reproduced reasonably well (Stohl et al. (2015)). Over the Arctic, models often underestimate both BC and sulfate aerosols due to problems with emissions (e.g. , vertical profiles of optical properties). The East asia region is a focus because, as fires), vertical redistribution, transport and loss processes such as wet deposition. Here, we present results from the evaluation of the ECLIPSE models over East Asia. As noted above, it is a region with this region was targeted due to its still high pollution levelswhere emissions still appear to be increasing but where , climate impacts and as a region where SLP mitigation options are also being actively considered. East Asian emissions originate

from several different sources including anthropogenic emissions, biomass burning, and dust episodes. In addition, pollution from Asia can be transported many thousand kilometers downwind to North America (e. g., (Jaffe et al., 1999)) as well as to the Arctic (e.g., Roiger et al. (2011)). The focus of this study is on summer It is also a region where significant uncertainties surround model estimates of radiative forcing. For example, Kinne et al. (2006) showed important underestimation of observed AOD by multiple models over East Asia in summer and pointed out that uncertainties in the direct radiative forcing could be larger than inter-model differences in AOD suggest. Even in the recent AeroCom model comparison, inter-model variation in radiative forcing is largest in this region (Myhre et al., 2013a). Samset et al. (2014) suggested that BC direct radiative forcing is overestimated by about 25% downwind of Asian emissions in the upper troposphere over the Pacific based on overestimation of modeled BC compared to aircraft observations.

The ECLIPSE model evaluation over East Asia focuses on the summer period (August and September) 2008; a period for which the). This was motivated by the availability of intensive observations from the CAREBEIJING 2008 measurement measurements campaign (Huang et al., 2010; Zhang et al., 2014) are available. The CAREBEIJING 2008 campaign took place in August and September 2008 in the context of the Beijing Olympic and Paralympic games. SLPs emissions mitigation was attempted between June 30^{th} to September 20^{th} 2008 (see the detailed mitigation plan in Wang et al. (2010)) and concerned polluting vehicles, and by the fact that severe ozone pollution episodes occur over NCP at this time of year even if the maximum is generally earlier in the late spring for trace gases (Naja and Akimoto, 2004; Li et al., 2007; He et al., 2008; Safieddine et al., 2013) and aerosols (Cao et al., 2004; Sun et al., 2004; Yang et al., 2005; Huang et al., 2006). During the summer months, whilst the monsoon circulation brings cleaner air from the Pacific Ocean into southern and eastern Asia reducing pollutant concentrations (Lin et al., 2009), (Kim et al., 2007). However, high pollution episodes with enhanced aerosol concentrations and decreasing visibility still occur in coastal regions, due to increases in relative humidity increasing aerosol sizes (Flowers et al., 2010). The monsoon flux also induces transport of high ozone concentrations inland (He et al., 2008). In order to assess model performance over East Asia for air quality, as well as the chemical and power plants within the Beijing province. The impacts of this pollutant emission mitigation have been reported in several studies. Wang and Xie (2009) and Zhou et al. (2010) observed reduction of 19 to 57% in CO and 28 to 52% in PM10 on-road emissions, whereas Wang et al. (2009b) reported decrease of 21% in summer climate, we use a variety of different datasets covering the urban, regional, and continental scales. Ozone, aerosol and precursor data at surface sites in urban and rural locations are used, together with CAREBEIJING aircraft data collected in the lower troposphere south of Beijing, to evaluate model performance in terms of local and regional pollution from major emission regions. Continental scale horizontal and vertical transport of ozone and aerosols, important for radiative impacts, are assessed downwind of the main emission regions using aerosol lidar data as well as satellite aerosol lidar and tropospheric ozone, CO and NO_2 column data. Emissions (polluting vehicles, chemical, power plants) in the Beijing province were mitigated from 30 June 2008 and 20 September 2008CO observation compared to 2006 and 2007 at a remote site 100 km away from Beijing and concluded, based on a model analysis, that ozone concentrations were reduced by 2-10 ppbv over the NCP region during the mitigation period. At a larger scale, Worden et al. (2012) used satellite observations to deduce a 11% reduction in CO emissions over Beijing. Although lower pollutant concentrations can be observed, particularly in the south and (see

the detailed mitigation plan in Wang et al. (2010)) in the context of the Beijing Olympic and Paralympic games. We examine the effects of these emission reductions on atmospheric composition using a regional model in ~~coastal regions due to the monsoon influences at this time of year (Kim et al., 2007), high pollution episodes still occur (Flowers et al., 2010). Improved estimation of the radiative impact of SLPs requires careful evaluation of model performance, in particular with respect to the vertical distribution of ozone, aerosols and their precursors. This study focuses on evaluating models at the surface and in the free troposphere, where pollutant distributions are important for air quality and climate, respectively~~order to assess potential influences on the model results compared to data collected in the Beijing region.

~~The models used in this study~~

The emissions, models and datasets used to assess model performance are described in Sect. ~~2~~ 2 as well as the meteorological situation during summer 2008. Evaluation of simulated ~~trace gas distributions on different temporal and spatial~~ ozone and its precursors on local, regional and continental scales are presented in Sect. ~~3.~~ 3. This includes comparison with the Infrared Atmospheric Sounding Interferometer (IASI) CO and ozone data, the Global Ozone Monitoring Experiment (GOME) NO_2 data and surface/aircraft data collected in the vicinity of Beijing and surface trace gas data collected at downwind sites in Korea and Japan. Comparison of modeled and observed trace gas correlations are used to draw conclusions about whether model discrepancies are due to emissions, chemical processing (VOC or NO_x limited ozone production) and/or transport. Comparisons between observed and modeled aerosol optical properties, as well as available surface/aircraft data on aerosol chemical composition, are discussed in Sect. ~~4. The summary and conclusions~~ 4. This includes comparison with Moderate resolution Imaging Spectrometer (MODIS) AOD, Cloud-Aerosol LIdar with Orthogonal Polarisation (CALIOP) as well as ground-based attenuated backscatter lidar profiles, and aerosol surface composition. Conclusions are given in Sect. ~~5.~~ 5.

2 Models, evaluation datasets and meteorological conditions

In this section, the global and regional models involved in this study are presented together with the different measurement datasets used to evaluate their performance including satellite data, ground-based, and airborne measurements. The meteorological conditions over East Asia during summer 2008 are also ~~briefly described~~discussed, including possible biases in model transport patterns.

2.1 ~~Models description~~Model descriptions and emission dataset

The main model characteristics are listed in Table 1. ~~Within the ECLIPSE project, a dedicated anthropogenic emission dataset was developed (ECLIPSE V4, Klimont et al. (2016)). Whilst ECLIPSE~~ Models were run with ECLIPSEv4a present-day anthropogenic emissions, including agricultural waste burning, for the year 2008 (Klimont et al., 2016). Whilst ECLIPSEv4a emissions are annual averages for most of the sectors, a seasonal cycle is applied to the domestic sector (Streets et al., 2003). ~~In addition, agricultural waste burning fires are included in the inventory.~~ Emission reductions associated with the mitigation strategies during the Olympic period mentioned in the ~~Introduction~~ Sect. 1 are not taken into account in the ECLIPSE anthro-

pogenic emissions. Wildfire emissions were taken from GFED 3.1 (van der Werf et al., 2010) and aircraft/shipping emissions were from the RCP 6.0 scenario (Lee et al. (2009) and Buhaug et al. (2009), respectively~~)~~ , whereas biogenic emissions were prescribed individually by each model (~~see~~ Table 1). Dust, sea salt and ~~DMS~~ dimethyl sulfide (DMS) emissions were also model-dependent. WRF-Chem provides online dust and sea-salt emissions but only the latter are used in the ECLIPSE simulations due to an overestimation of dust loads, as reported by Saide et al. (2012). The main dust sources in East Asia are located in dry regions of China and Mongolia, north of the Himalayas (Taklamakan, Gobi and Gurbantunggut deserts). Most of the dust events occur in spring (Huang et al., 2013) whilst in summer, due to the Asian summer monsoon flux, rather little dust is transported to coastal areas (Kim et al., 2007). Thus, neglecting this source in WRF-Chem summertime simulations is not expected to introduce large bias in ~~the~~ modeled aerosol loads. Global model simulations were conducted for 2008 with a one or two years spin-up (depending on the model) whereas the regional WRF-Chem simulation was for August and September 2008 with a 10 day spin-up using initial chemical and boundary conditions from the ~~Mozart model.~~MOZART-4 model (Emmons et al., 2010). Models were run over a range of horizontal and vertical resolutions ranging from around 50 km to 250 km and with 26 to 60 vertical levels.

2.2 Satellite observations

Several satellite datasets have been used in this evaluation since they provide useful information about ~~the spatial distribution~~ continental scale spatial distributions of pollutants and their precursors. The ~~Infrared Atmospheric Sounding Interferometer (IASI)~~ IASI sensor mounted onboard the MetOp-A platform has provided data since June 2007. It is a nadir-looking Fourier transform spectrometer working in the thermal infrared spectral range (645-2760 ~~cm^{-1}~~cm^{-1}) (Clerbaux et al., 2009) ~~. It~~ that can detect several trace gases including ozone and CO. The MetOp-A orbit is sun-synchronous and provides complete observation of the Earth's surface every day. However, clouds may affect the signal and lead to errors in the retrieved data. The software used for the retrieval of ~~the ozone global distribution~~ CO and ozone global distributions is the Fast Optimal Retrievals on Layers for IASI (~~FORLI-O3~~FORLI, Hurtmans et al. (2012)). ~~The second Global Ozone Monitoring Experiment (~~GOME-2 ~~,~~ (Munro et al. (2000)), also onboard the Metop-A satellite, is a nadir-looking spectrometer covering the spectral range between 240 and 790 nm at 0.2~~–0.4~~ –0.4 nm resolution. With its large swath of 1920 km, GOME-2 provides near global daily coverage. The GOME-2 sensor uses the Differential Optical Absorption Spectroscopy (DOAS) technique to observe the atmosphere~~and tropospheric NO_2~~ , and tropospheric NO_2 concentrations are retrieved using the algorithm developed by Boersma et al. (2004). AOD space-borne observations are collected by the ~~Moderate resolution Imaging Spectrometer (MODIS) instrument onboard 2~~ MODIS instrument onboard two satellites, Aqua and Terra, flying opposing orbits, providing global coverage of the Earth every 1-2 days. The MODIS level 3 products that are used in this study are described by Hubanks et al. (2008). Vertical distributions of aerosols and clouds and are probed with the ~~Cloud-Aerosol LIdar with orthogonal Polarisation (CALIOP)~~ CALIOP instrument mounted on the CALIPSO satellite, part of the A-train satellite constellation. CALIOP is a two wavelength (532 and 1064 nm) polarisation-sensitive lidar as described by Winker et al. (2007).

2.3 Ground-based data

~~The locations of the measurement stations~~

Surface data collected at urban, rural and remote sites in China, Japan and Korea was used in this study from SNU/EANET (Seoul National University / Acid Deposition Monitoring Network in East Asia) and Peking University (PKU) stations. Site locations are shown in Fig. 1 ~~, whereas the~~ and station coordinates are given in Table 2. ~~During~~ For example, during the CAREBEIJING 2008 campaign, SLP concentrations were measured at the air quality observatory of ~~Peking University (PKU)~~ PKU in Beijing which can be considered as a typical urban environment. Instrumentation deployed at the observatory measured ozone, ~~NO_x~~ NO_x (defined as the sum of NO and ~~NO_2~~ NO_2), and CO (Chou et al., 2011), as well as ~~PMs (PM~~particulate matter (PM: PM 2.5 and PM 10), ~~OC~~ organic carbon (OC) and BC (Huang et al., 2010). The Gosan observatory (Kim et al., 2005) is a long-term observatory located on the Jeju island, South Korea, measuring OC, BC, aerosol number size distributions (Flowers et al., 2010), and ~~NO_x, SO_2~~ NO_x, SO_2, CO and ozone concentrations. It is not influenced by local pollutant emissions and ~~experiences continental dust and/or pollution outflow from continental~~ samples air masses transported downwind of continental Asia (e.g., Kim et al. (2005)). ~~Aerosol~~ Models were also compared to vertical aerosol backscatter signals measured by the Japanese National Institute for Environmental Studies (NIES) ground-based lidar network (Shimizu et al., 2004; Sugimoto et al., 2008) ~~are also used to evaluate model results. The spatial resolution of this network covers~~ covering Japan with 10 lidars calibrated using a similar procedure. Backscatter data are available on-line on a hourly basis, allowing a robust validation of the models.

2.4 Airborne ~~dataset~~data

As part of the CAREBEIJING 2008 campaign, 12 scientific flights were performed over the area south of Beijing in the Hebei province. Flights followed linear routes at altitudes in the range 500-2100 m in order to sample both the boundary layer and the free troposphere. The instrumentation on board the aircraft is described by Zhang et al. (2014) and includes ozone, CO, ~~SO_2 and NO_x~~ SO_2 and NO_x samplers. The flight tracks are shown in blue in Fig. 1. One goal of the CAREBEIJING 2008 campaign was to examine the effects of additional local emission mitigation from June to September 2008 in Beijing province (Wang et al., 2010) which appear to have been significant locally. It also included surface measurements, showing, for example, Wang and Xie (2009) and Zhou et al. (2010) observed reductions of 19 to 57% in CO and 28 to 52% in PM10 on-road emissions in Beijing, whereas Wang et al. (2009b) reported a decrease of 21% in summer 2008 CO observations, compared to 2006 and 2007, at a site 100 km from the centre of Beijing and concluded, based on a model analysis, that ozone concentrations were reduced by 2-10 ppbv over the NCP region during the mitigation period. In contrast, Worden et al. (2012) deduced only a 11% reduction in CO emissions over Beijing based on analysis of satellite data. We address this point by using WRF-Chem model to run a sensitivity test with lower emissions (Sect. 3.6).

2.5 Meteorological context

~~During summer 2008, East Asia was mostly influenced by the Asian summer monsoon with dominant synoptic winds blowing from the south-eastern Pacific Ocean, as seen in Fig. 1. These winds transport clean marine air to the Chinese coast and thus contribute to reducing pollutant concentrations. He et al. (2008) presented seasonal distributions of ozone and the monsoon index over East Asia, showing how the impact of the monsoon decreases gradually from coastal polluted regions to inland areas and induces transport of high ozone concentrations inland. However, the monsoon also increases relative humidity in coastal regions increasing aerosol sizes and decreasing visibility. The representation of the monsoon transport varies between the models used in this study.This is illustrated in Fig. 1 which highlights the differences between the model meteorological fields, especially over the NCP region where~

The majority of ECLIPSE models were driven or nudged using various meteorological analyses from ECMWF (European center for medium-range weather forecast) ~~and~~ with only WRF-Chem being nudged using NCEP (National ~~centers for environmental predictions~~Centers for Environmental Predictions) FNL (final) fields~~show a lower relative humidity (RH) than NorESM. Also, NorESM predicts strong~~ . To illustrate average transport patterns during August and September 2008, we show surface relative humidity and winds over the region in Fig. 1. At this time of the year, the flow over the southern part of East Asia is influenced by the Asian summer monsoon with dominant synoptic winds blowing from the ~~west over the continent in contrast to the~~ south-east linked to the anticyclonic circulation over the Pacific Ocean to the east. This also leads to high relative humidity over the southern part of the region. ECMWF and NCEP ~~analyses used to force other models. The transport of pollution is partly influenced by the meteorological parameters (e.g., winds)and differences between them can partially explain the variability in the model results.~~wind fields are rather similar suggesting that differences in large-scale transport patterns are not the main cause of differences in trace gases and aerosols discussed in later sections. The NorESM climate model is an exception since it was forced with sea surface temperatures for 2008 rather than nudged to meteorology. It was included in this evaluation in order to provide consistency with companion studies where this model was used to estimate present-day and future emission impacts on air quality and climate (Stohl et al., 2015, e.g.). This model has a monsoon circulation which penetrates further over East Asia compared to the ECMWF and NCEP analyses resulting in higher surface relative humidities over this region. This may affect ozone photochemistry as well as aerosol formation (Sect. 3.6 and Sect. 4.3).

3 ~~Evaluation~~ Interpretation of differences in modeled trace gas distributions

In this section, ~~model results are first compared with the 0-20 km and~~ modeled ozone and precursors (CO, NO_2) are evaluated at local, regional and continental scales to examine model performance on scales relevant for regional air quality and regional/global climate impacts. Firstly, large-scale spatial distributions of modeled ozone, CO and NO_2 are compared to IASI and GOME-2 satellite data. Lower tropospheric ozone is evaluated against 0-6 km ~~ozone columns observed by IASI , as well as NO_2 tropospheric columns detected by~~ IASI columns and 0-20 km are used to assess whether differences in downward transport from the stratosphere could be influencing modeled ozone over East Asia. IASI CO and GOME-2 NO_2 are used to evaluate

performance over and downwind of emission regions. Secondly, ~~the ability of models to simulate~~ to evaluate modeled ozone and its precursors ~~(NO_x and CO) reported by ground-based measurements at~~ on a regional scale, results are compared to surface measurements from various Chinese and Korean stations ~~is evaluated. Thirdly, in order to examinemodeled trace gas vertical distributions, comparisons with aircraft observations from the CAREBEIJING campaign are presented.~~ as well as vertical distributions observed by aircraft during CAREBEIJING in the lower troposphere. Further analysis of trace gas ratios and ozone diurnal cycles is used to provide insights into whether modeled discrepancies are due to deficiencies in emissions, photochemical processing or transport in the models (Sect. 3.6). We also examine, using one model (WRF-Chem), the potential impact of emission reductions over Beijing during the study period.

3.1 ~~Ozone~~ IASI ozone columns

~~Day and night~~

Day and night-time observations of IASI ozone are used to evaluate the models. Due to the variation of the IASI sensor's sensitivity with altitude, modeled ozone values need to be smoothed using the following equation:

$$X_{\text{smooth}} = AK * X_{\text{model}} + (I - AK) * X_{\text{apriori}} \quad (1)$$

where AK is the averaging ~~kernels~~ kernel matrix, I is the identity matrix and X_{smooth}, X_{model}, and X_{apriori} are the smoothed, modeled, and a priori ozone profiles, respectively. AK and X_{apriori} are obtained when inverting the measured signal. The 0-20 km column is ~~then~~ retrieved by adding up the smoothed profiles over all altitudes. AK is a 40×40 matrix, and when it is multiplied by X_{model}, every layer has an influence on the 39 other layers. Here, the 0-20 km column ~~is defined as the 0-20 km column thus excluding the~~ excludes maximum ozone concentrations in the stratosphere. Nevertheless, an overestimation of the ~~ozone stratospheric~~ stratospheric ozone maximum by a model could lead to an overestimation in other layers ~~. Whilst the 0-20 km columns largely reflect the ozone distribution in the lower stratosphere and give some indication about model variability in this source region for the troposphere, the 0-6 km columns are mainly influenced by lower tropospheric ozone sources (Boynard et al., 2009)~~providing an indication of the amount of ozone transported from the stratosphere to the troposphere. It should be noted that the ~~WRF-chem~~ WRF-Chem ozone profiles were completed by climatological ozone profiles between 20 and 40 km, because the convolution by the averaging kernel requires a complete vertical profile, whereas the model is limited to 20 km in altitude.

3.1.1 ~~0-20 km column~~

~~Four statistical parameters (described extensively by Kim et al. (2013)) are used to analyze the agreement between the observations~~ IASI data are averaged on a $1° \times 1°$grid and model results ~~for the 0-20 km column: the correlation coefficient (R), the normalized mean bias (NMB) and error (NME), and the root mean square error (RMSE). This altitude range includes the lower stratosphere and ozone columns are thus influenced by the large ozone concentrations in this region. The good agreement between models and IASI 0-20 km columns is shown by the results presented in Table 3: the high correlation coefficients (R > 0.93,~~

except for WRF-Chem, 0.80) and the low NME (< 20%) suggest that the models are able to reproduce the ozone columns in summer over Asia.

3.1.1 0-6 km column

were scaled to this grid. Given that the IASI sensor is not particularly sensitive to near-surface trace gas concentrations (Boynard et al., 2009), we focus here on the layer between the ground and 6 km, which is also relevant for regional air quality impacts. This tropospheric layer can also be considered to be less influenced by the stratosphere and therefore a good indicator of ozone produced over and downwind of Asian emission regions (Boynard et al., 2009). Figure 2 shows August 2008 average IASI 0-6 km ozone columns and the smoothed columns using equation Equation 1. Statistical parameters (R, NMB, RMSE, NME)correlation coefficient (R), normalized mean bias (NMB) and error (NME), root mean square error (RMSE)) based on Fig. 2 are given in Table 4.

The IASI 0-6 km columns in Fig. 2 highlight large ozone concentrations over the eastern coast of China, covering the NCP and YRD regions north of 30°30° (5-6$\times10^{17}$ molec cm^{-2}). At lower latitudes, and particularly over the PRD region, ozone concentrations are lower (3-4$\times10^{17}$ molec.cm^{-2}molec cm^{-2}). The two northerly regions are known for their high emissions of ozone precursors but over the PRD region, as seen in Fig. 1 and discussed in the IntroductionSect. 1 (Safieddine et al., 2015), the monsoon flux partially cleans the atmosphere, increases ozone destruction due to higher humidities as well as transporting pollution northward. High ozone concentrations are also observed over Korea, the Sea of Japan, and in the north-eastern part of the evaluation domain which can be attributed to transport of ozone and its precursors from China, Korea and Japan (Naja and Akimoto, 2004). Although the clear separation between high and low concentrations along 30°N and along the Chinese coast is not reproduced by the different models , a significant gradient is nevertheless visible. Moderate ozone concentrations are also observed over the sea of Japan and the north-east part of the domain but this pattern is not reproduced by the models, possibly due to underestimated emissions.Correlation coefficients between the observations and model results reflect the partial restitution of the different patterns by the models. In general, global models tend to overestimate the lower tropospheric ozone columns south of 30°N, and underestimate them slightly to the north (NME around The ECLIPSE models have too much ozone over these regions compared to IASI. This is confirmed by further statistical analysis for this region (delimited in black in Fig. 2) provided in Table 4 with, for example, model mean NME of 24%. Ozone is also overestimated further downwind over the Pacific Ocean compared to IASI in many models. Tropospheric ozone columns also too high south of 30°N, even if concentrations are much lower in this region, and may indicate that simulated relative humidities are too low. Higher modeled ozone East Asia is not due to a general overestimation in the stratospheric ozone flux, since models show good agreement with 0-20 km IASI ozone columns (high correlation coefficients (R > 0.93, except for WRF-Chem, 0.80), low NME (< 20%). Tanimoto et al. (2005) observed a latitudinal gradient in summer due to the Asian monsoon flux. The results presented here suggest that the ECLIPSE models only partially simulate the monsoon flux. , as indicated in Table 3). Other reasons for these discrepancies are discussed in Sect. 3.6.

The area delimited in black in

3.2 IASI CO columns

In a similar manner to the IASI ozone columns, IASI CO columns are smoothed using equation 1 (George et al., 2015). ECLIPSE models are compared to average August 2008 IASI total CO columns in Fig. ~~2 which encompasses regions with high ozone concentrations, namely eastern China, Korea and~~ 3. In general, models underestimate CO over the Chinese emission regions and over the Pacific downwind from Japan. Underestimation of CO over eastern Asia has already been pointed out in previous studies and suggested as a cause for the general underestimation of CO in the Northern Hemisphere (Shindell et al., 2006). Improvements to simulated CO in winter have been noted following the introduction of a seasonal cycle in domestic combustion emissions (Stein et al., 2014) and also taken into account in this study, albeit not in the same manner. However, this cannot explain the underestimation in the ~~Sea of Japan was examined in more detail. Statistical parameters comparing the IASI and smoothed model columns over this region are given in Table 4. The results show that, in general the ECLIPSE models are able to simulate the observed patterns over these regions reasonably well (mean NME of 24%). The 0-6 km column provides more precise information than the 0-20 km column about horizontal pollution features. Biases between observations and model results are similar for 0-20 km and 0-6 km ozone columns, suggesting that the models are able to simulate adequately the vertical distribution of tropospheric ozone.~~ summer months shown here over Chinese emission regions nor the apparent CO overestimation over India (Sect. 3.6).

3.3 Tropospheric ~~NO_2~~ NO_2 columns

~~NO_2~~

NO_2 is a short-lived species produced largely as a result of rapid interconversion of NO emitted from anthropogenic activities and which can be spectroscopically observed. ~~NO_2~~ NO_2 photolysis is the primary source of tropospheric ozone. Investigating ~~how models represent NO_2 can help to understand~~ modeled NO_2 provides insights into discrepancies between simulated and observed ozone. Here, tropospheric ~~NO_2~~ NO_2 columns observed by GOME-2 are compared with the model results. Column retrievals do not include corrections for aerosol scattering, which are estimated to be less than 10% by Boersma et al. (2004). Monthly mean observed tropospheric ~~NO_2~~ NO_2 columns for August and September 2008, averaged on a regular ~~$1 \times 1°$~~ $1° \times 1°$ grid, are shown in Fig. 4, as well as the absolute differences between the simulated and the observed tropospheric ~~NO_2~~ NO_2 columns. Absolute differences are shown instead of the tropospheric columns to highlight ~~the significant bias in the remote region. Since NO_2 has~~ significant biases in remote regions. Since NO_2 has a lifetime of only about 1-2 days in the lower troposphere, highest concentrations are observed close to emission areas, i.e., around Beijing and the main cities (Shanghai, ~~Hong-Kong~~Hong Kong, Seoul, Tokyo). ~~Whilst the models reproduce the patterns over the emission regions,~~ HadGEM, WRF-Chem and NorESM overestimate, and OsloCTM2 and TM4-ECPL underestimate ~~the~~ NO_2 columns with NMBs of 53, 45, 29, -3 and -11 % and NMEs of 65, 64, 51, 40 and 38 %, respectively. The same biases are seen over the delimited emission area. In ~~general, biases vary between the models which may point to differences in model chemistry or boundary layer exchange~~terms of spatial patterns, we note that the models systematically underestimate NO_2 over the southern/~~regional transport, rather than emissions, influencing simulated NO_2 lifetimes.~~ eastern part of the NCP region, as well as over Korea and Japan, possibly

pointing to an underestimation in emissions over these regions. They also tend to overestimate emission over and north of the Beijing region.

3.4 ~~Trace gas surface concentrations~~

~~Due to the low sensitivity of the IASI sensor close to the surface, it is also necessary to evaluate the models at the surface~~

3.4 Surface trace gas concentrations

As well as evaluating the models on regional/continental scales, we also evaluate the results against surface data where air quality issues are important. ~~Surface~~ Daily average surface mixing ratios of ozone and its precursors, as well as ~~SO_2~~ SO_2 (an important anthropogenic aerosol precursor), ~~daily model results~~ are compared with ground-based observations at eight sites (SNU~~and PKU station locations~~ /EANET and PKU stations) shown in Fig. 1 ~~and coordinates given in Table 2) during~~ averaged over August and September 2008. The ~~three first~~ first three stations (Beijing, Incheon, and Seoul) are urban stations whereas the last five (Gosan, Kunsan, Kangwha, Mokpo, and Taean) are located at rural locations. Therefore, we evaluate models, not only at polluted locations but also at sites downwind from major emission regions or in regions where pollution levels are lower. We note that the observations at PKU may have been influenced by the mitigation strategies put in place during the study period although we do not find very large differences between the measurements at PKU compared to Incheon and Seoul.

Fig. 5 shows box and whisker plots for modeled and observed ~~NO_2 mixing ratios .~~ NO_2 mixing ratios at these sites. There is significant variability in modeled NO_2 compared the observations at polluted and rural sites. This could be caused by differences in model vertical resolution near the surface although no correlation was found between the height of the first model layers and pollutant concentrations. While HadGEM and TM4-ECPL are able to reproduce the magnitude of ~~NO_2~~ NO_2 surface concentrations at both urban and rural sites~~.~~ , EMEP and WRF-Chem ~~results show better correspondence~~ show better agreement with measured rural concentrations ~~, whereas they~~ and tend to overestimate ~~NO_2~~ NO_2 in urban areas~~, although at the Beijing site, emissions could have been influenced by the mitigation strategies put in place during this period~~. OsloCTM2 has difficulties reproducing concentrations at both types of site and NorESM ~~underestimates~~ slightly underestimates surface NO_2 surface ~~NO_2 surface~~ concentrations in ~~general. Given that the models have been run with the same emissions, these differences are likely to be due to differences in loss of NO_2 by OH to form nitric acid in high NO_x environments and subsequent dry or wet deposition. These results are in broad agreement with the tropospheric NO_2 columns shown in Fig. 4.Modeled CO is also compared to data from eight stations~~ urban areas. Several models (NorESM, OsloCTM2, and TM4-ECPL) underestimate CO at urban locations whereas HadGEM overestimates CO as shown in Fig. 5. In general, ~~observed CO is underestimated by the models~~ all models underestimate observed CO at rural stations ~~. Such underestimation over East Asia has already been noted in a multi-model evaluation study by Shindell et al. (2006) and, more globally, in the Northern Hemisphere by Stein et al. (2014). In addition, CO is generally underestimated downwind from source regions, as shown by Shindell et al. (2008) and more recently by Monks et al. (2015). Similarly to the other models (NorESM, OsloCTM2, and TM4-ECPL), the regional model WRF-Chem shows good agreement with observed CO at urban sites but underestimates CO at rural stations. This may also~~

point to shortcomings in model chemical schemes leading to model CO lifetimes which are too short.Comparisons with surface ozone data are also presented in confirming the discrepancies found compared to IASI CO data. With regard to ozone (Fig. 5. Not surprisingly, higher ozone), higher mixing ratios are observed at the rural stations due to photochemical rural stations compared to polluted urban sites, due to less ozone titration and a switch to photochemical ozone production downwind from source regions. Differences This gradient between the urban and rural stations are locations is reproduced by EMEP , HadGEM and WRF-Chem, whereas TM4-ECPL, OsloCTM2 and NorESM simulate rather constant and too high values of but excessive ozone at both urban and rural sites. Global model overestimation of surface ozone is likely to be linked to the underestimation of CO and NO_2. SO_2 mixing ratios are Ozone in the HadGEM model is too low at urban sites. Reasons for these discrepancies are discussed further in Sect. 3.6. Comparison with observed SO_2 mixing ratios shows that models tend to overestimate concentrations both at urban and rural locations (also discussed in Sect. 4.2together with sulphate.).

3.5 Trace gas vertical distributions

Modeled vertical distributions for $NO_2$$NO_2$, CO, ozone, and $SO_2$$SO_2$ (hourly or 3-hourly profiles , depending on the model, averaged over the measurement period) are compared with observations from the CAREBEIJING 2008 airborne campaign in Fig. 6 . The observations are the averages of collected south of the main urban center of Beijing. Observed data are averages over the 12 flights performed between August 28 th and September August and 25 th September 2008 binned by altitude between 500 and 2500 m, providing a good insight into useful information about pollutant concentrations in the boundary layer (BL) and lower free troposphere pollutant distributions. Three flight routes covering the area 38-40°°N and 114-118°°E from Tianjin to Shijiazhuang were flown repeatedly. Observations were collected primarily over rural regions but nevertheless show high concentrations of pollutants representative of the large NCP region. Since the dominant winds in summer 2008 were blowing from the South, we estimate the influence of Zhang et al. (2014) showed that the flights sampling polluted air masses originated from urban areas south of the flight locations and suggested a limited influence from emission mitigation measures applied in the Beijing province on the observations to be limited. Ozone precursors are also Beijing province at this time. Observed ozone precursors are elevated up to about 1.5 km showing that the entire boundary layer is was influenced by pollution. In general, concentrations are lower than observed at the urban surface sites as discussed earlier. Maximum although maximum ozone concentrations of more than 100 ppbv were observed in some certain air masses. Using this dataset, Zhang et al. (2014) showed that, not surprisingly, the main polluted air masses originate from the urban areas south of the sampling locations. Observed NO_2 Model results were extracted along the flight paths corresponding to 2 or 3 model pixels (depending on the model) using hourly (or 3-hourly) output. This allows a fairer evaluation against the observations especially since trace gases have important diurnal cycles. Whilst the model results are an average over fairly large spatial scales, such a comparison provides useful insights into the vertical distribution of pollutants simulated by the models over a region which more representative of the less polluted background. Observed NO_2 is underestimated by the models at all vertical levelsaltitudes (except HadGEM). This result is consistent with the satellite comparison in Fig. 4 where tropospheric NO_2 columns over the southern Beijing area (where the CAREBEIJING flights were performed) NO_2 columns south of Beijing are underestimated by all the models (except HadGEM). Three models (OsloCTM2, TM4-ECPL, and WRF-Chem) also underestimate the CO mixing ratios, especially

several models. Certain models, and, in particular OsloCTM2 and NorESM, also underestimate CO between 500 and 1000 m where observed CO reached ~~0.4~~400 ppbv. This underestimation is consistent with the surface comparisons ~~at rural stations presented~~ in Fig. 5 and ~~with the~~ IASI CO tropospheric columns ~~(not shown).~~ . Comparison with airborne ozone vertical profiles shows that EMEP, TM4-ECPL, and WRF-Chem are able to capture the high concentrations observed below about 750 m~~albeit with a slight overestimation. All the other models~~ , whereas other models tend to underestimate ozone below this altitude ~~(a result that is somewhat different from the comparison with surface data where models tended to overestimate ozone at rural locations). OsloCTM2 performs better above 750 m whilst other models overestimate concentrations possibly due to too vigorous exchange in the boundary layer~~ .

3.6 Discussion

The comparisons presented in the previous sub-sections show that the model performances vary considerably. In this section we examine, in more detail, possible reasons for these discrepancies described earlier and making use of the various observations in a more synergistic manner, for example, by examining observed/modeled trace gas ratios and ozone diurnal cycles. Discrepancies may be due to differences in model resolution, transport processes such as boundary layer exchange as well as ~~too much photochemical ozone production. Solazzo et al. (2013) also identified a similar overestimation in the boundary layer in summer (as part of a regional model evaluation) and suggested this could be due to surface processes. In general, the model performances are consistent with the results presented in the previous section (0-6 km ozone columns).~~ photochemical processing or loss by deposition. The sensitivity of modeled pollutants to reducing emissions over the Beijing region is investigated using WRF-Chem in order to assess the impact of additional emission mitigation during the study period.

~~SO_2 vertical distributions are discussed in Sect. 4.2.~~

Deviations of observed trace gas ratios compared to emitted ratios can be used to determine the extent of chemical or dynamical processing that has taken place (Wang et al., 2005). In our study, observed ratios (CO:NO_x and SO_2:NO_x) at urban sites deviate from the emitted ratios. For example, at the Beijing site the CO:NO_x emission ratio is 10.7 ppbv ppbv^{-1} compared to an observed ratio of 16.7 ppbv ppbv^{-1}. This indicates that there has been stronger processing of NO_x compared to CO which is not surprising given their different chemical lifetimes (few hours compared to several weeks). It may also suggest more active mixing with cleaner air masses lower in NO_x compared to CO The latter is a longer-lived pollutant and also has significant secondary sources from VOC oxidation. Modeled ratios are generally less scattered than the observations and either lie close to emitted ratios or between the emitted and observed ratios. In models lying close to the emitted ratios (e.g. TM4-ECPL, NorESM, WRF-Chem, OsloCTM2) this points to a lack of chemical processing, particularly with respect to NO_x and may have implications for modeled ozone, as discussed hereafter. In the case of SO_2, models generally overestimate concentrations at polluted sites (and many rural sites). Over Beijing, observed SO_2:NO_x ratios are lower (0.12 ppbv ppbv^{-1}) than emission ratios (1.5 ppbv ppbv^{-1}). Models (WRF-Chem, TM4-ECPL, NorESM, EMEP) lie between observed and emitted ratios. A possible cause is that SO_2 emissions from power plants, which occur outside urban areas, such as Beijing, are placed in coarse model grid cells encompassing both urban and rural areas thereby mixing emissions from a variety of sources.

This explains why there is better agreement with the CAREBEIJING vertical profile data collected south of Beijing, where near surface concentrations were higher than those measured at the Beijing urban site. These discrepancies may also be due to emission reductions associated with the Beijing Olympics (see later discussion), although we find the same overestimation of observed SO_2 at Incheon and Seoul in most models.

3.7 ~~Discussion~~

~~EMEP, as well as~~ Significant variability is seen in the comparison between model and observed CO at polluted locations. Models (OsloCTM2, TM4-ECPL and NorESM) that significantly underestimate CO at most polluted sites also overestimate ozone indicating faster chemistry simulated in the models than occurring in urban environments. The opposite is true for HadGEM which has very low ozone and significantly overestimates CO pointing to an overestimate of ozone titration by high NO_x levels, associated with low OH radicals and thus weak CO chemical loss. This is also illustrated in Fig. 7 which compares averaged simulated ozone (for available models) with observed diurnal cycles of ozone in Beijing. The observations show a clear early afternoon maximum even if levels were slightly lower, on average during August 2008, than other years (Zhang et al., 2014). Model variability is large with several models overestimating the daytime maximum (EMEP, NorESM, WRF-Chem, ~~have a tendency to overestimate NO_2 at polluted urban locations, and given that these two models were run at rather high resolution, may point to an overestimation in the ECLIPSE NO_x emissions~~ TM4-EPCL). HadGEM has a very flat diurnal cycle with no daytime maxima consistent with an underestimation of photochemical activity in this model As noted above, the ECLIPSE models also tend to overestimate ozone at rural sites as well as downwind over the Pacific. This may be due to a variety of factors including excessive photochemical production and lack of, or excessive, NO_x titration over polluted regions as well as photochemical production during transport. Examination of the ozone:NO_z (NO_z = NO_y-NO_x) ratios can be used to examine whether a region is under a VOC or NO_x limited regime with a ratio of less than 25 indicating a VOC limited regime (Tie et al., 2013). Analysis of data collected at the PKU site during CAREBEIJING in August 2008 showed that ozone production during a high ozone episode (peaks around 150 ppbv) was due to VOC-limited ozone production until late morning followed by additional NONO_x limited production in the early afternoon (Chou et al., 2011) whereas VOC-limited ozone production prevailed during lower periods with lower observed ozone. Previous studies have noted that major emission regions in China are under VOC-limited regimes (**?**). Here, we have been able to use NO_y=NO_x+HNO_3+PAN from 3 models to examine average behavior in these models. WRF-Chem, which agrees well with the surface observations at polluted and rural sites, is largely under a VOC limited regime (ratio less than 25) over the main emission regions. However, as can be seen in Fig. 7, WRF-Chem overpredicts daytime ozone and has very low predicted nighttime ozone. The latter is due to high NO_x at night brought about by lack of processing of NO_x emissions in this model, also suggested by the analysis of CO:NO_x ratios. TM4-ECPL is also under a VOC limited regime but this model overestimates ozone at urban and rural surface sites. In this modeled NO:NO_2 ratios (lowest model level) are a factor of 2 higher than observed ratios (less than 0.5) in Beijing. As suggested earlier, this indicates insufficient conversion of NO emissions to NO_2) leading to a lack of ozone titration which may be linked to the VOC chemistry shifting the NO:NO_2 balance resulting in ozone rather than NO_z (e.g. HNO_3) formation.

In contrast, the NorESM model, which also overestimates ozone at all surface sites (e.g. Fig. 7), is in a NO_x limited regime over polluted areas. This model has too much daytime NO_2 compared to surface observations in Beijing, for example (not shown). This leads to too much photochemical ozone production over emission regions which is transported downwind over Korea/Japan and the Pacific (surface sites and IASI ozone). ~~However, comparison against the aircraft profiles, collected in the Beijingregion, shows the models are consistently too low making it difficult to conclude on this point. A more consistent pattern is found when comparing to CO data, with a general underestimation of observed concentrations at the surface and in the lower troposphere. This is more marked at rural locationssuggesting potential problems either with transport of polluted air masses from the boundary layer into the free troposphere or that the modeled lifetime of CO (against photochemical OH loss) is too short rather than necessarily an underestimation in emissions. A recent study by Monks et al. (2015) points to these factors to explain the underestimation of CO at mid-latitudes and downwind in the Arctic whereas Stein et al. (2014) noted that the introduction of a seasonal cycle in CO emissions and improvements in the CO dry deposition scheme reduces model biases. There is a tendency for models to overestimate ozone downwind of emission regions at rural sites and compared to aircraft data.~~ This may ~~be due to the general overestimation of background NO_x as shown by comparisons with satellite data.~~ also be linked to the simulation of the monsoon inflow over East Asia which penetrates too far to the north over NCP in this model leading to dilution of emissions with less polluted air masses.

The ozone discrepancies discussed above may also be due to discrepancies in the ECLIPSE emissions. While this is difficult to diagnose explicitly, model evaluation against satellite GOME-2 data, representing NO_2 over wider spatial scales, provides some consistent insights. Models tend to underestimate NO_2 over the southern and eastern part of the main Chinese NCP emission region, consistent with the evaluation against CAREBEIJING aircraft data. On the other hand, ~~underestimation of NO_2 at polluted sites, appears to result in significant overestimation of ozone in some models (OsloCTM2, TM4-ECPL) likely leading to too little ozone titration in polluted regions. Whilst it is difficult for global models~~background NO_2 is generally overestimated over the Chinese coastal region, around and to the north of Beijing, which may contribute to the overestimation of ozone downwind of the main emission regions. These spatially distributed discrepancies occur across a region with strong concentration gradients leading to over- and underestimations at surface sites. A more systematic underestimation of NO_2 over Korea and Japan by the models is found compared to the GOME-2 data suggesting that emissions over these regions may be underestimated.

The ECLIPSE models also systematically underestimate CO downwind compared to surface data over Korea, Japan, and compared to IASI CO data over Japan and downwind over the north-western Pacific Ocean. Whilst inclusion of additional seasonality in the ECLIPSE emissions (already included for domestic combustion), might improve agreement in winter and spring months (Stein et al., 2014), this is unlikely to explain these summertime differences. Low model CO appears to be linked to the clear overestimation in modeled ozone at rural sites and compared to IASI 0-6 km column data. Excessive ozone resulting in too much destruction of CO may suggest that modeled CO lifetimes are too short. This hypothesis is consistent with the findings of Monks et al. (2015) who concluded that, in models run with the same emissions, differences in

OH (chemical schemes) are a more likely cause of the systematic CO underestimation in the Northern Hemisphere and the Arctic than differences in vertical transport. Indeed, we find that surface August mean modeled OH (not shown) is higher in the NorESM model (due to the penetration of the monsoon flux) compared to, for example, TM4-EPCL and WRF-Chem over the main Chinese emission regions. In contrast, excessive modeled CO over the central Pacific, where concentrations are low, may be due to the position of the Pacific anticyclone in the meteorological analyses used by the majority of models. A shift in the position of the anticyclone to the south could result in this pattern of negative (positive) biases over the north (south) Pacific potentially as a result of transport that is too zonal. This may also explain low modeled CO in the Arctic noted by Monks et al. (2015).

To assess the possible impact of emission mitigation measures in Beijing during the period analyzed in this study, the WRF-Chem model was run for 2 weeks (1 to ~~capture pollutant distributions and gradients between urban and rural environments, these results point to the need to run models at higher resolution in order to correctly simulate the change between ozone titration and production regimes over large polluted conurbations. Evaluation against aircraft data also suggests that models have difficulties simulating high ozone concentrations (more than 100 ppbv) over the Beijing region in summer. This has implications for air quality considerations as well as the amount of ozone (and precursors)transported downwind.~~15 August 2008) with reduced pollutant emissions from the transport, industrial and solvent use sectors, following the mitigation strategy during the Olympics described in Wang et al. (2010). For example, emissions of all species in the transport sector were reduced by 75% in Beijing and 20% in the area 200 km from Beijing, corresponding to eight model grid cells around Beijing in this model. Emissions linked to the industrial sector or to solvents were reduced by 50% in the same region. Most pollutant concentrations are reduced resulting, for example, in lower CO, by about 30 ppbv, locally in and around Beijing in the emission reduction run compared to the base run. This results in ozone reductions of up to 6-7 ppbv in the region of Beijing. Based on these results, it appears that these reduction measures cannot explain the discrepancies between the models and the observations discussed earlier.

Overall, the evaluation of the ECLIPSE trace gas distributions points to excessive ozone production in many models. Potential causes vary between models and are linked to model treatments of NO_x/NO_y partitioning and VOC chemistry as well as physical factors, as noted for the NorESM model. This leads to systematic overestimation of ozone downwind of main Chinese emissions regions coupled to a general underestimation in CO concentrations in the same outflow regions. Comparison with a combination of satellite data and surface data at rural sites enables more robust conclusions to be drawn whereas comparisons at urban sites are less conclusive due to large variability in model results and difficulties for global models to reproduce fine-scale variations. This overestimation of ozone has implication for the ability of models to correctly assess regional air quality and climate impacts. Ozone anthropogenic forcing is sensitive to the altitude distribution of ozone perturbations from different emissions (e.g. Stevenson et al. (2013)).

4 ~~Evaluation~~ Interpretation of differences in modeled aerosol distributions

In this section, model results are evaluated against satellite observations from MODIS and CALIOP instruments, measuring AOD and attenuated backscatter, respectively. MODIS AOD allows a comparison of the total aerosol load integrated over the atmospheric column whereas CALIOP signals are used to evaluate vertical aerosol distributions. Simulated aerosols are also compared to observations (BC, sulphate, and OC~~are also compared to observations~~) at Beijing and Gosan ground-based stations as well as with vertical profiles from aerosol ~~LIDAR~~ lidar observations at 10 stations in the Japanese NIES network (white open triangles in Fig. 1). In Sect.4.3 we discuss reasons for differences between modeled and observed aerosol distributions.

4.1 Aerosol optical properties

4.1.1 Aerosol optical depth

AOD is determined as the aerosol extinction coefficient integrated over the whole atmospheric column. Since the aerosol extinction coefficient is mostly linked to the aerosol surface distribution (and to a lesser extent to the aerosol complex refractive index), large values of AOD can be observed in cases of high concentrations of fine mode aerosol particles, e.g. ~~,~~pollution over cities (Wang et al., 2011a). MODIS AOD fields at 550 nm were retrieved from daily observations averaged over August and September 2008 ~~,~~and taking into account missing observations primarily due to the presence of clouds within the column. Days with missing observations were removed from the model results at specific locations. The model results are bi-dimensionally interpolated on the 1×~~1°~~ 1°MODIS grid. Fig. 8 (top panels) shows average maps of observed and simulated ~~AODs~~ AOD at 550 nm. In general, the models correctly represent the main features of the spatial AOD distribution, including the large values over the NCP area and ~~over~~ northern India. However, ~~AODs are~~ AOD is not reproduced equally accurately by the models, especially over these two regions. Absolute differences between the models and MODIS are also shown in Fig. 8 (bottom panels). HadGEM and, to a lesser extent, NorESM and TM4-ECPL overestimate AOD background values. In addition, HadGEM and EMEP overestimate ~~AODs~~ AOD over NCP whereas they are underestimated by ECHAM6-HAM2, OsloCTM2, and WRF-Chem.

~~Several reasons might explain such discrepancies. Different treatments of aerosol emissions within the models can influence the AOD, especially, assumptions about the size distribution of particle emissions. All aerosols, including dust and sea salt particles can affect observed AOD, and biases in the simulated aerosol concentrations and sizes impact AOD estimates. Aerosol removal processes (i.e., dry and wet deposition) also play an important role in the aerosol loading and resulting AODs.~~

In order to further investigate model skill at simulating AOD, two specific regions with high AOD values are selected within the domain and are ~~shown~~ indicated in the top left panel of Fig. 8. The first region is located over northern India and is well known for the significant accumulation of pollutants at this time of the year, driven by the Indian monsoon. This accumulation is due to large local emissions, and the effect of dominant southerly winds causing the transport of pollution up to the Himalayas which acts as a natural barrier (Lawrence and Lelieveld, 2010). The second region encompasses the main emission areas in eastern China, including several megacities such as Beijing, Shanghai, and ~~Hong-Kong~~Hong Kong. Whilst this region is influenced by dust episodes coming from the north-eastern Asian deserts (e.g., Huang et al. (2013)) in spring, the monsoon

flux inhibits such events in summer.

Fig. 9 compares MODIS and model mean/percentile ~~AODs~~ AOD over these regions, and the rest of the domain, during August and September 2008. In terms of AOD variability, the agreement is generally better over eastern China than over northern India. More specifically, over eastern China, NorESM and TM4-ECPL capture the observed variability with a deviation of less than 10% whereas it is somewhat overestimated by HadGEM and EMEP and underestimated by ECHAM6-HAM2 and WRF-Chem. Over northern India, observed variability is reproduced by ~~the~~ HadGEM, OsloCTM2 and WRF-Chem ~~models~~ with deviations from the MODIS observations of less than 25%, whereas it is underestimated by the other models. The corresponding statistical parameters are summarized in Table 6 for the two regions and in Table 7 for the entire domain.

4.1.2 Aerosol backscatter coefficient

Evaluation against satellite observations

In this section, the ECLIPSE models are evaluated against vertical distributions of attenuated backscatter at 532 nm from CALIOP, averaged over a ~~3 × 5°~~3°×5° grid over Asia for August and September 2008. As indicated in Table 1, most of the models calculated the aerosol extinction coefficient (α) rather than the aerosol backscatter (β). Though CALIPSO level 3 data from the operational algorithm includes α, important uncertainties are associated with these retrievals. This is because α retrievals rely on inversion of lidar signals, which requires knowledge of the so-called Lidar ratio ~~S =~~ $\frac{\alpha_{aer}}{\beta_{aer}}$ $S = \frac{\alpha_{aer}}{\beta_{aer}}$, dependent on the aerosol type. Omar et al. (2010) showed that a low SNR (Signal to Noise Ratio) can lead to mis-classification and lack of aerosol layer identification, especially close to the surface. Liu et al. (2009) noted cloud contamination in backscatter and α profiles, whereas Young and Vaughan (2009) pointed out potentially erroneous assumptions in the lidar ratio S used in α retrievals. Finally, Winker et al. (2009) highlighted calibration coefficient biases in the daytime attenuated backscatter profiles. ~~To verify the~~

To verify possible aerosol mis-classification, an alternative product based on the CALIOP level 1 data, and presented by Ancellet et al. (2014), is used. This product is based on level 1 backscatter signals filtered for clouds using CALIPSO level 2 cloud masks. In this retrieval, 3 brightness temperatures (8, 10, 12 ~~um~~μm), measured by the infrared interferometer on CALIPSO, the cloud layer depolarization ratio and the color ratio are used as additional requirements. The final product described in Ancellet et al. (2014) is unitless and is called the apparent (or attenuated) scattering ratio (R_{app}). This product is not affected by errors associated with the lidar signal inversion. To allow a fair comparison, model results must be converted to R_{app} using:

$$R_{app}(z) = \frac{\beta(z)}{\beta_{mol}(z)} \times \exp\left(-2 \int_{z}^{z_{ref}} \alpha_{aer}(z')dz'\right). \quad (2)$$

β and α can be described as the sum of molecular (β_{mol} and α_{mol}) and aerosol (β_{aer} and α_{aer}) signals which describe the backscatter and extinction associated with trace gases and aerosols, respectively. z_{ref} is an altitude where only the molecular signal is observed. For some models, α only is provided. In this case, R_{app} is calculated using:

$$R_{app}(z) = \frac{\alpha_{aer,model}(z) \times BER + \beta_{mol}(z)}{\beta_{mol}(z)} \times \exp(-2 \int_{z}^{z_{ref}} \alpha_{mol}(z')dz'), \quad (3)$$

where BER is the backscatter to extinction ratio ($BER = \frac{\beta_{aer}}{\alpha_{aer}}$) and is fixed to 0.02 ~~sr^{-1}~~ sr^{-1} which is a common value observed over east-Asia (Cattrall et al., 2005; Xie et al., 2008; Chiang et al., 2008). BER is the only assumption made in the R_{app} calculation.

The distribution pattern of ~~CALIPSO-derived~~ CALIOP-derived R_{app} between 0 and 2 km (Fig. 10a) highlights three major features over Asia, consistent with the MODIS observations presented in Fig. 8. ~~They~~ Enhancements are associated with polluted regions over eastern China and northern India where anthropogenic emissions are significant, and background pollution over the desert region, north-west of the Himalayas. The ECLIPSE models reproduce the location of anthropogenic ~~plumes~~ air masses over eastern China but underestimate the magnitude by 5-50%. One exception is EMEP which overestimates the backscatter signal by more than 50%. Only OsloCTM2 and EMEP simulate the observed pattern over northern India, albeit with a slight overestimation (20-40%). The signal over the Tibetan plateau desert, which is mostly due to dust particles, is not simulated by the models. This result suggests that all models lack a source of crustal aerosols, because soil erosion, in this particular region with complex orography that is not well represented in models run at coarse resolution. Between 2 and 4 km (Fig. 10b), CALIOP detected elevated aerosols over the Tibetan plateau and eastern China. Again, none of the models are able to reproduce the signal over the desert region. However, the models (except NorESM and HadGEM) are able to capture a higher R_{app} over eastern China. Whereas most models overestimate the observed signals, it is slightly under predicted by WRF-Chem (10%). Two models (WRF-Chem and ECHAM6-HAM2) simulate aerosols at this altitude range over northern India, which probably ~~corresponding~~ corresponds to those detected by CALIOP between 0 and 2 km. Above 4 km (not shown), CALIOP only observes a significant signal over the Tibetan plateau region ~~North~~ north of the Himalayas, whereas WRF-Chem and ECHAM6-HAM2 simulate backscatter signal over the northern India and all the models, with the exception of NorESM and HadGEM, simulate some aerosols in the middle troposphere over eastern China.

~~This suggests that these models are overestimating transport of pollution into the free troposphere or that loss by wet scavenging is insufficient. This finding is in agreement with previous work noting overestimation of observed aerosol concentrations in the free troposphere (Koffi et al., 2012; Samset et al., 2014). We also note that, compared to CALIPSO, some of the models (NorESM, and to a lesser extent EMEP and OsloCTM2), simulate high amounts of aerosols below 2 km in the northeast of the domain (south of the Kamtchatka peninsula). This is due to elevated sulphate in these models (not shown).~~

Evaluation against ground-based observations

Good model skill in simulating aerosol vertical distributions is essential for reliable aerosol radiative forcing estimations (Boucher et al., 2013) and assessment of air quality impacts. Whilst evaluation against CALIOP is made with the data averaged over large grid boxes due to the scarcity of satellite overpasses, simulated optical properties are also compared with aerosol lidar measurements collected at sites in the Japanese NIES network. In NIES lidar network, full overlap between laser and the field of view of telescope is achieved above 500 m. However, compensated data are provided with geometrical form factor, that is empirically determined. Thus the lowest height of useful data is around 150 m. In order to allow a fair comparison, the simulated extinction and observed backscatter are converted to R_{app}, as described in Sect. 4.1.2, starting from an altitude of 150 m. All the NIES stations are at urban locations and the R_{app} profiles calculated from observations reveal large aerosol loads in the boundary layer and up to 3 km at certain stations.

The R_{app} mean profiles at each station are shown in Fig. 11. In general, ECHAM6-HAM2~~and NorESM~~, NorESM and HadGEM underestimate the R_{app} between the surface and 2 km whereas average profiles (shape and intensity) from EMEP, OsloCTM2 and WRF-Chem are in a fair agreement with the observations ~~. Over~~ over the same altitude range. Above 2 km, the models adequately simulate R_{app} values. These results are not always in agreement with the model comparison against MODIS and CALIOP observations. For example, the EMEP model overestimates the ~~spaceborne~~ space-borne derived backscatter signal over Japan (~~see~~ Fig. 10a and 10b) whereas it is in agreement with the ground-based lidar observations. ECHAM6-HAM2 and NorESM also show ambiguous results since the backscatter signal observed by CALIOP is overestimated by both models whereas they underestimate the signals provided by the ground-based lidars. These discrepancies may be due to (i) CALIOP uncertainties at lower altitudes ~~, especially over complex topography, and~~ particularly because aerosol products are only retrieved when clouds are not present above aerosol layers, (ii) low model resolution making it difficult for models to capture lidar profiles obtained in urban areas (NIES), and (iii) complex topography, not well resolved by global models. On the other hand, CALIOP signals are averaged over urban, rural and ~~marine~~ background regions.

~~The~~

A more quantitative parameter providing information about the profile shape is the mean altitude Z_{mean} of the NIES lidar profiles ~~, which provides information on the profile shape, is~~ calculated for each profile by a weighting function ~~following:~~ from 150 m up to 8 km, following:

$$Z_{mean} = \frac{\sum_{level=1}^{n}(R_{app,i}-1)^2 \times Z_i}{\sum_{i=1}^{n}(R_{app,i}-1)^2} \frac{\sum_{i=1}^{n}(R_{app,i}-1)^2 \times Z_i}{\sum_{i=1}^{n}(R_{app,i}-1)^2}. \quad (4)$$

where n is the number of vertical levels. Z_{mean} is overestimated by the models. Modeled overestimation of this quantity ~~implies~~ is caused by an underestimation of low altitude signals and/or an overestimation of signals at higher altitudes. ~~In general,~~ EMEP simulates Z_{mean} ~~is overestimated by the models. EMEP and ECHAM6-HAM2 simulate Z_{mean}~~ with an error of less than ~~300~~50 m, ECHAM-6-HAM2, OsloCTM2 and WRF-Chem errors are around ~~1 km, and finally~~ 0.5 km. Finally the

NorESM mean error is ~~2 km.~~ 1.5 km, giving a rather high Z_{mean} value, not only because it has quite some high values of the scattering coefficient in the upper troposphere (comparison limited to 8 km), but also because NorESM strongly underestimates aerosols in the boundary layer. This results in an overestimation of the mean height of the aerosol layer(s). Such biases are likely to be due to coarse model resolutions that are unable to adequately describe variations between rural, maritime ~~, and peri-urban~~ and urban areas, leading to an underestimation of ~~the high backscatter value~~ high backscatter values usually observed in the boundary layer in ~~the urban area~~ urban areas where the NIES lidar are operating.

4.2 Aerosol composition

In this section, modeled aerosol components which are important for estimation of anthropogenic radiative forcing (BC, OC, and sulphate), and PM estimation for air quality, are compared with ~~the~~ in-situ ground-based observations at Beijing and Gosan as shown in Fig. 12. As noted earlier, pollutant concentrations in Beijing are mainly influenced by local emissions, whereas pollution at Gosan is transported from the Asian continent (Kim et al., 2007). In addition, as noted earlier, anthropogenic emissions in the Beijing area were reduced during summer 2008, with ~~probable~~ local impacts on observed pollutant ~~amounts (Wang et al., 2009b, 2010).~~ levels (Wang et al., 2009b, 2010). The effect of the emission reduction in Beijing area is discussed in Sect. 4.3. BC originates from primary emissions due to incomplete combustion, whereas OC and sulphate are emitted from primary sources or formed as secondary products following oxidation of precursors. BC concentrations are mainly influenced by emissions and deposition and less influenced by chemical processing. On the other hand, OC and sulphate aerosols are more hydrophilic and may react with gaseous species and interact with cloud droplets. Their mass therefore evolves as a function of gas condensation at their surface, in addition to primary emissions, oxidation and ~~deposition.~~ wet/dry deposition. In models with aerosol schemes that consider internal mixing of aerosols (HadGEM, NorESM, WRF-Chem), particles containing BC can change from a more hydrophobic to a more hydrophilic state. Mean observed concentrations of BC, OC and sulphate during August and September 2008 are 2.0, 15.4, and 12.1 ~~$\mu g\, m^{-3}$~~ $\mu g\, m^{-3}$ at the Beijing site compared to 0.18, 1.3 and 7.4 ~~$\mu g\, m^{-3}$~~ $\mu g\, m^{-3}$, respectively at Gosan, the latter being consistent with observations from this site reported by Sun et al. (2004). ~~In general, the models capture the observed BC mass concentrations at Gosan but most models overpredict BC over Beijing, possibly a result of emission measures in place from mid-August to mid-September 2008.~~ OsloCTM2 simulates well the ~~gradient in~~ transition from high to low concentrations between polluted and downwind locations, whereas NorESM, which simulates polluted concentrations reasonably well, has ~~too little~~ very high BC downwind at Gosan.

ECHAM6-HAM2, TM4-ECPL, and WRF-Chem capture observed OC concentrations fairly well in Beijing but they are underestimated at Gosan. NorESM and OsloCTM2 have very low (factor ~~$\times 3$~~$\times 3$) OC over polluted Beijing and OsloCTM2 also underpredicts OC downwind over Gosan whereas EMEP largely overestimates OC in Gosan. These discrepancies, whilst based on comparisons with rather limited data, support global model evaluations (e.g., Tsigaridis et al. (2014)) showing that models have problems simulating OC and has implications with regard to estimates of radiative forcing from these aerosols. This also applies to radiative forcing estimates due to sulphate aerosols which, as shown in Fig. 12, is one of the most abundant

aerosol component measured at the sites considered here (the fraction of organics is also high). At the polluted Beijing site, observed sulphate concentrations are largely over-predicted by ECHAM6-HAM2, EMEP, and TM4-ECPL. This is linked to the overestimation of ~~SO_2~~ SO_2 as shown in Fig. 5. ECHAM6-HAM2 also overestimates sulphate at Gosan, ~~whereas they are lower than observed in~~ in contrast to the EMEP results which are lower than observed, suggesting that sulphate may be lost too fast in this model. ~~This also appears to be the case for NorESMwhich has very low sulphate at Gosan.~~ HadGEM, ~~OsloCTM2,~~ OsloCTM2, and WRF-Chem capture the concentrations reasonably well at both sites. Given that HadGEM overestimates SO_2 this suggests insufficient loss processes in this model.

4.3 Discussion

MODIS-derived AODs over the Asian region in summer 2008 highlighted that, not surprisingly, elevated AODs are observed over larger cities in eastern China and northern India, where significant accumulation of local pollutants is found close to the Himalayas, due to dominant southerly winds driven by the summer monsoon. The agreement between modeled and observed AODs is generally better over eastern China (where observed AOD is 0.45 ± 0.15) than over northern India (where observed AOD is about 0.4 ± 0.1). However, larger variability can be detected in the model AODs. Several reasons might explain such discrepancies. Different treatments of aerosol emissions within the models can influence simulated AODs, especially, assumptions about the size distribution of particle emissions. All aerosols, including dust and sea salt particles can affect observed AOD, and biases in simulated aerosol concentrations and sizes impact AOD estimates. Aerosol removal processes (i.e., dry and wet deposition) also play an important role in determining modeled aerosol loading and resulting AODs. In this study, ECHAM6-HAM2 shows very strong underestimation, not only in regions where high AODs are reported, but also in more remote regions where background aerosols dominate total mass concentrations. Because all the models use the same emissions, this suggests that this model has aerosol lifetimes that are too short, probably due to a strongly overestimated deposition efficiency. HadGEM and EMEP overestimate AODs over eastern China, but not elsewhere. This suggests that this deficiency is not due to problems in horizontal advection but rather due to a lack of deposition in the atmospheric boundary layer. The HadGEM overestimation can be also ascribed to a large overestimation of sulphate aerosols aloft, linked to a strong overestimation of SO_2 (Fig. 6). In addition, EMEP-derived AODs are strongly underestimated in northern India. Similarly, but to a lesser extent, NorESM and TM4-ECPL slightly underestimate the aerosol loading over the same area. This is an indication that the effect of pollution accumulation due to the Indian monsoon is not adequately represented. This is particularly visible for NorESM, which has lighter winds associated with the Indian monsoon (Fig. 1). WRF-Chem simulates AODs in agreement with observations over India, but underestimates over China, owing partly to a missing source of dust in dry regions of China in this model. The comparison of model results to MODIS aerosol products also highlighted that the model mean provides an excellent result, reproducing the main features of aerosol pollution in East Asia, as well as aerosol abundances over the main source regions. Using results from an ensemble of models to answer air quality-related questions or to study aerosol radiative effects is therefore recommended. The strong underestimation by ECLIPSE models of aerosol loadings identified over northern India is in agreement with the work of Gadhavi et al. (2015), who showed that BC concentrations are strongly underestimated

in southern India even when aerosol removal processes in one model were completely switched off. In our study, the fact that observed AODs in northern India are larger than those simulated by most ECLIPSE models suggests that the emissions of BC and precursors of other aerosols are underestimated for India in the ECLIPSE emission data set. This could be related to the rapid recent growth of emissions in India (Klimont et al., 2013), which may be underestimated in the inventories, for example higher emissions from kerosene lamps were identified (Lam et al., 2012) and included in the next generation ECLIPSEv5 dataset increasing the BC estimate for India by about 25% (Klimont et al., 2016), as well as with problems capturing the true spatial distribution of emissions in India.

The analysis of aerosol vertical distributions using CALIOP retrievals highlighted an underestimation in the lowest layers and an overestimation in more elevated layers for most models, except for EMEP. These results suggest that models overestimate transport of aerosol pollution into the free troposphere linked to deficiencies in model treatments of boundary layer exchange, convection, or too much vertical diffusion. Loss by wet scavenging, especially in the boundary layer, may also be insufficient. These findings are confirmed by the comparison to the NIES lidar network above 2 km, representative of aerosols downwind in altitude : the observations derived from this network indeed indicate an overestimation of the mean altitude of aerosols layers, suggesting an underestimation of aerosol deposition efficiency in the mid-troposphere during transport from urban areas in China to sites located downwind. This is in agreement with previous work noting overestimation of observed aerosol concentrations in the free (upper) troposphere (Koffi et al., 2012; Samset et al., 2014). For instance, Samset et al. (2014) evaluated model simulations performed over longer periods against aircraft measurements and found that the models systematically overpredicted BC concentrations in the remote upper troposphere over the Pacific Ocean. They concluded that the BC lifetime in the models is too long. As mentioned in Sect. 3.6, CO lifetimes appear to be too short, whereas aerosol lifetimes appear to be too long. This suggests a clear influence of wet deposition rather than chemical processing on aerosols transported downwind from East Asia in the lower troposphere. We also note that, compared to CALIOP data, certain models (NorESM, and to a lesser extent EMEP and OsloCTM2), simulate high amounts of aerosols below 2 km in the northeast of the domain (south of the Kamtchatka peninsula). This is due to elevated sulphate concentrations simulated in these models (not shown). Concerning EMEP, the overestimation of the backscatter signal is found throughout the tropospheric column supporting the suggested overestimation in aerosol lifetimes linked to an underestimation in wet deposition processes. Because lidar measurements are particularly sensitive to the presence of scattering aerosols, the overestimation of lidar signals in elevated layers points towards an overestimation (respectively, underestimation) of the aerosol scattering effect at altitude (respectively, in the planetary boundary layer) above the Asian region. In terms of aerosol-radiation interactions, the aerosol vertical profile is important for absorbing aerosols like black carbon, but in a complex way making it important to correctly simulate aerosol vertical profiles (Samset and Myhre, 2015). Furthermore, simulation of excessive aerosols between 2 and 4 km may trigger the artificial activation of aerosols as Cloud Condensation Nuclei (CCN) leading to erroneous cloud droplet formation. This would also contribute to an overestimation of the aerosol cooling effect due to the low-level clouds over the Asian continent. Since ECLIPSE models mainly only represent interactions with liquid clouds, the fact that the simulated aerosol layers aloft with enhanced concentrations are too high aloft and may end up above these clouds suggests that aerosol-cloud interactions will be

underestimated. Correct simulation of the aerosol vertical profile is critical for aerosol-cloud interactions.

At the surface, the model-mean overestimates BC in Beijing, even if there is a strong divergence of model results. Observed BC concentrations are quite low ($2 \pm 1 \mu g\,m^{-3}$) in August-September 2008, but the model-mean overpredicts concentrations by a factor of 3. Only one model (OsloCTM2) agrees well with the observations at this site. Three models simulate surface sulphate concentrations reasonably well over Beijing (WRF-Chem, HadGEM, OsloCTM2) with an average value of $10 \pm 2\ \mu g\,m^{-3}$. The other models strongly overestimate surface sulphate concentrations by a factor 3, driven by the general overestimation in SO_2 concentrations over Beijing. Good agreement with sulphate data in WRF-Chem, HadGEM and OsloCTM2 may be fortuitous since the evaluation of trace gases suggested excessive/lacking oxidizing chemistry in OsloCTM2 and HadGEM, respectively. On the other hand, measurements report OC concentrations of the order of $14 \pm 6 \mu g\,m^{-3}$ over Beijing. With the exception of two models (OsloCTM2, NorESM) that underpredict OC concentrations, the model-mean agrees rather well with the observations. These results are an improvement compared to the study of Tsigaridis et al. (2014) who reported a systematic under-prediction of organic aerosols (OA) near the surface as well as large model divergence in the middle and upper troposphere. They attributed these discrepancies to missing or underestimated OA sources, the removal parameterizations as well as uncertainties in the temperature-dependent partitioning of secondary OA in the models. At a rural location in South Korea (Gosan), the agreement is much better for the three individual components of the aerosol (BC, OC, sulphate). A very good agreement is indeed found between the model-mean and the measurements. The overestimation found for BC is due to only one model (NorESM) that strongly overestimates the surface concentrations. The BC/SO_4 ratio observed at Beijing is almost constant (~ 0.2) with enhanced values detected episodically (02-03 August, 16-17 August, 01-02 September), where the ratio can reach 1. All models reproduce this ratio reasonably well (not shown), but two (HadGEM and EMEP) show high oscillations between 0.1 and 6 (mean value of ~ 2). Over Gosan, the observed BC/SO_4 ratio is lower (~ 0.1) underlining that Gosan is a more remote site from local sources. Models also present a good agreement, except EMEP (~ 0.4) and TM4-ECPL (~ 0.2), which overestimate the ratio. Such discrepancies may affect model responses to emission perturbations and thus radiative forcing.

Results from the WRF-Chem ~~capture the concentrations reasonably well at both two sites. These findings are similar to those found when evaluating modeled optical properties (AOD and R_{app}) over eastern China~~simulations with reduced emissions due to additional mitigation measures in the Beijing area (Gao et al., 2011), discussed earlier, show that the measures taken for Olympic Games leads to small reductions in surface BC, OC and sulphate concentrations by 0.3, 1 and $1 \mu g\,m^{-3}$, respectively. This cannot explain discrepancies between model results and the observations, and especially the overestimation of surface BC and sulphate concentrations in Beijing. As a consequence, the general model overestimation at the surface close to the anthropogenic sources, and the good agreement downwind of the sources, suggests an overestimation of emissions close to local sources in Beijing. But, it also confirms that the generally good agreement found compared to MODIS AOD and the overestimation compared to CALIPSO attenuated backscatter coefficients above the planetary boundary layer are mostly due to an overestimation of vertical transport or an insufficient deposition process. Different studies highlight the potential role of a

poor representation of secondary OA production during transport to explain the underestimation of organic aerosols close to the surface (Tsigaridis et al., 2014). This is not observed here, based on the comparison with Gosan data, and is not in agreement with the overestimation of aerosols detected at altitude. As a consequence, poor representation of secondary OA production in models during transport does not appear to be a dominant factor over Asia in summer.

5 Summary

The ability of ~~six global~~ chemical-aerosol/chemistry-climate models ~~and one regional chemical aerosol model~~ to simulate distributions of short-lived ~~climate~~ pollutants is evaluated over Asia during summer 2008 using ~~satellite and in-situ observations. The models were~~ results from models run with the same 2008 ECLIPSE anthropogenic emissions and the same biomass burning dataset. ~~Where possible models werealso nudged with relevant meteorological analyses .~~Models were, in general, nudged with meteorological analyses for the study period. Model performance has been evaluated using a variety of datasets in order to assess models in different environments and over different spatial and vertical scales. We note again that these models have been used to estimate present-day air quality and climate impacts of short-lived pollutants for the present-day and future scenarios (Stohl et al., 2015). To examine ozone (and its precursors) and aerosols over major emission regions, model results were compared to surface observations at polluted and rural locations, aircraft trace gas data collected south of Beijing and satellite data. Vertically resolved aerosol lidar data collected downwind over Japan, and satellite data, were used to assess model behaviour on regional and continental scales in the lower troposphere as well as over the total atmospheric column. The assessment of model performance over different scales is important for radiative forcing estimates.

~~Ozone precursors (NO_2 and CO) are in general underestimated by the models in the troposphere , with the notable exception of surface NO_2 concentrations possibly suggesting that these emission are too high. The CO underestimation , mostly seen in rural areas, is~~

Models show systematic positive biases in ozone, especially at rural surface locations, and compared to satellite data downwind of major Chinese emission regions. The general underestimation of CO over and downwind of emissions is linked to this, most likely due ~~to too shortlifetimes (against photochemical OH loss) in models~~ excessive destruction by OH, suggesting that CO lifetimes are too short. Reasons for ozone differences varies between models but are linked to model ability to simulate VOC and NO_x regimes in polluted and less polluted environments. This may also be linked to inter-model spatial variability in compared to NO_2 surface data and NO_2 satellite column data. The latter, however, indicates a possible underestimation in NO_x emissions over Korea and Japan as well as ~~underestimation of transport from the boundary layer to the free troposphere. Ozone columns and vertical profiles are generally well captured. In contrast, global models have difficulties simulating ozone gradients at the surface between urban and rural environments~~under (over)-estimation of emissions to the south/east (west) of the Chinese NCP emission region. These findings point to the need to employ adequate model resolution ~~in order~~ to improve simulated responses to emissions when moving from ozone titration to ozone production regimes within large polluted conurbations, their surroundings and downwind. ~~This has implications for the use of global models to assess regional air quality as~~

well as climate impacts of tropospheric ozone .Overestimation of Asian ozone and its transport downwind implies that radiative forcing from this source may be overestimated. Sensitivity analyses, based on one model, suggest that emission mitigation over Beijing cannot explain these discrepancies.

Model results are compared to aerosol optical depths, as well as aerosol scattering ratios observed by MODIS and CALIOP, respectively, showing enhanced aerosol concentrations over eastern China and northern India.Over both regions, AOD predictions are model-dependent: only certain models are able to adequately reproduce summertime average MODIS AODs over eastern China (NorESM, TM4-ECPL and WRF-Chem), and over northern India (HadGEM, OsloCTM2 and WRF-Chem). In general, similar results are found when comparing to CALIOP aerosol scattering ratios. Aerosol vertical distributions between 0 and 4 km over the desert regions are not really captured at all, whereas the spatial distributions over eastern China and northern India are reasonably well simulated by most models . However, important model biases are seen in the signal intensity, suggesting that models are unable to simulate the correct amount of aerosol species with high scattering efficiency like sulphate and to a lesser extent OC. Evaluation against ground-based aerosol lidar profiles at urban locations over Japan show inconsistencies with the CALIOP comparison which can be attributed to differences in spatial representativeness of the different measurement types

Satellite-derived AOD measurements were reproduced quite well by the models over China even if evaluation of individual aerosol components over Asia overestimate ECLIPSE model-mean surface BC and sulphate aerosols in urban China in summer 2008. The effect of the short-term mitigation measures taken during the Olympic Games in summer 2008 is too weak to explain divergences between the models and observations. Our results rather point to an overestimation of emissions close to the surface in urban areas, particularly for SO_2. A potential reason for this is the fact that the spatial distribution of power plant emissions has been changing dramatically in the last decades in China (Liu et al., 2015), a change that has not been captured well in the ECLIPSEv4a dataset. A strong underestimation by ECLIPSE models of aerosol loadings has been identified over northern India, suggesting that the emissions of BC and precursors of other aerosols are underestimated in the ECLIPSEv4a inventory. Improvements have already been included, such as higher emissions from kerosene lamps, in the ECLIPSEv5 dataset. Model deficiencies in the representation of pollution accumulation due to the Indian monsoon may also play a role. The underestimation of the scattering aerosols in the lower troposphere, above the boundary layer, suggests too much vertical transport of pollutants towards the free troposphere and/or insufficient deposition in the boundary layer, leading to overestimated aerosol residence times in models.

Overall

In summary, the ECLIPSE model evaluation highlights significant differences between the models and observations, even when models are run using the same emissions . Better over East Asia. Nevertheless, an important finding is that overall the Earth System Models show a similar level of performance as the Chemistry Transport Models, which is encouraging for the further use of the Earth System Models for determining both climate and air quality impacts. Somewhat better general agreement is found for trace gases compared to aerosols, for which model simulations are very variable. For both trace gases and aerosols, models have difficulties reproducing gradients between urban and rural (downwind) locations and vertical distributions. Improved model resolution as well as improved understanding and model treatments of processes affecting pol-

lutant lifetimes are needed. Model evaluations using a variety of ~~observation types~~ observations are required so that difference aspects of model behavior can be tested. Results from this study suggest that significant uncertainties still exist in chemistry-climate simulations which has implications for the use of such models in the assessment of radiative effects of short-lived climate forcers on climate and regional/global air quality.

Acknowledgements. The authors would like to thank the European Union for funding the FP7 ECLIPSE project (grant agreement no. 282688). Jerome Fast (PNNL) is acknowledged for providing advice about the WRF-Chem simulations. This work was granted access to the HPC resources of IDRIS under the allocation 2014-017141 made by GENCI (Grand Equipement National de Calcul Intensif). The MODIS data used in this study were acquired as part of the NASA's Earth-Sun System Division and archived/distributed by the MODIS Adaptive Processing System (MODAPS). We also acknowledge the free use of tropospheric ~~NO_2~~ NO_2 column data from the GOME-2 sensor from www.temis.nl. IASI is a joint mission of EUMETSAT and the Centre National ~~d1000~~d' Etudes Spatiales (CNES, France). We are grateful to Juliette Hadji-Lazaro (LATMOS), Pierre Coheur and Daniel Hurtmans (ULB) for providing the IASI O3-FORLI data. S.-W. Kim was supported by the KMA~~R~~ R&D (~~CATER 2012–3020~~KMIPA 2015 – 2011) and by the Korean Ministry of Environment as "Climate Change Correspondence".

References

Ackerman, A. S., Toon, O. B., Stevens, D. E., Heymsfield, A. J., Ramanathan, V., and Welton, E. J.: Reduction of Tropical Cloudiness by Soot, Science, 288, 1042–1047, doi:10.1126/science.288.5468.1042, http://www.sciencemag.org/content/288/5468/1042.abstract, 2000.

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, doi:10.1126/science.245.4923.1227, http://www.sciencemag.org/content/245/4923/1227.abstract, 1989.

Allen, R. J. and Landuyt, W.: The Vertical Distribution of Black Carbon in CMIP5 Models: Comparison to Observations and the Importance of Convective Transport, Journal of Geophysical Research: Atmospheres, pp. n/a–n/a, doi:10.1002/2014JD021595, http://dx.doi.org/10.1002/2014JD021595, 2014.

Ancellet, G., Pelon, J., Blanchard, Y., Quennehen, B., Bazureau, A., Law, K. S., and Schwarzenboeck, A.: Transport of aerosol to the Arctic: analysis of CALIOP and French aircraft data during the spring 2008 POLARCAT campaign, Atmospheric Chemistry and Physics Discussions, 14, 5721–5769, doi:10.5194/acpd-14-5721-2014, http://www.atmos-chem-phys-discuss.net/14/5721/2014/, 2014.

Anenberg, S., Schwartz, J., Shindell, D., Amann, M., Faluvegi, G., Klimont, Z., Janssens-Maenhout, G., Pozzoli, L., Van Dingenen, R., Vignati, E., Emberson, L., Muller, N. Z., West, J. J., Williams, M., Demkine, V., Hicks, W. K., Kuylenstierna, J., Raes, F., and Ramanathan, V.: Global Air Quality and Health Co-benefits of Mitigating Near-Term Climate Change through Methane and Black Carbon Emission Controls., Environ Health Perspect, p. 831:839, doi:http://dx.doi.org/10.1289/ehp.1104301, 2012.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate, Geoscientific Model Development, 6, 687–720, doi:10.5194/gmd-6-687-2013, http://www.geosci-model-dev.net/6/687/2013/, 2013.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2003JD003962, http://dx.doi.org/10.1029/2003JD003962, 2004.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S.: Bounding the role of black carbon in the climate system: A scientific assessment, Journal of Geophysical Research: Atmospheres, 118, 5380–5552, doi:10.1002/jgrd.50171, http://dx.doi.org/10.1002/jgrd.50171, 2013.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, book section 7, p. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.016, www.climatechange2013.org, 2013.

Boynard, A., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Turquety, S., George, M., Hadji-Lazaro, J., Keim, C., and Meyer-Arnek, J.: Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations, Atmospheric Chemistry and Physics, 9, 6255–6271, doi:10.5194/acp-9-6255-2009, http://www.atmos-chem-phys.net/9/6255/2009/, 2009.

Boynard, A., Clerbaux, C., Clarisse, L., Safieddine, S., Pommier, M., Van Damme, M., Bauduin, S., Oudot, C., Hadji-Lazaro, J., Hurtmans, D., and Coheur, P.-F.: First simultaneous space measurements of atmospheric pollutants in the boundary layer from IASI: A case study in the North China Plain, Geophysical Research Letters, 41, 645–651, doi:10.1002/2013GL058333, http://dx.doi.org/10.1002/2013GL058333, 2014.

Buhaug, O., Corbett, J., Endresen, O., Eyring, V., Faber, J., Hanayama, S., Lee, D.S.and Lee, D., Lindstad, H., Markowska, A., Mjelde, A., Nelissen, D., Nilsen, J., Palsson, C., Winebrake, J., Wu, W., and Yoshida, K.: Second IMO GHG study 2009, Tech. rep., International Maritime Organization (IMO) London, UK, 2009.

Cao, J., Lee, S., Ho, K., Zou, S., Fung, K., Li, Y., Watson, J. G., and Chow, J. C.: Spatial and seasonal variations of atmospheric organic carbon and elemental carbon in pearl river delta region, china, Atmospheric Environment, 38, 4447 – 4456, doi:http://dx.doi.org/10.1016/j.atmosenv.2004.05.016, http://www.sciencedirect.com/science/article/pii/S1352231004004923, 2004.

Cattrall, C., Reagan, J., Thome, K., and Dubovik, O.: Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, doi:10.1029/2004JD005124, http://dx.doi.org/10.1029/2004JD005124, 2005.

Chan, C. K. and Yao, X.: Air pollution in mega cities in china, Atmospheric Environment, 42, 1 – 42, doi:http://dx.doi.org/10.1016/j.atmosenv.2007.09.003, http://www.sciencedirect.com/science/article/pii/S1352231007007911, 2008.

Chen, J., Zhao, C. S., Ma, N., Liu, P. F., Göbel, T., Hallbauer, E., Deng, Z. Z., Ran, L., Xu, W. Y., Liang, Z., Liu, H. J., Yan, P., Zhou, X. J., and Wiedensohler, A.: A parameterization of low visibilities for hazy days in the North China Plain, Atmospheric Chemistry and Physics, 12, 4935–4950, doi:10.5194/acp-12-4935-2012, http://www.atmos-chem-phys.net/12/4935/2012/, 2012.

Cherian, R., Quaas, J., Salzmann, M., and Wild, M.: Pollution trends over Europe constrain global aerosol forcing as simulated by climate models, Geophysical Research Letters, 41, 2176–2181, doi:10.1002/2013GL058715, http://dx.doi.org/10.1002/2013GL058715, 2014.

Chiang, C.-W., Das, S. K., and Nee, J.-B.: An iterative calculation to derive extinction-to-backscatter ratio based on lidar measurements, Journal of Quantitative Spectroscopy and Radiative Transfer, 109, 1187 – 1195, doi:http://dx.doi.org/10.1016/j.jqsrt.2007.10.011, http://www.sciencedirect.com/science/article/pii/S0022407307003019, 2008.

Chou, C. C.-K., Tsai, C.-Y., Chang, C.-C., Lin, P.-H., Liu, S. C., and Zhu, T.: Photochemical production of ozone in Beijing during the 2008 Olympic Games, Atmospheric Chemistry and Physics, 11, 9825–9837, doi:10.5194/acp-11-9825-2011, http://www.atmos-chem-phys.net/11/9825/2011/, 2011.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmospheric Chemistry and Physics, 9, 6041–6054, doi:10.5194/acp-9-6041-2009, http://www.atmos-chem-phys.net/9/6041/2009/, 2009.

Daskalakis, N., Myriokefalitakis, S., and Kanakidou, M.: Sensitivity of tropospheric loads and lifetimes of short lived pollutants to fire emissions, Atmospheric Chemistry and Physics Discussions, 14, 22 639–22 676, doi:10.5194/acpd-14-22639-2014, http://www.atmos-chem-phys-discuss.net/14/22639/2014/, 2014.

Ding, A. J., Wang, T., Thouret, V., Cammas, J.-P., and Nédélec, P.: Tropospheric ozone climatology over Beijing: analysis of aircraft data from the MOZAIC program, Atmospheric Chemistry and Physics, 8, 1–13, doi:10.5194/acp-8-1-2008, http://www.atmos-chem-phys.net/8/1/2008/, 2008.

Dufour, G., Eremenko, M., Orphal, J., and Flaud, J.-M.: IASI observations of seasonal and day-to-day variations of tropospheric ozone over three highly populated areas of China: Beijing, Shanghai, and Hong Kong, Atmospheric Chemistry and Physics, 10, 3787–3801, doi:10.5194/acp-10-3787-2010, http://www.atmos-chem-phys.net/10/3787/2010/, 2010.

Eckhardt, S., Quennehen, B., Olivié, D. J. L., Berntsen, T. K., Cherian, R., Christensen, J. H., Collins, W., Crepinsek, S., Daskalakis, N., Flanner, M., Herber, A., Heyes, C., Hodnebrog, Ø., Huang, L., Kanakidou, M., Klimont, Z., Langner, J., Law, K. S., Lund, M. T., Mahmood, R., Massling, A., Myriokefalitakis, S., Nielsen, I. E., Nøjgaard, J. K., Quaas, J., Quinn, P. K., Raut, J.-C., Rumbold, S. T., Schulz, M., Sharma, S., Skeie, R. B., Skov, H., Uttal, T., von Salzen, K., and Stohl, A.: Current model capabilities for simulating black

carbon and sulfate concentrations in the Arctic atmosphere: a multi-model evaluation using a comprehensive measurement data set, Atmospheric Chemistry and Physics, 15, 9413–9433, doi:10.5194/acp-15-9413-2015, http://www.atmos-chem-phys.net/15/9413/2015/, 2015.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geoscientific Model Development, 3, 43–67, doi:10.5194/gmd-3-43-2010, http://www.geosci-model-dev.net/3/43/2010/, 2010.

Flowers, B. A., Dubey, M. K., Mazzoleni, C., Stone, E. A., Schauer, J. J., Kim, S.-W., and Yoon, S. C.: Optical-chemical-microphysical relationships and closure studies for mixed carbonaceous aerosols observed at Jeju Island; 3-laser photoacoustic spectrometer, particle sizing, and filter analysis, Atmospheric Chemistry and Physics, 10, 10 387–10 398, doi:10.5194/acp-10-10387-2010, http://www.atmos-chem-phys.net/10/10387/2010/, 2010.

Fuelberg, H. E., Kiley, C. M., Hannan, J. R., Westberg, D. J., Avery, M. A., and Newell, R. E.: Meteorological conditions and transport pathways during the Transport and Chemical Evolution over the Pacific (TRACE-P) experiment, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, doi:10.1029/2002JD003092, http://dx.doi.org/10.1029/2002JD003092, 2003.

Gadhavi, H. S., Renuka, K., Ravi Kiran, V., Jayaraman, A., Stohl, A., Klimont, Z., and Beig, G.: Evaluation of black carbon emission inventories using a Lagrangian dispersion model â€“ a case study over southern India, Atmospheric Chemistry and Physics, 15, 1447–1461, doi:10.5194/acp-15-1447-2015, http://www.atmos-chem-phys.net/15/1447/2015/, 2015.

Gao, L., Zhang, M., and Han, Z.: Model analysis of seasonal variations in tropospheric ozone and carbon monoxide over East Asia, Advances in Atmospheric Sciences, 26, 312–318, doi:10.1007/s00376-009-0312-9, http://dx.doi.org/10.1007/s00376-009-0312-9, 2009.

Gao, Y., Liu, X., Zhao, C., and Zhang, M.: Emission controls versus meteorological conditions in determining aerosol concentrations in Beijing during the 2008 Olympic Games, Atmospheric Chemistry and Physics, 11, 12 437–12 451, doi:10.5194/acp-11-12437-2011, http://www.atmos-chem-phys.net/11/12437/2011/, 2011.

George, M., Clerbaux, C., Bouarar, I., Coheur, P.-F., Deeter, M. N., Edwards, D. P., Francis, G., Gille, J. C., Hadji-Lazaro, J., Hurtmans, D., Inness, A., Mao, D., and Worden, H. M.: An examination of the long-term CO records from MOPITT and IASI: comparison of retrieval methodology, Atmospheric Measurement Techniques, 8, 4313–4328, doi:10.5194/amt-8-4313-2015, http://www.atmos-meas-tech.net/8/4313/2015/, 2015.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957 – 6975, doi:http://dx.doi.org/10.1016/j.atmosenv.2005.04.027, http://www.sciencedirect.com/science/article/pii/S1352231005003560, 2005.

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., Mckay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., and Zimmerman, P.: A global model of natural volatile organic compound emissions, Journal of Geophysical Research: Atmospheres, 100, 8873–8892, doi:10.1029/94JD02950, http://dx.doi.org/10.1029/94JD02950, 1995.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmospheric Chemistry and Physics, 6, 3181–3210, doi:10.5194/acp-6-3181-2006, http://www.atmos-chem-phys.net/6/3181/2006/, 2006.

Han, Z., Sakurai, T., Ueda, H., Carmichael, G., Streets, D., Hayami, H., Wang, Z., Holloway, T., Engardt, M., Hozumi, Y., Park, S., Kajino, M., Sartelet, K., Fung, C., Bennet, C., Thongboonchoo, N., Tang, Y., Chang, A., Matsuda, K., and Amann, M.: Mics-asia ii: model intercomparison and evaluation of ozone and relevant species, Atmospheric Environment, 42, 3491

– 3509, doi:http://dx.doi.org/10.1016/j.atmosenv.2007.07.031, http://www.sciencedirect.com/science/article/pii/S1352231007006589, mICS-ASIA II, 2008.

Haywood, J. M. and Shine, K. P.: The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget, Geophysical Research Letters, 22, 603–606, doi:10.1029/95GL00075, http://dx.doi.org/10.1029/95GL00075, 1995.

He, Y. J., Uno, I., Wang, Z. F., Pochanart, P., Li, J., and Akimoto, H.: Significant impact of the East Asia monsoon on ozone seasonal behavior in the boundary layer of Eastern China and the west Pacific region, Atmospheric Chemistry and Physics, 8, 7543–7555, doi:10.5194/acp-8-7543-2008, http://www.atmos-chem-phys.net/8/7543/2008/, 2008.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geoscientific Model Development, 4, 223–253, doi:10.5194/gmd-4-223-2011, http://www.geosci-model-dev.net/4/223/2011/, 2011.

Huang, X.-F., He, L.-Y., Hu, M., and Zhang, Y.-H.: Annual variation of particulate organic compounds in pm2.5 in the urban atmosphere of beijing, Atmospheric Environment, 40, 2449 – 2458, doi:http://dx.doi.org/10.1016/j.atmosenv.2005.12.039, http://www.sciencedirect.com/science/article/pii/S1352231006000033, 2006.

Huang, X.-F., He, L.-Y., Hu, M., Canagaratna, M. R., Sun, Y., Zhang, Q., Zhu, T., Xue, L., Zeng, L.-W., Liu, X.-G., Zhang, Y.-H., Jayne, J. T., Ng, N. L., and Worsnop, D. R.: Highly time-resolved chemical characterization of atmospheric submicron particles during 2008 Beijing Olympic Games using an Aerodyne High-Resolution Aerosol Mass Spectrometer, Atmospheric Chemistry and Physics, 10, 8933–8945, doi:10.5194/acp-10-8933-2010, http://www.atmos-chem-phys.net/10/8933/2010/, 2010.

Huang, X.-X., Wang, T.-J., Jiang, F., Liao, J.-B., Cai, Y.-F., Yin, C.-Q., Zhu, J.-L., and Han, Y.: Studies on a Severe Dust Storm in East Asia and Its Impact on the Air Quality of Nanjing, China, Aerosol and air quality research, 13, 179–193, doi:http://dx.doi.org/10.4209/aaqr.2012.05.0108, 2013.

Hubanks, P. A., King, M. D., Latnick, S. P., and Pincus, R.: MODIS Algorithm Theoretical Basis Document No. ATBD-MOD-30 for Level-3 Global Gridded Atmosphere Products (08_D3, 08_E3, 08_M3), Tech. rep., Wyle Information Systems LLC, Lanham MD, 2008.

Hurtmans, D., Coheur, P.-F., Wespes, L., Clarisse, L., Scharf, O., Clerbaux, C., Hadji-Lazaro, J., George, M., and Turquety, S.: FORLI radiative transfer and retrieval code for IASI, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 1391–1408, 2012.

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjansson, J. E., Medhaug, I., Sand, M., and Seierstad, I. A.: The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections, Geoscientific Model Development, 6, 389–415, doi:10.5194/gmd-6-389-2013, http://www.geosci-model-dev.net/6/389/2013/, 2013.

Jackson, S. C.: Parallel Pursuit of Near-Term and Long-Term Climate Mitigation, Science, 326, 526–527, doi:10.1126/science.1177042, http://www.sciencemag.org/content/326/5952/526.short, 2009.

Jaffe, D., Anderson, T., Covert, D., Kotchenruther, R., Trost, B., Danielson, J., Simpson, W., Berntsen, T., Karlsdottir, S., Blake, D., Harris, J., Carmichael, G., and Uno, I.: Transport of Asian air pollution to North America, Geophysical Research Letters, 26, 711–714, doi:10.1029/1999GL900100, http://dx.doi.org/10.1029/1999GL900100, 1999.

Kanakidou, M., Duce, R. A., Prospero, J. M., Baker, A. R., Benitez-Nelson, C., Dentener, F. J., Hunter, K. A., Liss, P. S., Mahowald, N., Okin, G. S., Sarin, M., Tsigaridis, K., Uematsu, M., Zamora, L. M., and Zhu, T.: Atmospheric fluxes of organic N and P to the global ocean, Global Biogeochemical Cycles, 26, n/a–n/a, doi:10.1029/2011GB004277, http://dx.doi.org/10.1029/2011GB004277, 2012.

Kim, S.-W., Yoon, S.-C., Jefferson, A., Ogren, J. A., Dutton, E. G., Won, J.-G., Ghim, Y. S., Lee, B.-I., and Han, J.-S.: Aerosol optical, chemical and physical properties at gosan, korea during asian dust and pollution episodes in 2001, Atmospheric Environment, 39, 39 – 50, doi:http://dx.doi.org/10.1016/j.atmosenv.2004.09.056, http://www.sciencedirect.com/science/article/pii/S1352231004009471, 2005.

Kim, S.-W., Yoon, S.-C., Kim, J., and Kim, S.-Y.: Seasonal and monthly variations of columnar aerosol optical properties over east Asia determined from multi-year MODIS, LIDAR, and AERONET Sun/sky radiometer measurements, Atmospheric Environment, 41, 1634 – 1651, doi:http://dx.doi.org/10.1016/j.atmosenv.2006.10.044, http://www.sciencedirect.com/science/article/pii/S1352231006010545, 2007.

Kim, Y., Sartelet, K., Raut, J.-C., and Chazette, P.: Evaluation of the Weather Research and Forecast/Urban Model Over Greater Paris, Boundary-Layer Meteorology, 149, 105–132, doi:10.1007/s10546-013-9838-6, http://dx.doi.org/10.1007/s10546-013-9838-6, 2013.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment – optical properties in aerosol component modules of global models, Atmospheric Chemistry and Physics, 6, 1815–1834, doi:10.5194/acp-6-1815-2006, http://www.atmos-chem-phys.net/6/1815/2006/, 2006.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., and Schulz, M.: Aerosol–climate interactions in the Norwegian Earth System Model – NorESM1-M, Geoscientific Model Development, 6, 207–244, doi:10.5194/gmd-6-207-2013, http://www.geosci-model-dev.net/6/207/2013/, 2013.

Klimont, Z., Smith, S. J., and Cofala, J.: The last decade of global anthropogenic sulfur dioxide: 2000–2011 emissions, Environmental Research Letters, 8, 014 003, doi:10.1088/1748-9326/8/1/014003, 2013.

Klimont, Z., Hoglund, L., Heyes, C., Rafaj, P., Schoepp, W., Cofala, J., Borken-Kleefeld, J., Purohit, P., Kupiainen, K., Winiwarter, W., Amann, M., Zhao, B., Wang, S., and Bertok, I. an Sander, R.: Global scenarios of pollutants and methane: 1990–2050, In preparation, 2016.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., Bauer, S., Berntsen, T., Bond, T. C., Boucher, O., Chin, M., Clarke, A., De Luca, N., Dentener, F., Diehl, T., Dubovik, O., Easter, R., Fahey, D. W., Feichter, J., Fillmore, D., Freitag, S., Ghan, S., Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevaring, A., Klimont, Z., Kondo, Y., Krol, M., Liu, X., Miller, R., Montanaro, V., Moteki, N., Myhre, G., Penner, J. E., Perlwitz, J., Pitari, G., Reddy, S., Sahu, L., Sakamoto, H., Schuster, G., Schwarz, J. P., Seland, Ø., Stier, P., Takegawa, N., Takemura, T., Textor, C., van Aardenne, J. A., and Zhao, Y.: Evaluation of black carbon estimations in global aerosol models, Atmospheric Chemistry and Physics, 9, 9001–9026, doi:10.5194/acp-9-9001-2009, http://www.atmos-chem-phys.net/9/9001/2009/, 2009.

Koffi, B., Schulz, M., Bréon, F.-M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W. D., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S., Horowitz, L. W., Iversen, T., Kirkevåg, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T.: Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, doi:10.1029/2011JD016858, http://dx.doi.org/10.1029/2011JD016858, 2012.

Lam, N. L., Chen, Y., Weyant, C., Venkataraman, C., Sadavarte, P., Johnson, M. A., Smith, K. R., Brem, B. T., Arineitwe, J., Ellis, J. E., and Bond, T. C.: Household Light Makes Global Heat: High Black Carbon Emissions From Kerosene Wick Lamps, Environmental Science & Technology, 46, 13 531–13 538, doi:10.1021/es302697h, http://dx.doi.org/10.1021/es302697h, pMID: 23163320, 2012.

Lawrence, M. G. and Lelieveld, J.: Atmospheric pollutant outflow from southern Asia: a review, Atmospheric Chemistry and Physics, 10, 11 017–11 096, doi:10.5194/acp-10-11017-2010, http://www.atmos-chem-phys.net/10/11017/2010/, 2010.

Lee, D. S., Fahey, D. W., Forster, P. M., Newton, P. J., Wit, R. C., Lim, L. L., Owen, B., and Sausen, R.: Aviation and global climate change in the 21st century, Atmospheric Environment, 43, 3520 – 3537, doi:http://dx.doi.org/10.1016/j.atmosenv.2009.04.024, http://www.sciencedirect.com/science/article/pii/S1352231009003574, 2009.

Li, J., Wang, Z., Akimoto, H., Gao, C., Pochanart, P., and Wang, X.: Modeling study of ozone seasonal cycle in lower troposphere over east Asia, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2006JD008209, http://dx.doi.org/10.1029/2006JD008209, 2007.

Li, J., Wang, Z., Huang, H., Hu, M., Meng, F., Sun, Y., Wang, X., Wang, Y., and Wang, Q.: Assessing the effects of trans-boundary aerosol transport between various city clusters on regional haze episodes in spring over East China, Tellus B, 65, doi:10.3402/tellusb.v65i0.20052, http://www.tellusb.net/index.php/tellusb/article/view/20052, 2013.

Lin, M., Holloway, T., Oki, T., Streets, D. G., and Richter, A.: Multi-scale model analysis of boundary layer ozone over East Asia, Atmospheric Chemistry and Physics, 9, 3277–3301, doi:10.5194/acp-9-3277-2009, http://www.atmos-chem-phys.net/9/3277/2009/, 2009.

Lin, W., Xu, X., Zhang, X., and Tang, J.: Contributions of pollutants from North China Plain to surface ozone at the Shangdianzi GAW Station, Atmospheric Chemistry and Physics, 8, 5889–5898, doi:10.5194/acp-8-5889-2008, http://www.atmos-chem-phys.net/8/5889/2008/, 2008.

Liu, F., Zhang, Q., Tong, D., Zheng, B., Li, M., Huo, H., and He, K. B.: High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010, Atmospheric Chemistry and Physics, 15, 13 299–13 317, doi:10.5194/acp-15-13299-2015, http://www.atmos-chem-phys.net/15/13299/2015/, 2015.

Liu, H., Jacob, D. J., Chan, L. Y., Oltmans, S. J., Bey, I., Yantosca, R. M., Harris, J. M., Duncan, B. N., and Martin, R. V.: Sources of tropospheric ozone along the Asian Pacific Rim: An analysis of ozonesonde observations, Journal of Geophysical Research: Atmospheres, 107, ACH 3–1–ACH 3–19, doi:10.1029/2001JD002005, http://dx.doi.org/10.1029/2001JD002005, 2002.

Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C., and Hostetler, C.: The CALIPSO Lidar Cloud and Aerosol Discrimination: Version 2 Algorithm and Initial Assessment of Performance, J. Atmos. Oceanic Technol., 26, 1198–1213, doi:doi: 10.1175/2008JTECHA1229.1, 2009.

Ma, J., Chen, Y., Wang, W., Yan, P., Liu, H., Yang, S., Hu, Z., and Lelieveld, J.: Strong air pollution causes widespread haze-clouds over China, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2009JD013065, http://dx.doi.org/10.1029/2009JD013065, 2010.

Ma, J., Xu, X., Zhao, C., and Yan, P.: A review of atmospheric chemistry research in China: Photochemical smog, haze pollution, and gas-aerosol interactions, Advances in Atmospheric Sciences, 29, 1006–1026, doi:10.1007/s00376-012-1188-7, http://dx.doi.org/10.1007/s00376-012-1188-7, 2012.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geoscientific Model Development, 3, 519–551, doi:10.5194/gmd-3-519-2010, http://www.geosci-model-dev.net/3/519/2010/, 2010.

McCormick, R. A. and Ludwig, J. H.: Climate Modification by Atmospheric Aerosols, Science, 156, 1358–1359, doi:10.1126/science.156.3780.1358, http://www.sciencemag.org/content/156/3780/1358.abstract, 1967.

Monks, S. A., Arnold, S. R., Emmons, L. K., Law, K. S., Turquety, S., Duncan, B. N., Flemming, J., Huijnen, V., Tilmes, S., Langner, J., Mao, J., Long, Y., Thomas, J. L., Steenrod, S. D., Raut, J. C., Wilson, C., Chipperfield, M. P., Diskin, G. S., Weinheimer, A., Schlager, H., and Ancellet, G.: Multi-model study of chemical and physical controls on transport of anthropogenic and biomass burning pollution to the

Arctic, Atmospheric Chemistry and Physics, 15, 3575–3603, doi:10.5194/acp-15-3575-2015, http://www.atmos-chem-phys.net/15/3575/2015/, 2015.

Munro, R., Perez Albiñana, A., Callies, J., Corpaccioli, E., Eisinger, M., Lefebvre, A., and Hahne, A.: Expectations for GOME-2 on the METOP Satellites, in: ERS-Envisat Symposium 'Looking Down to Earth in the New Millennium', European Space Agency (Special Publication) ESA SP, Gothenburg, p. 221–227, 2000.

Myhre, G., Berglen, T. F., Johnsrud, M., Hoyle, C. R., Berntsen, T. K., Christopher, S. A., Fahey, D. W., Isaksen, I. S. A., Jones, T. A., Kahn, R. A., Loeb, N., Quinn, P., Remer, L., Schwarz, J. P., and Yttri, K. E.: Modelled radiative forcing of the direct aerosol effect with multi-observation evaluation, Atmospheric Chemistry and Physics, 9, 1365–1392, doi:10.5194/acp-9-1365-2009, http://www.atmos-chem-phys.net/9/1365/2009/, 2009.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J.-H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmospheric Chemistry and Physics, 13, 1853–1877, doi:10.5194/acp-13-1853-2013, http://www.atmos-chem-phys.net/13/1853/2013/, 2013a.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing, book section 8, p. 659–740, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.018, www.climatechange2013.org, 2013b.

Myriokefalitakis, S., Tsigaridis, K., Mihalopoulos, N., Sciare, J., Nenes, A., Kawamura, K., Segers, A., and Kanakidou, M.: In-cloud oxalate formation in the global troposphere: a 3-D modeling study, Atmospheric Chemistry and Physics, 11, 5761–5782, doi:10.5194/acp-11-5761-2011, http://www.atmos-chem-phys.net/11/5761/2011/, 2011.

Naja, M. and Akimoto, H.: Contribution of regional pollution and long-range transport to the Asia-Pacific region: Analysis of long-term ozonesonde data over Japan, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2004JD004687, http://dx.doi.org/10.1029/2004JD004687, 2004.

Nawahda, A., Yamashita, K., Ohara, T., Kurokawa, J., and Yamaji, K.: Evaluation of Premature Mortality Caused by Exposure to PM2.5 and Ozone in East Asia: 2000, 2005, 2020, Water, Air, & Soil Pollution, 223, 3445–3459, doi:10.1007/s11270-012-1123-7, http://dx.doi.org/10.1007/s11270-012-1123-7, 2012.

Omar, A., Liu, Z., Vaughan, M., Thornhill, K., Kittaka, C., Ismail, S., Hu, Y., Chen, G., Powell, K., Winker, D., Trepte, C., Winstead, E., and Anderson, B.: Extinction-to-backscatter ratios of Saharan dust layers derived from in situ measurements and CALIPSO overflights during NAMMA, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2010JD014223, http://dx.doi.org/10.1029/2010JD014223, 2010.

Penner, J. E., Prather, M. J., Isaksen, I. S. A., Fuglestvedt, J. S., Klimont, Z., and Stevenson, D. S.: Short-lived uncertainty?, Nature Geosciences, 3, 587–588, doi:10.1038/ngeo932, 2010.

Pruppacher, H. R. and Klett, J. D.: Microphysics of clouds and precipitation, Atmospheric and oceanographic sciences library, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2nd edn., 1997.

Ramanathan, V. and Carmichael, G.: Global and regional climate changes due to black carbon, Nature Geosci, 1, 221–227, doi:10.1038/ngeo156, 2008.

Richter, A., Burrows, J. P., Nusz, H., Granier, C., and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129–132, doi:10.1038/nature04092, http://dx.doi.org/10.1038/nature04092, 2005.

Rogelj, J., Schaeffer, M., Meinshausen, M., Shindell, D. T., Hare, W., Klimont, Z., Velders, G. J. M., Amann, M., and Schellnhuber, H. J.: Disentangling the effects of CO2 and short-lived climate forcer mitigation, Proceedings of the National Academy of Sciences, 111, 16 325–16 330, doi:10.1073/pnas.1415631111, http://www.pnas.org/content/111/46/16325.abstract, 2014.

Roiger, A., Schlager, H., Schäfler, A., Huntrieser, H., Scheibe, M., Aufmhoff, H., Cooper, O. R., Sodemann, H., Stohl, A., Burkhart, J., Lazzara, M., Schiller, C., Law, K. S., and Arnold, F.: In-situ observation of Asian pollution transported into the Arctic lowermost stratosphere, Atmospheric Chemistry and Physics, 11, 10 975–10 994, doi:10.5194/acp-11-10975-2011, http://www.atmos-chem-phys.net/11/10975/2011/, 2011.

Safieddine, S., Clerbaux, C., George, M., Hadji-Lazaro, J., Hurtmans, D., Coheur, P.-F., Wespes, C., Loyola, D., Valks, P., and Hao, N.: Tropospheric ozone and nitrogen dioxide measurements in urban and rural regions as seen by IASI and GOME-2, Journal of Geophysical Research: Atmospheres, 118, 10,555–10,566, doi:10.1002/jgrd.50669, http://dx.doi.org/10.1002/jgrd.50669, 2013.

Safieddine, S., Boynard, A., Hao, N., Huang, F., Wang, L., Ji, D., Barret, B., Ghude, S. D., Coheur, P.-F., Hurtmans, D., and Clerbaux, C.: Tropospheric Ozone Variability during the East Asian Summer Monsoon as Observed by Satellite (IASI), Aircraft (MOZAIC) and Ground Stations, Atmospheric Chemistry and Physics Discussions, 15, 31 925–31 950, doi:10.5194/acpd-15-31925-2015, http://www.atmos-chem-phys-discuss.net/15/31925/2015/, 2015.

Saide, P. E., Spak, S. N., Carmichael, G. R., Mena-Carrasco, M. A., Yang, Q., Howell, S., Leon, D. C., Snider, J. R., Bandy, A. R., Collett, J. L., Benedict, K. B., de Szoeke, S. P., Hawkins, L. N., Allen, G., Crawford, I., Crosier, J., and Springston, S. R.: Evaluating WRF-Chem aerosol indirect effects in Southeast Pacific marine stratocumulus during VOCALS-REx, Atmospheric Chemistry and Physics, 12, 3045–3064, doi:10.5194/acp-12-3045-2012, http://www.atmos-chem-phys.net/12/3045/2012/, 2012.

Samset, B. H. and Myhre, G.: Climate response to externally mixed black carbon as a function of altitude, Journal of Geophysical Research: Atmospheres, 120, 2913–2927, doi:10.1002/2014JD022849, http://dx.doi.org/10.1002/2014JD022849, 2014JD022849, 2015.

Samset, B. H., Myhre, G., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Penner, J. E., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and Zhang, K.: Black carbon vertical profiles strongly affect its radiative forcing uncertainty, Atmospheric Chemistry and Physics, 13, 2423–2434, doi:10.5194/acp-13-2423-2013, http://www.atmos-chem-phys.net/13/2423/2013/, 2013.

Samset, B. H., Myhre, G., Herber, A., Kondo, Y., Li, S.-M., Moteki, N., Koike, M., Oshima, N., Schwarz, J. P., Balkanski, Y., Bauer, S. E., Bellouin, N., Berntsen, T. K., Bian, H., Chin, M., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Penner, J. E., Schulz, M., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and Zhang, K.: Modelled black carbon radiative forcing and atmospheric lifetime in AeroCom Phase II constrained by aircraft observations, Atmospheric Chemistry and Physics, 14, 12 465–12 477, doi:10.5194/acp-14-12465-2014, http://www.atmos-chem-phys.net/14/12465/2014/, 2014.

Shao, M., Tang, X., Zhang, Y., and Li, W.: City clusters in China: air and surface water pollution, Frontiers in Ecology and the Environment, 4, 353–361, doi:10.1890/1540-9295(2006)004[0353:CCICAA]2.0.CO;2, 2006.

Shimizu, A., Sugimoto, N., Matsui, I., Arao, K., Uno, I., Murayama, T., Kagawa, N., Aoki, K., Uchiyama, A., and Yamazaki, A.: Continuous observations of Asian dust and other aerosols by polarization lidars in China and Japan during ACE-Asia, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2002JD003253, http://dx.doi.org/10.1029/2002JD003253, 2004.

Shindell, D., Kuylenstierna, J. C. I., Vignati, E., van Dingenen, R., Amann, M., Klimont, Z., Anenberg, S. C., Muller, N., Janssens-Maenhout, G., Raes, F., Schwartz, J., Faluvegi, G., Pozzoli, L., Kupiainen, K., Höglund-Isaksson, L., Emberson, L., Streets, D., Ra-

manathan, V., Hicks, K., Oanh, N. T. K., Milly, G., Williams, M., Demkine, V., and Fowler, D.: Simultaneously Mitigating Near-Term Climate Change and Improving Human Health and Food Security, Science, 335, 183–189, doi:10.1126/science.1210026, http://www.sciencemag.org/content/335/6065/183.abstract, 2012.

Shindell, D. T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J.-F., Pétron, G., Dentener, F. J., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Lawrence, M. G., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., Unger, N., van Noije, T. P. C., and Zeng, G.: Multimodel simulations of carbon monoxide: Comparison with observations and projected near-future changes, Journal of Geophysical Research: Atmospheres, 111, n/a–n/a, doi:10.1029/2006JD007100, http://dx.doi.org/10.1029/2006JD007100, 2006.

Shindell, D. T., Chin, M., Dentener, F., Doherty, R. M., Faluvegi, G., Fiore, A. M., Hess, P., Koch, D. M., MacKenzie, I. A., Sanderson, M. G., Schultz, M. G., Schulz, M., Stevenson, D. S., Teich, H., Textor, C., Wild, O., Bergmann, D. J., Bey, I., Bian, H., Cuvelier, C., Duncan, B. N., Folberth, G., Horowitz, L. W., Jonson, J., Kaminski, J. W., Marmer, E., Park, R., Pringle, K. J., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T. J., and Zuber, A.: A multi-model assessment of pollution transport to the Arctic, Atmospheric Chemistry and Physics, 8, 5353–5372, doi:10.5194/acp-8-5353-2008, http://www.atmos-chem-phys.net/8/5353/2008/, 2008.

Shoemaker, J. K., Schrag, D. P., Molina, M. J., and Ramanathan, V.: What Role for Short-Lived Climate Pollutants in Mitigation Policy?, Science, 342, 1323–1324, doi:10.1126/science.1240162, http://www.sciencemag.org/content/342/6164/1323.short, 2013.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, A., and Wind, P.: The EMEP MSC-W chemical transport model & technical description, Atmospheric Chemistry and Physics, 12, 7825–7865, doi:10.5194/acp-12-7825-2012, http://www.atmos-chem-phys.net/12/7825/2012/, 2012.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmospheric Chemistry and Physics, 14, 9317–9341, doi:10.5194/acp-14-9317-2014, http://www.atmos-chem-phys.net/14/9317/2014/, 2014.

Skeie, R. B., Berntsen, T. K., Myhre, G., Tanaka, K., Kvalevåg, M. M., and Hoyle, C. R.: Anthropogenic radiative forcing time series from pre-industrial times until 2010, Atmospheric Chemistry and Physics, 11, 11 827–11 857, doi:10.5194/acp-11-11827-2011, http://www.atmos-chem-phys.net/11/11827/2011/, 2011.

Smith, S. J. and Mizrahi, A.: Near-term climate mitigation by short-lived forcers, Proceedings of the National Academy of Sciences, 110, 14 202–14 206, doi:10.1073/pnas.1308470110, 2013.

Solazzo, E., Bianconi, R., Pirovano, G., Moran, M. D., Vautard, R., Hogrefe, C., Appel, K. W., Matthias, V., Grossi, P., Bessagnet, B., Brandt, J., Chemel, C., Christensen, J. H., Forkel, R., Francis, X. V., Hansen, A. B., McKeen, S., Nopmongcol, U., Prank, M., Sartelet, K. N., Segers, A., Silver, J. D., Yarwood, G., Werhahn, J., Zhang, J., Rao, S. T., and Galmarini, S.: Evaluating the capability of regional-scale air quality models to capture the vertical distribution of pollutants, Geoscientific Model Development, 6, 791–818, doi:10.5194/gmd-6-791-2013, http://www.geosci-model-dev.net/6/791/2013/, 2013.

Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., and Clerbaux, C.: On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, Atmospheric Chemistry and Physics, 14, 9295–9316, doi:10.5194/acp-14-9295-2014, http://www.atmos-chem-phys.net/14/9295/2014/, 2014.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, Journal of Advances in Modeling Earth Systems, 5, 146–172, doi:10.1002/jame.20015, http://dx.doi.org/10.1002/jame.20015, 2013.

Stevenson, D. S., Dentener, F. J., Schultz, M. G., Ellingsen, K., van Noije, T. P. C., Wild, O., Zeng, G., Amann, M., Atherton, C. S., Bell, N., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. 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., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Shindell, D. T., Strahan, S. E., Sudo, K., and Szopa, S.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, Journal of Geophysical Research: Atmospheres, 111, n/a–n/a, doi:10.1029/2005JD006338, http://dx.doi.org/10.1029/2005JD006338, 2006.

Stevenson, D. S., Young, P. J., Naik, V., Lamarque, J.-F., Shindell, D. T., Voulgarakis, A., Skeie, R. B., Dalsoren, S. B., Myhre, G., Berntsen, T. K., Folberth, G. A., Rumbold, S. T., Collins, W. J., MacKenzie, I. A., Doherty, R. M., Zeng, G., van Noije, T. P. C., Strunk, A., Bergmann, D., Cameron-Smith, P., Plummer, D. A., Strode, S. A., Horowitz, L., Lee, Y. H., Szopa, S., Sudo, K., Nagashima, T., Josse, B., Cionni, I., Righi, M., Eyring, V., Conley, A., Bowman, K. W., Wild, O., and Archibald, A.: Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmospheric Chemistry and Physics, 13, 3063–3085, doi:10.5194/acp-13-3063-2013, http://www.atmos-chem-phys.net/13/3063/2013/, 2013.

Stohl, A., Aamaas, B., Amann, M., Baker, L. H., Bellouin, N., Berntsen, T. K., Boucher, O., Cherian, R., Collins, W., Daskalakis, N., Dusinska, M., Eckhardt, S., Fuglestvedt, J. S., Harju, M., Heyes, C., Hodnebrog, Ø., Hao, J., Im, U., Kanakidou, M., Klimont, Z., Kupiainen, K., Law, K. S., Lund, M. T., Maas, R., MacIntosh, C. R., Myhre, G., Myriokefalitakis, S., Olivié, D., Quaas, J., Quennehen, B., Raut, J.-C., Rumbold, S. T., Samset, B. H., Schulz, M., Seland, Ø., Shine, K. P., Skeie, R. B., Wang, S., Yttri, K. E., and Zhu, T.: Evaluating the climate and air quality impacts of short-lived pollutants, Atmospheric Chemistry and Physics, 15, 10 529–10 566, doi:10.5194/acp-15-10529-2015, http://www.atmos-chem-phys.net/15/10529/2015/, 2015.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, doi:10.1029/2002JD003093, http://dx.doi.org/10.1029/2002JD003093, 2003.

Sugimoto, N., Matsui, I., Shimizu, A., Nishizawa, T., Hara, Y., Xie, C., Uno, I., Yumimoto, K., Wang, Z., and Yoon, S.-C.: Lidar network observations of tropospheric aerosols, in: Proc. SPIE, vol. 7153, pp. 71 530A–71 530A–13, doi:10.1117/12.806540, http://dx.doi.org/10.1117/12.806540, 2008.

Sun, Y., Zhuang, G., Wang, Y., Han, L., Guo, J., Dan, M., Zhang, W., Wang, Z., and Hao, Z.: The air-borne particulate pollution in Beijing: concentration, composition, distribution and sources, Atmospheric Environment, 38, 5991 – 6004, doi:http://dx.doi.org/10.1016/j.atmosenv.2004.07.009, http://www.sciencedirect.com/science/article/pii/S1352231004006934, 2004.

Tanimoto, H., Sawa, Y., Matsueda, H., Uno, I., Ohara, T., Yamaji, K., Kurokawa, J.-i., and Yonemura, S.: Significant latitudinal gradient in the surface ozone spring maximum over East Asia, Geophysical Research Letters, 32, n/a–n/a, doi:10.1029/2005GL023514, http://dx.doi.org/10.1029/2005GL023514, 2005.

Tie, X., Geng, F., Guenther, A., Cao, J., Greenberg, J., Zhang, R., Apel, E., Li, G., Weinheimer, A., Chen, J., and Cai, C.: Megacity impacts on regional ozone formation: observations and WRF-Chem modeling for the MIRAGE-Shanghai field campaign, Atmospheric Chemistry and Physics, 13, 5655–5669, doi:10.5194/acp-13-5655-2013, http://www.atmos-chem-phys.net/13/5655/2013/, 2013.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., Artaxo, P., Bahadur, R., Balkanski, Y., Bauer, S. E., Bellouin, N., Benedetti, A., Bergman, T., Berntsen, T. K., Beukes, J. P., Bian, H., Carslaw, K. S., Chin, M., Curci, G., Diehl, T., Easter, R. C., Ghan, S. J., Gong, S. L., Hodzic, A., Hoyle, C. R., Iversen, T., Jathar, S., Jimenez, J. L., Kaiser, J. W., Kirkevåg, A., Koch, D., Kokkola, H., Lee, Y. H., Lin, G., Liu, X., Luo, G., Ma, X., Mann, G. W., Mihalopoulos, N., Morcrette, J.-J., Müller, J.-F., Myhre, G., Myriokefalitakis, S., Ng, S., O'Donnell, D., Penner, J. E., Pozzoli, L., Pringle, K. J., Russell, L. M., Schulz, M., Sciare, J., Seland, Ø., Shindell, D. T., Sillman, S., Skeie, R. B., Spracklen, D., Stavrakou, T., Steenrod, S. D., Takemura, T., Tiitta, P., Tilmes, S., Tost, H., van Noije, T., van Zyl, P. G., von Salzen, K., Yu, F., Wang, Z., Wang, Z., Zaveri, R. A., Zhang, H., Zhang, K., Zhang, Q., and Zhang, X.: The AeroCom evaluation and intercomparison of organic aerosol in global models, Atmospheric Chemistry and Physics Discussions, 14, 6027–6161, doi:10.5194/acpd-14-6027-2014, http://www.atmos-chem-phys-discuss.net/14/6027/2014/, 2014.

Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, J. Atmos. Sci., 34, 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmospheric Chemistry and Physics, 10, 11 707–11 735, doi:10.5194/acp-10-11707-2010, http://www.atmos-chem-phys.net/10/11707/2010/, 2010.

Wang, B., Shao, M., Lu, S. H., Yuan, B., Zhao, Y., Wang, M., Zhang, S. Q., and Wu, D.: Variation of ambient non-methane hydrocarbons in Beijing city in summer 2008, Atmospheric Chemistry and Physics, 10, 5911–5923, doi:10.5194/acp-10-5911-2010, http://www.atmos-chem-phys.net/10/5911/2010/, 2010.

Wang, R., Tao, S., Shen, H., Huang, Y., Chen, H., Balkanski, Y., Boucher, O., Ciais, P., Shen, G., Li, W., Zhang, Y., Chen, Y., Lin, N., Su, S., Li, B., Liu, J., and Liu, W.: Trend in Global Black Carbon Emissions from 1960 to 2007, Environmental Science & Technology, 48, 6780–6787, doi:10.1021/es5021422, 2014.

Wang, S., Fang, L., Gu, X., Yu, T., and Gao, J.: Comparison of aerosol optical properties from Beijing and Kanpur, Atmospheric Environment, 45, 7406 – 7414, doi:http://dx.doi.org/10.1016/j.atmosenv.2011.06.055, http://www.sciencedirect.com/science/article/pii/S1352231011006674, 2011a.

Wang, T. and Xie, S.: Assessment of traffic-related air pollution in the urban streets before and during the 2008 Beijing Olympic Games traffic control period, Atmospheric Environment, 43, 5682 – 5690, doi:http://dx.doi.org/10.1016/j.atmosenv.2009.07.034, http://www.sciencedirect.com/science/article/pii/S1352231009006359, 2009.

Wang, T., Wei, X. L., Ding, A. J., Poon, C. N., Lam, K. S., Li, Y. S., Chan, L. Y., and Anson, M.: Increasing surface ozone concentrations in the background atmosphere of Southern China, 1994–2007, Atmospheric Chemistry and Physics, 9, 6217–6227, doi:10.5194/acp-9-6217-2009, http://www.atmos-chem-phys.net/9/6217/2009/, 2009a.

Wang, X., Carmichael, G., Chen, D., Tang, Y., and Wang, T.: Impacts of different emission sources on air quality during March 2001 in the Pearl River Delta (PRD) region, Atmospheric Environment, 39, 5227 – 5241, doi:http://dx.doi.org/10.1016/j.atmosenv.2005.04.035, http://www.sciencedirect.com/science/article/pii/S1352231005004206, 2005.

Wang, Y., Hao, J., McElroy, M. B., Munger, J. W., Ma, H., Chen, D., and Nielsen, C. P.: Ozone air quality during the 2008 Beijing Olympics: effectiveness of emission restrictions, Atmospheric Chemistry and Physics, 9, 5237–5251, doi:10.5194/acp-9-5237-2009, http://www.atmos-chem-phys.net/9/5237/2009/, 2009b.

Wang, Y., Zhang, Y., Hao, J., and Luo, M.: Seasonal and spatial variability of surface ozone over China: contributions from background and domestic pollution, Atmospheric Chemistry and Physics, 11, 3511–3525, doi:10.5194/acp-11-3511-2011, http://www.atmos-chem-phys.net/11/3511/2011/, 2011b.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2007GL030135, http://dx.doi.org/10.1029/2007GL030135, 2007.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, J. Atmos. Oceanic Technol., 26, 2310–2323, doi:doi: 10.1175/2008JTECHA1281.1, 2009.

Worden, H. M., Cheng, Y., Pfister, G., Carmichael, G. R., Zhang, Q., Streets, D. G., Deeter, M., Edwards, D. P., Gille, J. C., and Worden, J. R.: Satellite-based estimates of reduced CO and CO2 emissions due to traffic restrictions during the 2008 Beijing Olympics, Geophysical Research Letters, 39, n/a–n/a, doi:10.1029/2012GL052395, http://dx.doi.org/10.1029/2012GL052395, 2012.

Xie, C., Nishizawa, T., Sugimoto, N., Matsui, I., and Wang, Z.: Characteristics of aerosol optical properties in pollution and Asian dust episodes over Beijing, China, Appl. Opt., 47, 4945–4951, doi:10.1364/AO.47.004945, http://ao.osa.org/abstract.cfm?URI=ao-47-27-4945, 2008.

Xu, X., Lin, W., Wang, T., Yan, P., Tang, J., Meng, Z., and Wang, Y.: Long-term trend of surface ozone at a regional background station in eastern China 1991–2006: enhanced variability, Atmospheric Chemistry and Physics, 8, 2595–2607, doi:10.5194/acp-8-2595-2008, http://www.atmos-chem-phys.net/8/2595/2008/, 2008.

Yang, F., He, K., Ye, B., Chen, X., Cha, L., Cadle, S. H., Chan, T., and Mulawa, P. A.: One-year record of organic and elemental carbon in fine particles in downtown Beijing and Shanghai, Atmospheric Chemistry and Physics, 5, 1449–1457, doi:10.5194/acp-5-1449-2005, http://www.atmos-chem-phys.net/5/1449/2005/, 2005.

Yang, F., Tan, J., Zhao, Q., Du, Z., He, K., Ma, Y., Duan, F., Chen, G., and Zhao, Q.: Characteristics of $PM_{2.5}$ speciation in representative megacities and across China, Atmospheric Chemistry and Physics, 11, 5207–5219, doi:10.5194/acp-11-5207-2011, http://www.atmos-chem-phys.net/11/5207/2011/, 2011.

Young, S. A. and Vaughan, M. A.: The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) Data: Algorithm Description, J. Atmos. Oceanic Technol., 26, 1105–1119, doi:doi: 10.1175/2008JTECHA1221.1, 2009.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J.: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, Atmospheric Chemistry and Physics, 12, 8911–8949, doi:10.5194/acp-12-8911-2012, http://www.atmos-chem-phys.net/12/8911/2012/, 2012.

Zhang, M., Uno, I., Zhang, R., Han, Z., Wang, Z., and Pu, Y.: Evaluation of the models-3 community multi-scale air quality (cmaq) modeling system with observations obtained during the trace-p experiment: comparison of ozone and its related species, Atmospheric Environment, 40, 4874 – 4882, doi:http://dx.doi.org/10.1016/j.atmosenv.2005.06.063, http://www.sciencedirect.com/science/article/pii/S1352231006000914, special issue on Model Evaluation: Evaluation of Urban and Regional Eulerian Air Quality Models, 2006.

Zhang, W., Zhu, T., Yang, W., Bai, Z., Sun, Y. L., Xu, Y., Yin, B., and Zhao, X.: Airborne measurements of gas and particle pollutants during CAREBeijing-2008, Atmospheric Chemistry and Physics, 14, 301–316, doi:10.5194/acp-14-301-2014, http://www.atmos-chem-phys.net/14/301/2014/, 2014.

Zhao, P., Zhang, X., Xu, X., and Zhao, X.: Long-term visibility trends and characteristics in the region of beijing, tianjin, and hebei, china, Atmospheric Research, 101, 711 – 718, doi:http://dx.doi.org/10.1016/j.atmosres.2011.04.019, http://www.sciencedirect.com/science/article/pii/S016980951100130X, international Conference on Nucleation and Atmospheric Aerosols (Part 2) ICNAA 2009, 2011.

Zhou, Y., Wu, Y., Yang, L., Fu, L., He, K., Wang, S., Hao, J., Chen, J., and Li, C.: The impact of transportation control measures on emission reductions during the 2008 Olympic Games in Beijing, China, Atmospheric Environment, 44, 285 – 293, doi:http://dx.doi.org/10.1016/j.atmosenv.2009.10.040, http://www.sciencedirect.com/science/article/pii/S1352231009009157, 2010.

Models	Met. fields	Horizontal Res.	Vertical levels	Aerosol parameterisation	Biogenic emissions
ECHAM6-HAM2[10]	Nudged to ECMWF	1.8 x 1.8°	31	7 aerosol modes (HAM2)	Guenther 1990 (Guenther et al., 1995)
EMEP[4]	ECMWF	1 x 1°	20	fine and coarse mode	Guenther 1990 (Guenther et al., 1995)
HadGEM3[5,11]	Nudged to ECMWF	1.25 x 1.875°	63	GLOMAP-mode scheme	Guenther 1995 (Guenther et al., 1995)
NorESM[3,4]	online, SSTs	1.9 x 2.5°	26	mass/species (13 modes)	MEGAN v2 (Guenther et al., 2006)
OsloCTM2[3]	ECMWF-IFS	2.81 x 2.81°	60	bulk aerosol scheme	Guenther 1990 (Guenther et al., 1995)
TM4-ECPL[6,7]	ECMWF-ERA INTERIM	2 x 3°	34	fine and coarse modes	MEGAN-MACC (Sindelarova et al., 2014)
WRF-Chem[1]	NCEP FNL	50 x 50 km	49	8 bins/species 40-10000 nm	MEGAN v2 (Guenther et al., 2006)

Models	Trace gases	Aerosol comp.	Optical param.	References
ECHAM6-HAM2	SO_2	BC, OC, SO_4	AOD, α_{550}	Stevens et al. (2013); Zhang et al. (2012)
EMEP	CO,O_3,NO_2,SO_2	BC, OC, SO_4	AOD, α_{550}	Simpson et al. (2012)
HadGEM3	CO,O_3,NO_2,SO_2	BC, SO_4	AOD	Mann et al. (2010); Hewitt et al. (2011)
NorESM	CO,O_3,NO_2,SO_2	BC, OC, SO_4	AOD, β_{550}	Kirkevåg et al. (2013); Iversen et al. (2013); Bentsen et al. (2013)
OsloCTM2	CO,O_3,NO_2	BC, OC, SO_4	AOD, α_{550}	Myhre et al. (2009); Skeie et al. (2011)
TM4-ECPL	CO,O_3,NO_2,SO_2	BC, OC, SO_4	AOD	Myriokefalitakis et al. (2011); Kanakidou et al. (2012); Daskalakis et al. (2014)
WRF-Chem	CO,O_3,NO_2,SO_2	BC, OC, SO_4	AOD, β_{550}	Grell et al. (2005)

Table 1. ECLIPSE model description including meteorological fields used to nudge simulations (where applicable), spatial resolution, aerosol schemes, and biogenic emissions. Trace gases, aerosol species and optical parameters output by the models are provided together with references for the different models. Institutes responsible for each model are indicated with the indices of the author affiliations.

Station	Country	longitude (°E)	latitude (°N)	Available parameters
Beijing	China	116.3	40.0	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2, BC, OC, ~~SO_4~~SO_4
Gosan		126.2	33.3	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2, BC, OC, ~~SO_4~~SO_4
Inchon		126.6	37.5	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Kangwha		126.3	37.7	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Kunsan	South Korea	126.7	36.0	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Mokpo		126.4	34.8	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Seoul		127.0	37.6	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Taean		126.4	39.7	CO, ~~NO_2, O_3, SO_2~~NO_2, O_3, SO_2
Chiba		140.1	36.0	β_{532}
Fukue		128.7	32.7	β_{532}
Matsue		133.0	35.5	β_{532}
Nagasaki		130.0	32.9	β_{532}
Niigata		138.9	37.8	β_{532}
Osaka	Japan	135.6	34.6	β_{532}
Sapporo		141.3	43.1	β_{532}
Tokyo		139.7	35.7	β_{532}
Toyama		137.1	36.7	β_{532}
Tsukuba		140.1	36.0	β_{532}

Table 2. Coordinates of stations used in this study and available parameters.

Models	R	NMB (%)	RMSE (~~molec.cm^{-2}~~molec cm^{-2})	NME (%)
HadGEM	0.95	9.4	2.50×10^{17}	10.4
NorESM	0.93	10.1	2.86×10^{17}	11.8
OsloCTM2	0.96	-0.01	1.48×10^{17}	5.4
TM4-ECPL	0.98	9.6	2.90×10^{17}	10.7
WRF-Chem	0.80	-17.5	5.36×10^{17}	18.1
Model mean	0.97	~~25.0~~2.5	1.30×10^{17}	5.2

Table 3. Statistical parameters (correlation coefficient R, normalized mean bias NMB, root mean square error RMSE, and normalized mean error NME) based on spatial variations for model simulations of 0-20 km ozone column over Asia in August 2008 compared to the IASI ozone-FORLI observations.

	Asia				High ozone region			
Models	R	NMB (%)	RMSE (molec cm^{-2})	NME (%)	R	NMB (%)	RMSE (molec cm^{-2})	NME (%)
HadGEM	0.76	15.4	0.93×10^{17}	18.5	0.72	12.1	1.09×10^{17}	24.3
NorESM	0.82	16.7	0.88×10^{17}	19.0	0.62	7.0	1.28×10^{17}	28.3
OsloCTM2	0.56	3.0	1.03×10^{17}	23.0	0.51	-3.9	1.40×10^{17}	32.9
TM4-ECPL	0.63	8.9	0.93×10^{17}	19.8	0.65	7.5	1.16×10^{17}	24.7
WRF-Chem	0.69	30.2	1.45×10^{17}	31.8	0.66	23.6	1.46×10^{17}	35.2
Model mean	0.76	13.5	0.87×10^{17}	18.1	0.72	12.2	1.10×10^{17}	24.1

Table 4. Statistical parameters (correlation coefficient R, normalized mean bias NMB, root mean square error RMSE, and normalized mean error NME) based on spatial variations calculated using simulations of the 0-6 km ozone columns over Asia (left parameters) and over the high concentration columns delimited in black in Fig. 2 (right parameters) in August 2008 by the ECLIPSE models compared to the IASI ozone-FORLI observations.

	Asia				Chinese emissions area			
Models	R	NMB (%)	RMSE (molec cm^{-2})	NME (%)	R	NMB (%)	RMSE (molec cm^{-2})	NME (%)
HadGEM	0.76	53.1	1.91×10^{15}	65.4	0.78	30.1	2.90×10^{15}	59.5
NorESM	0.80	29.1	1.27×10^{15}	50.7	0.75	8.1	2.63×10^{15}	47.9
OsloCTM2	0.76	-3.1	1.19×10^{15}	40.1	0.68	-19.8	2.80×10^{15}	51.4
TM4-ECPL	0.81	-11.6	1.09×10^{15}	38.5	0.81	-18.1	2.23×10^{15}	38.7
WRF-Chem	0.77	45.3	1.31×10^{15}	64.5	0.79	13.2	2.33×10^{15}	53.9
Model mean	0.84	22.6	1.04×10^{15}	41.1	0.83	-4.1	2.14×10^{15}	37.8

Table 5. Statistical parameters (correlation coefficient R, normalized mean bias NMB, root mean square error RMSE, and normalized mean error NME) based on spatial variations for model simulations of ~~NO_2~~ NO_2 tropospheric columns over Asia and over Chinese emission area shown in Fig. 4 averaged over August and September 2008 compared to GOME-2 satellite observations.

Eastern China								
Models	mean (%)	median (%)	75^{th}-25^{th} (%)	whiskers (%)	R	NMB (%)	RMSE	NME (%)
ECHAM6-HAM2	-47.9	-48.4	-37.2	-68.0	0.70	-36.0	0.21	38.9
EMEP	24.5	13.2	54.1	43.7	0.71	41.2	0.30	49.5
HadGEM	29.6	25.7	34.2	51.2	0.73	29.1	0.20	37.9
NorESM	-2.1	5.3	-3.8	-2.3	0.39	-1.9	0.20	37.7
OsloCTM2	-17.4	-20.7	-6.0	-25.8	0.63	-14.1	0.19	30.2
TM4-ECPL	-2.8	3.7	-6.1	-5.7	0.69	-2.0	0.15	25.9
WRF-Chem	-16.8	-20.4	5.8	-17.5	0.73	12.2	0.20	30.5
Model mean	-4.7	-4.1	-3.4	-5.6	0.82	4.1	0.12	20.8
Northern India								
Models	mean (%)	median (%)	75^{th}-25^{th} (%)	whiskers (%)	R	NMB (%)	RMSE	NME (%)
ECHAM6-HAM2	-46.7	-47.4	-58.5	-63.2	0.62	-46.1	0.17	46.2
EMEP	-56.4	-68.9	-11.6	-59.0	0.56	-35.5	0.18	51.2
HadGEM	-2.9	3.0	11.9	2.6	0.78	24.8	0.13	32.6
NorESM	-36.8	-36.0	-78.7	-61.0	0.55	-23.9	0.13	33.2
OsloCTM2	13.2	-5.0	12.3	22.3	0.33	39.8	0.29	64.9
TM4-ECPL	-33.2	-32.7	-55.6	-47.6	0.72	-26.0	0.12	30.4
WRF-Chem	-17.5	-13.5	-7.0	-15.3	0.77	-2.1	0.10	26.0
Model mean	-25.8	-22.7	-43.0	-35.4	0.78	-9.9	0.08	21.9

Table 6. Deviation (in %) of model results from MODIS AODs observations based on spatial variations: mean, median, differences between 75^{th}-25^{th} and whiskers. For each parameter, the value corresponds to the following relationship: $\frac{AOD_{model}-AOD_{MODIS}}{AOD_{MODIS}}$ and statistical parameters (correlation coefficient R, normalized mean bias NMB, root mean square error RMSE, and normalized mean error NME) are calculated using monthly mean AODs over eastern China and northern India in August and September 2008.

Rest of the domain				
Models	R	NMB (%)	RMSE	NME (%)
ECHAM6-HAM2	0.66	-55.3	0.17	57.3
EMEP	0.70	-29.3	0.16	50.4
HadGEM	0.72	21.4	0.13	41.1
NorESM	0.62	-3.3	0.13	37.5
OsloCTM2	0.62	-31.1	0.15	43.5
TM4-ECPL	0.71	-0.49	0.12	32.6
WRF-Chem	0.63	-27.1	0.14	39.9
Model mean	0.78	-17.8	0.10	28.4

Table 7. Correlation coefficient R, normalized mean bias NMB, root mean square error RMSE, and normalized mean error NME) based on spatial variations calculated from monthly mean differences between MODIS and simulated AODs over the Asian domain for August and September 2008.

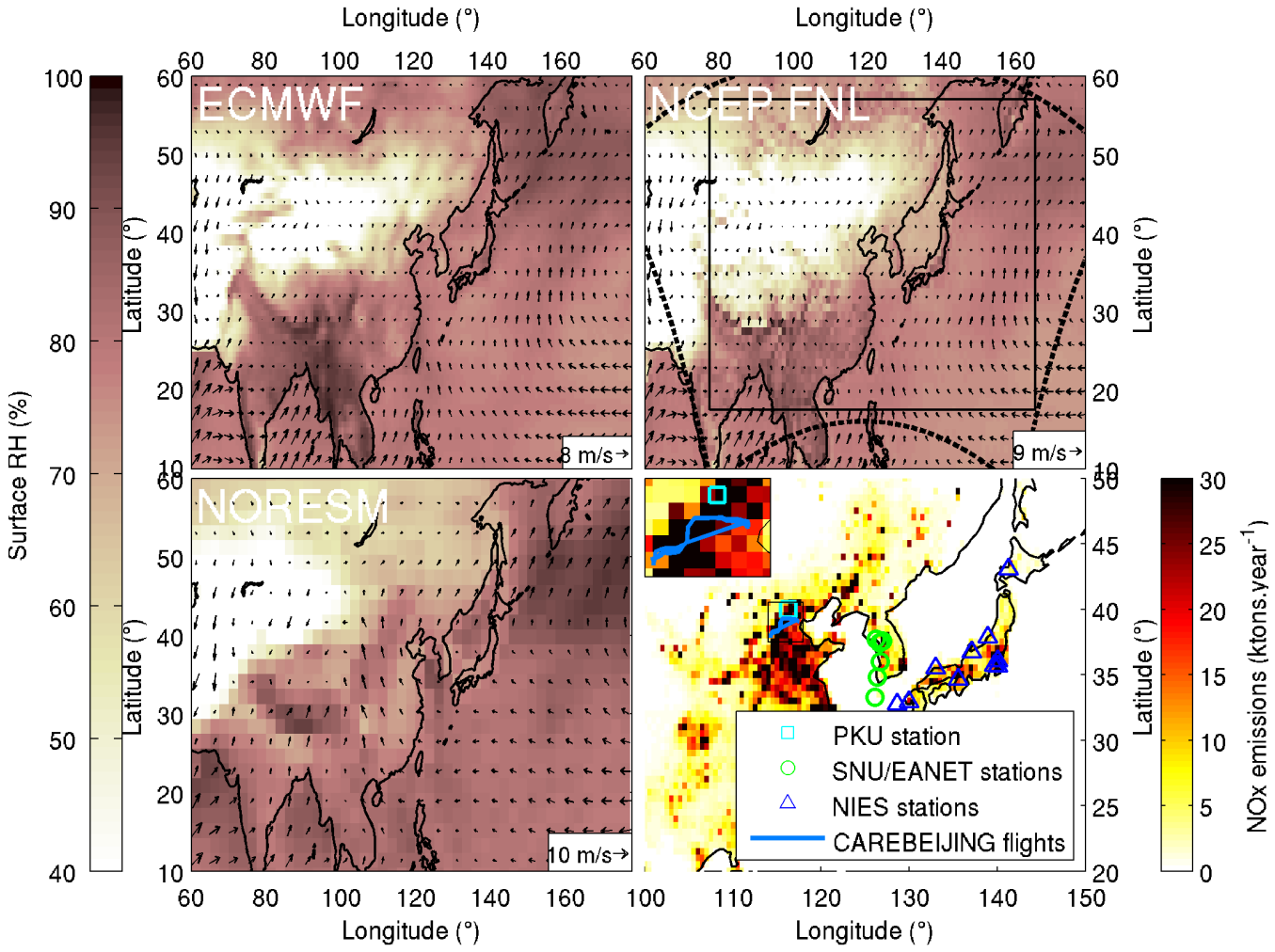

Figure 1. Map of the Asian region showing mean surface relative humidity (%) and surface wind speed (m/s) and direction (°) for August and September 2008 from ECMWF (top left panel), NCEP (National Centers for Environmental Prediction) FNL (final, top right panel) , and NorESM (bottom left panel), and the ~~NO_x~~ NO_x emissions over east-Asia (bottom right panel). The WRF-Chem domain (dashed line) and the satellite data comparison domain (thick black line) are shown in the top right panel whereas the ground-based stations (cyan square, green circles, blue triangles), and the CAREBEIJING flight tracks (blue lines) used in this study are shown in the bottom right panel.

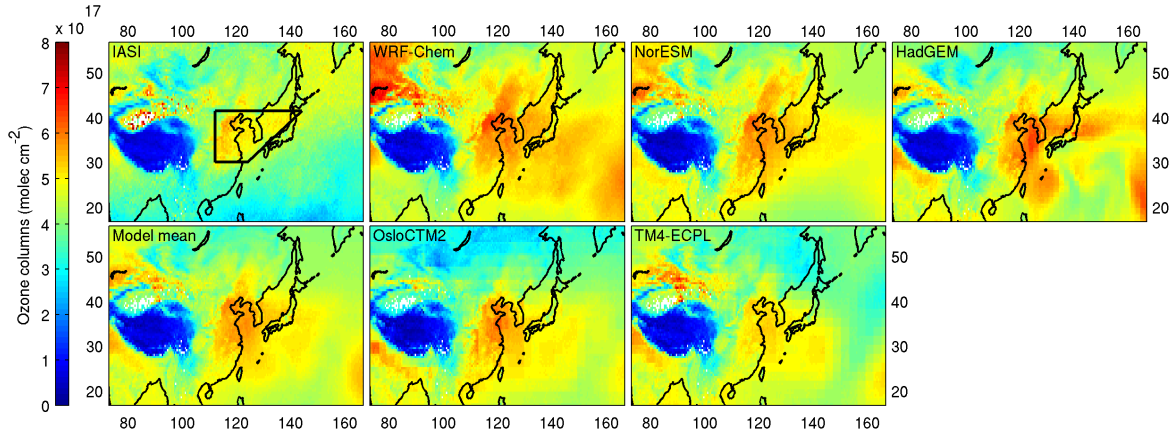

Figure 2. Average 0-6 km ozone columns (~~molec.cm^{-2}~~ molec cm^{-2}) over Asia in August 2008 observed by the IASI satellite (left panel) and simulated by the models (models names are given in each relevant panel). The black polygon delimits the region discussed in detail in the text.

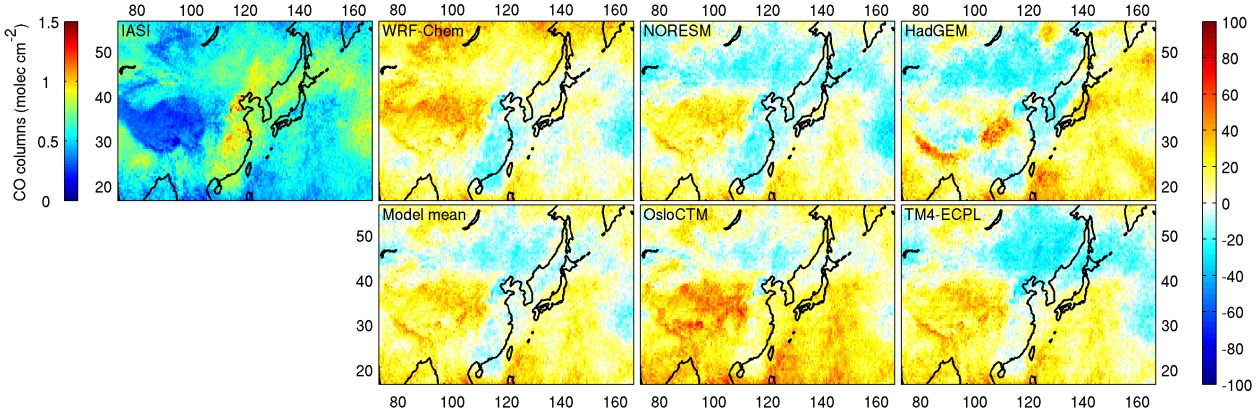

Figure 3. Average total CO columns (molec cm^{-2}) over Asia in August 2008 observed by the IASI satellite (left panel) and relative differences between columns observed by IASI and simulated by the models (models names are given in each relevant panel).

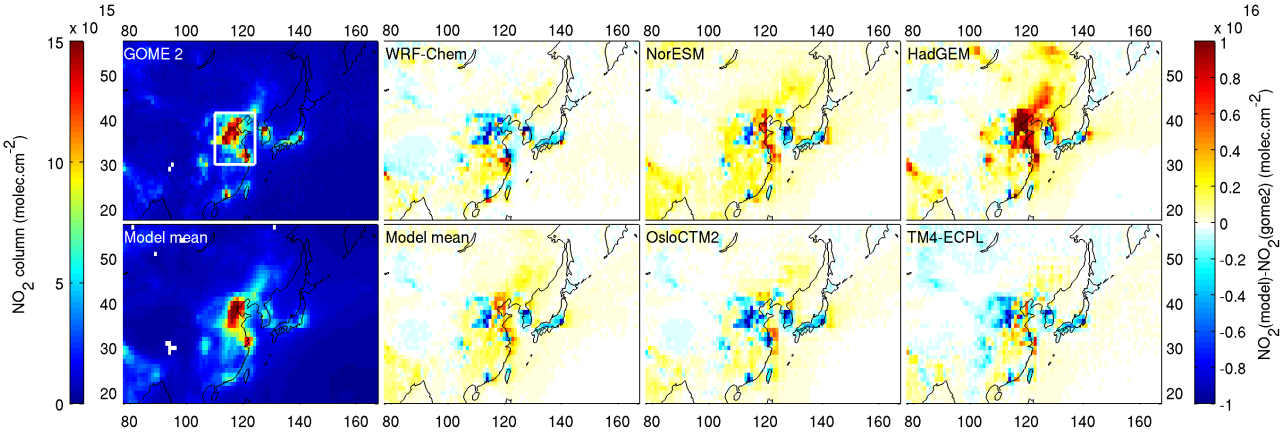

Figure 4. Mean tropospheric ~~NO_2~~ NO_2 columns in ~~molec.cm^{-2}~~ molec cm^{-2} (between the ground and the tropopause height given in the GOME-2 product) in August and September 2008 over Asia as observed by the GOME-2 satellite and absolute differences between GOME-2 and model simulations (model names are given in the panels). Model mean tropospheric columns are also presented in the bottom left panel. The white square denotes the emission region discussed in the text.

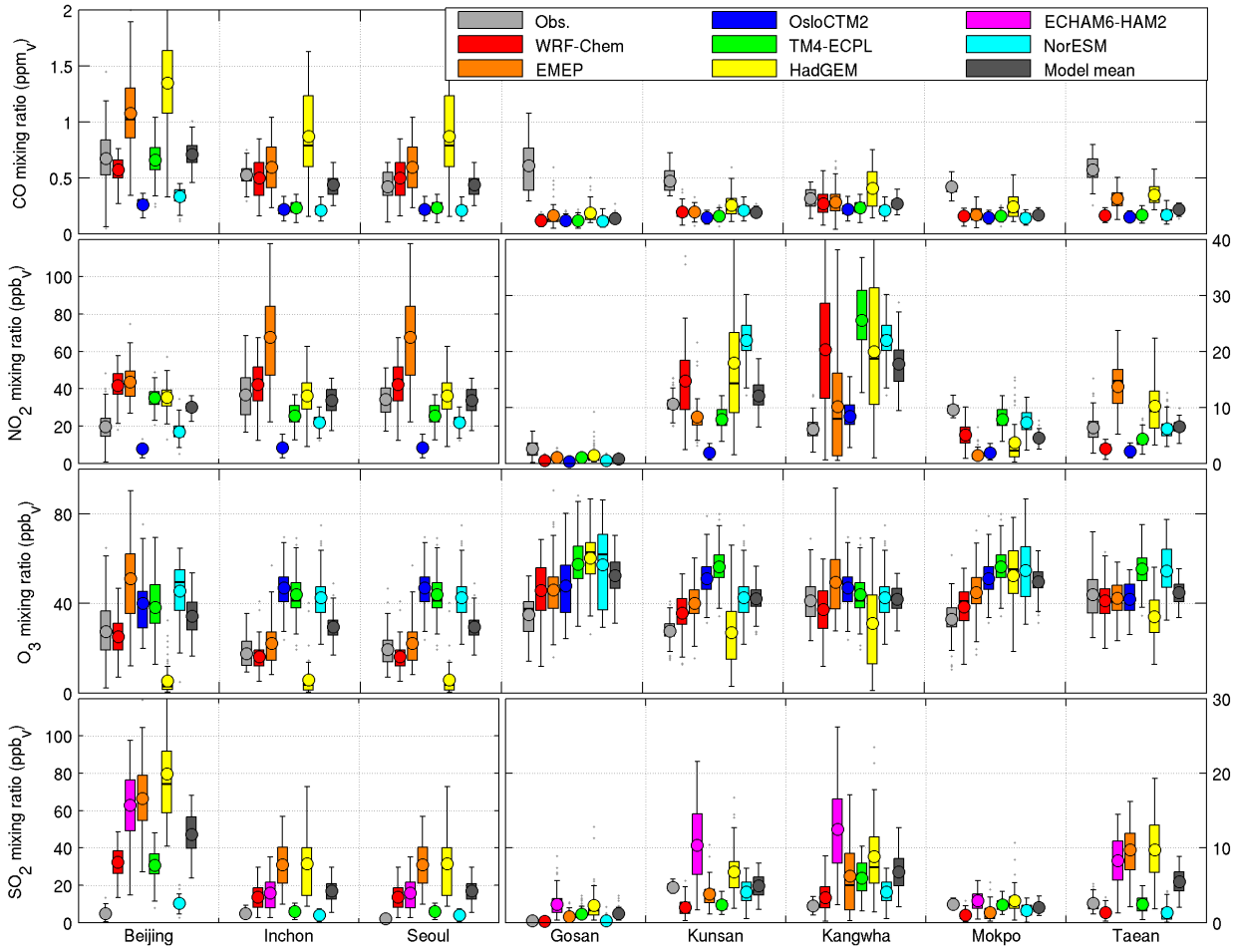

Figure 5. Box and whisker plots showing (from top to bottom) CO, ~~NO_2~~NO_2, ozone, and ~~SO_2~~ SO_2 mean (circle), median (central line), 25^{th} and 75^{th} percentile (box edges) concentrations in volume mixing ratio (ppbv), during August and September 2008 as observed and simulated at eight surface sites in East Asia. The whiskers encompass values from ~~$25^{th} - 1.5 * (75^{th} - 25^{th})$~~ $25^{th} - 1.5x(75^{th} - 25^{th})$ to the ~~$75^{th} + 1.5 * (75^{th} - 25^{th})$~~ $75^{th} + 1.5x(75^{th} - 25^{th})$. This range covers more than 99% of a normally distributed dataset. Note the different scales between urban (3 left) and rural (5 right) stations for ~~NO_2~~ NO_2 and ~~SO_2~~ SO_2 panels.

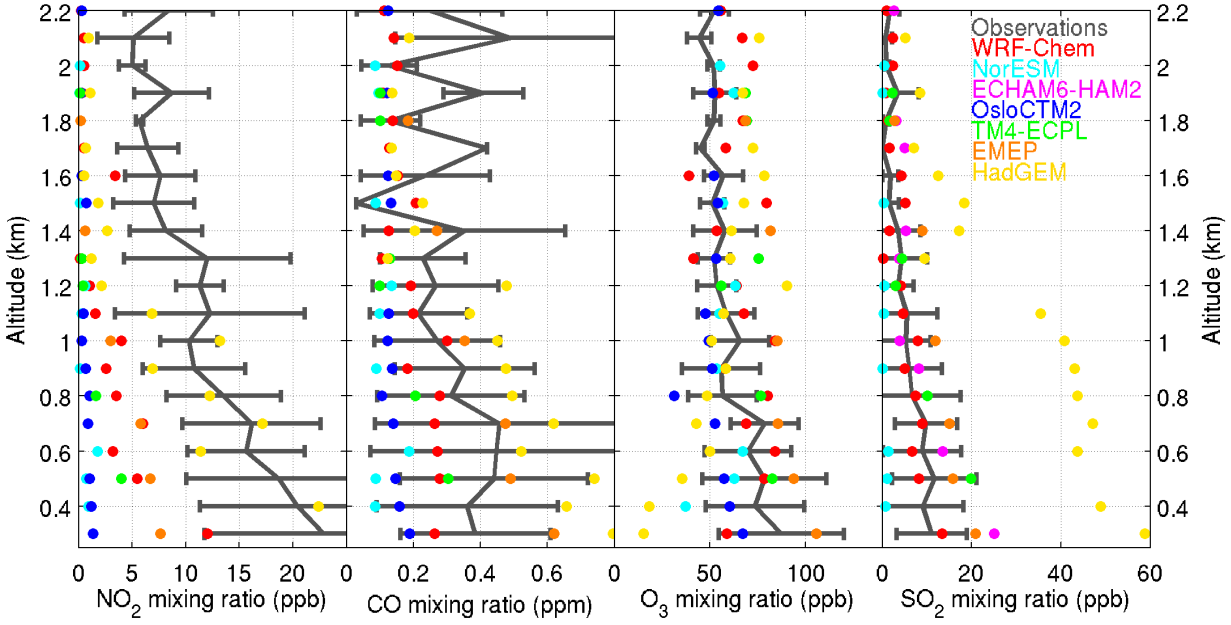

Figure 6. Mean vertical profiles of ozone, ~~NO_2~~NO_2, CO, and ~~SO_2~~ SO_2 observed over China during the CAREBEIJING 2008 airborne campaign and simulated by the ECLIPSE models.

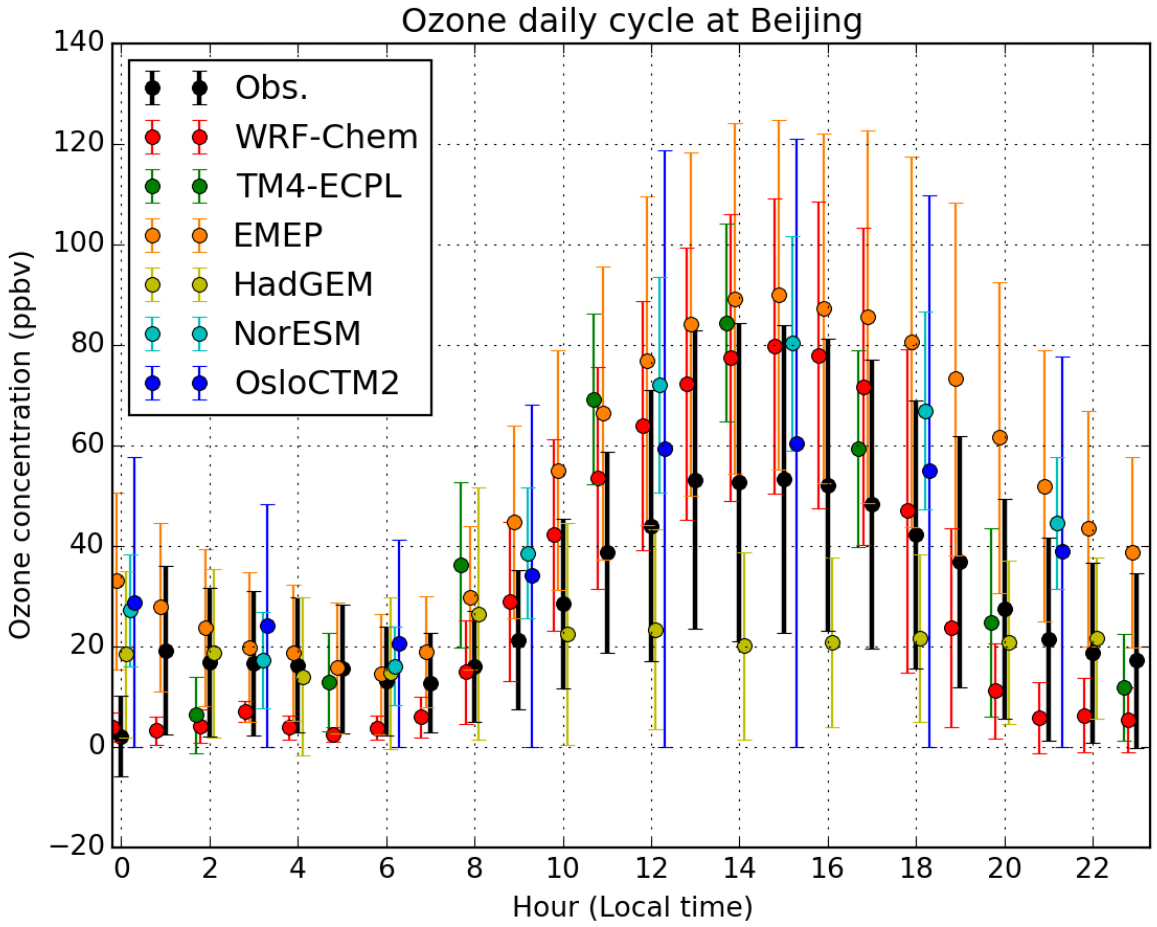

Figure 7. Ozone diurnal cycle as observed and as simulated by the ECLIPSE models in Beijing averaged over August and September 2008. Whiskers represent two standard deviations.

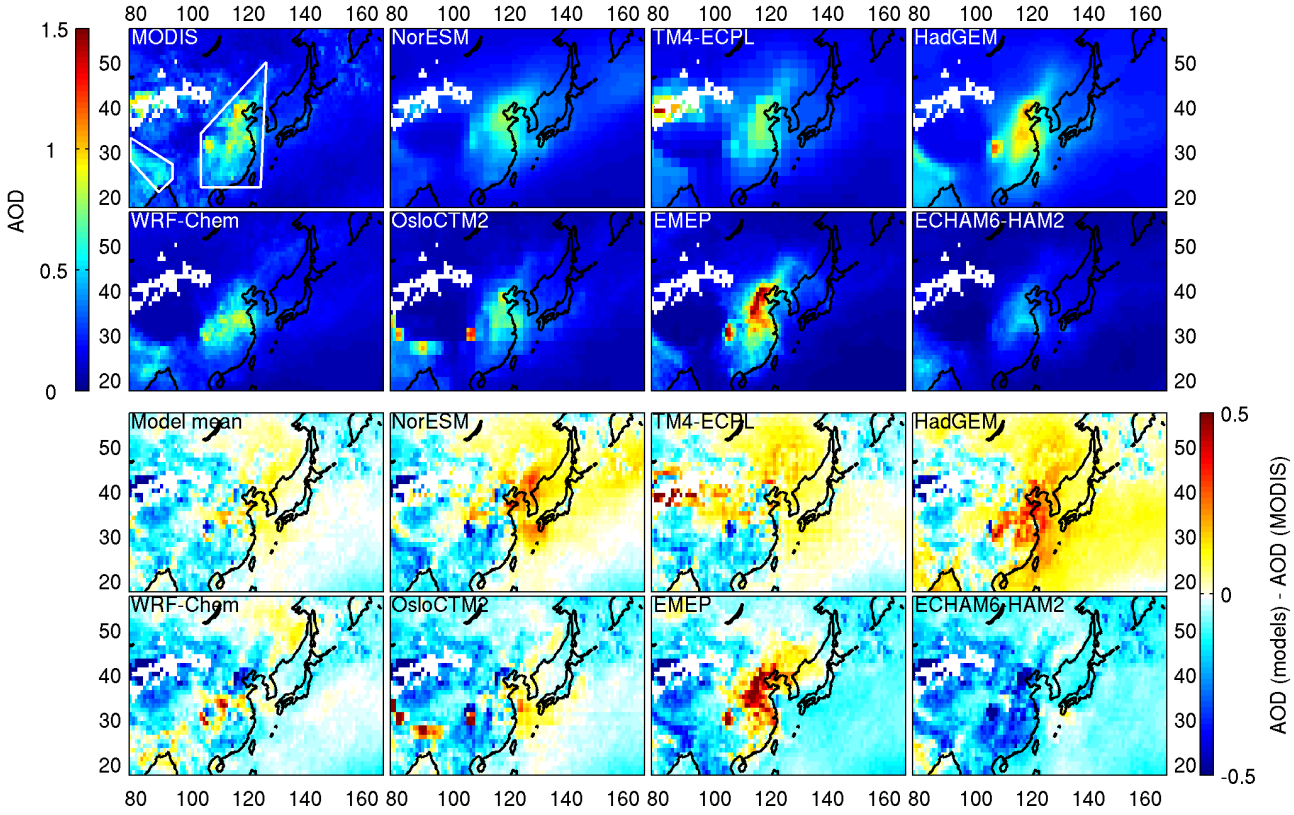

Figure 8. (8 top panels) Mean AODs observed by MODIS and simulated by the ECLIPSE models for August and September 2008, and (8 bottom panels) absolute differences between the simulated and the MODIS AODs (models names are given in each relevant panel). White polygons mark out the two regions discussed in the text.

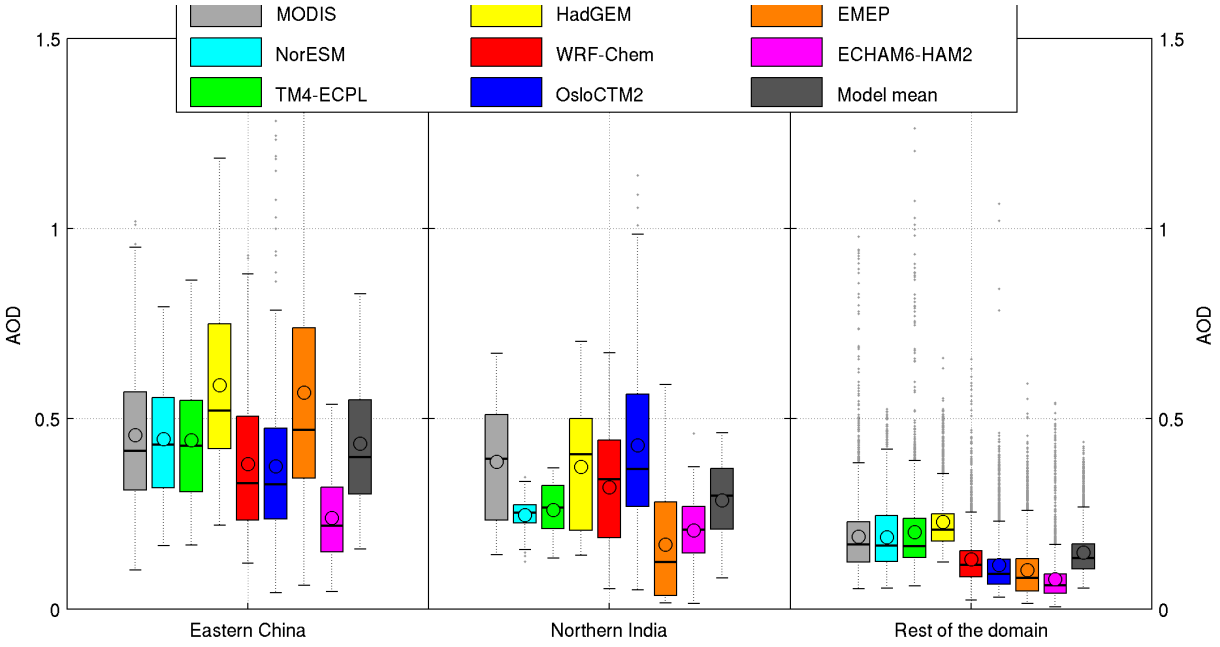

Figure 9. Box plots showing the mean AODs (circle), median (central line), 25^{th} and 75^{th} percentiles (box edges), and the extreme data not considered as outliers (whiskers) during August and September 2008 as observed by MODIS and simulated by the ECLIPSE models over northern India, eastern China and the rest of the domain.

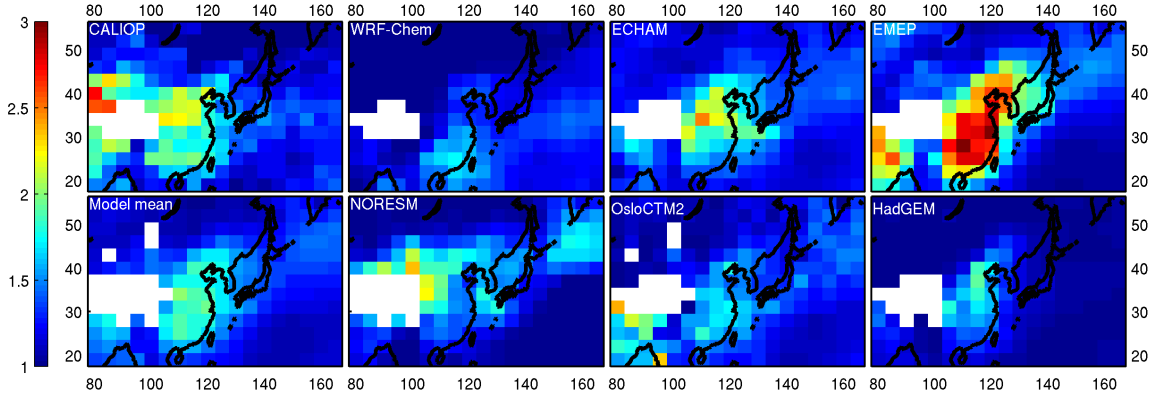

Figure 10a. Comparison between R_{app} over Asia in August and September 2008 averaged over a 0-2 km layer derived from CALIOP data and simulated by the ECLIPSE models. White boxes indicate missing observations due to ground elevation.

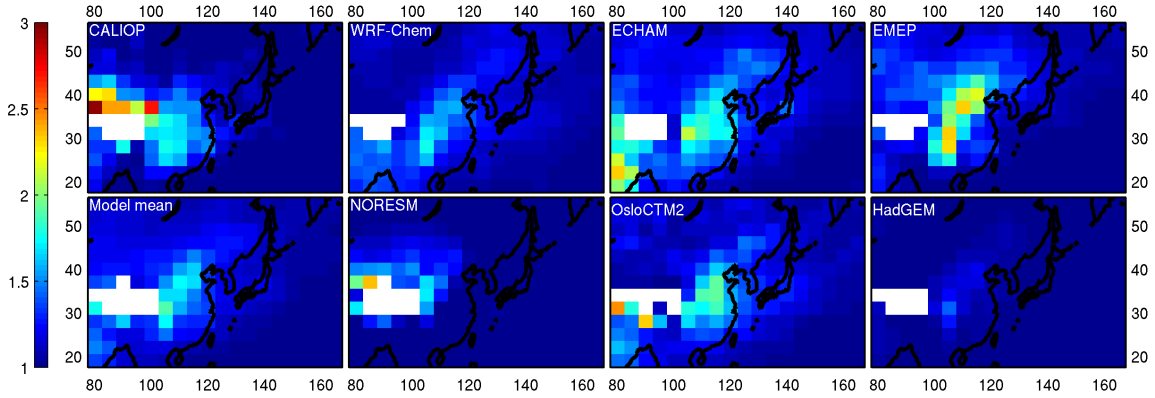

Figure 10b. Comparison between R_{app} over Asia in August and September 2008 averaged over a 2-4 km layer derived from CALIOP data and simulated by the ECLIPSE models. White boxes indicate missing observations due to ground elevation.

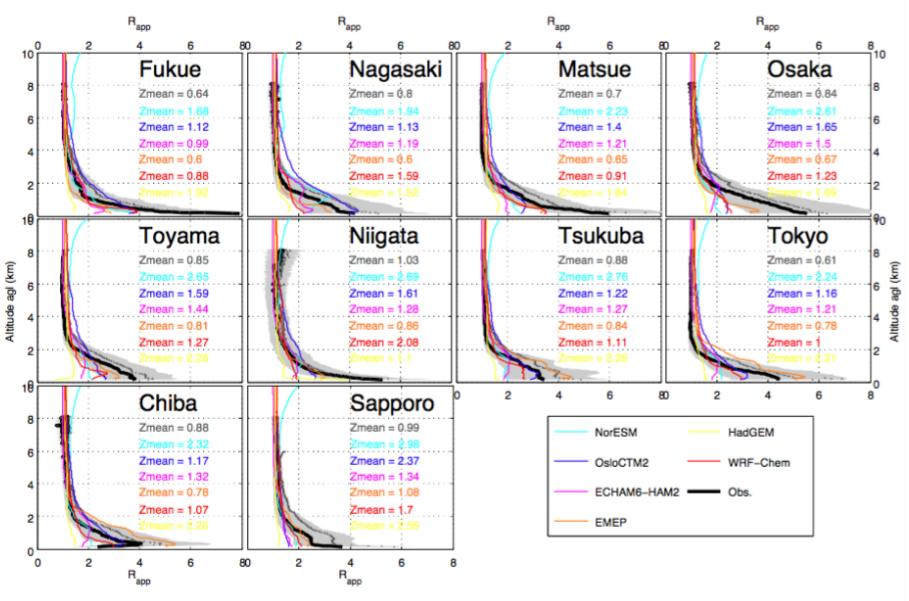

Figure 11. Comparison of mean (grey dots), median (black line), 25^{th} and 75^{th} percentiles (grey area) of R_{app} profiles observed at 10 NIES aerosol lidar stations over Japan (location shown in Fig. 1) and mean R_{app} (colored lines) simulated by the ECLIPSE models.

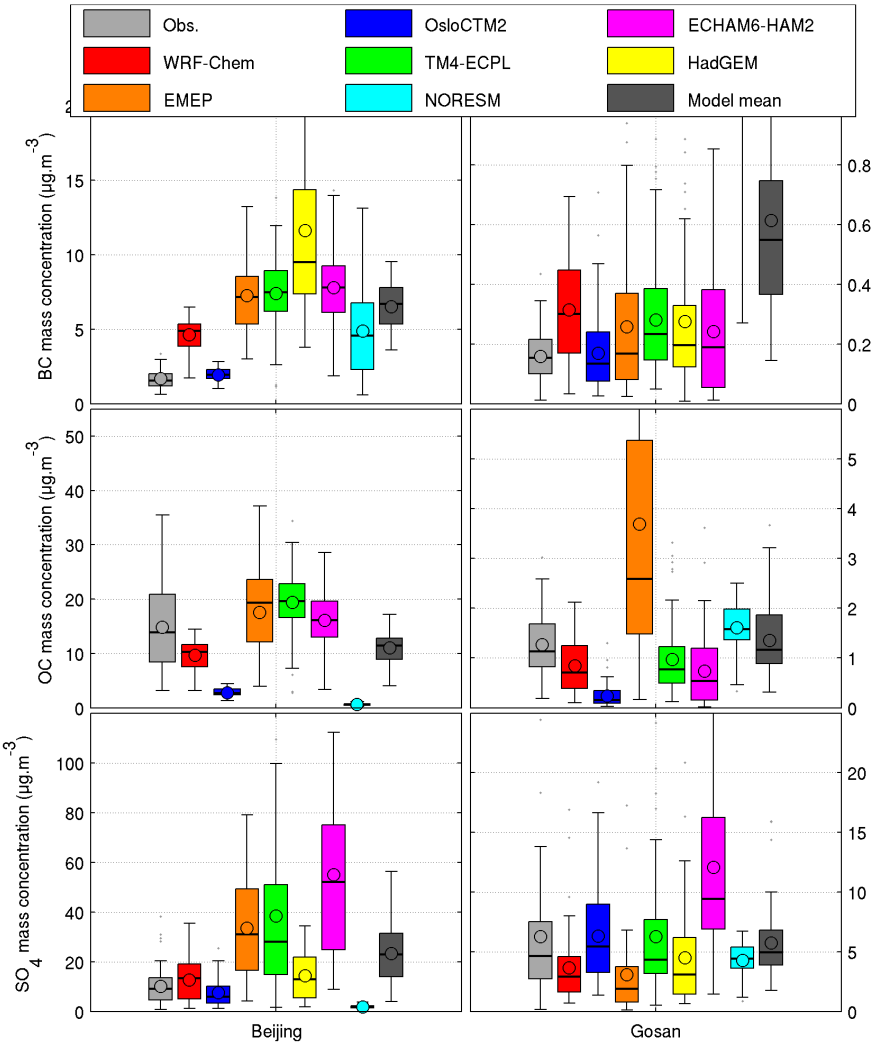

Figure 12. Comparison of observed and simulated box and whisker plots for BC, OC and sulphate: mean (circle), median (central line), 25^{th} and 75^{th} percentiles (box edges). The whiskers encompass values from $25^{th} - 1.5 * (75^{th} - 25^{th})$ to the $75^{th} + 1.5 * (75^{th} - 25^{th})$. This range covers more than 99% of a normally distributed dataset. Note the different scales between urban and rural stations for ~~NO_2~~ NO_2 and ~~SO_2~~ SO_2 panels.