Authors' response to the comments by Anonymous Referees #1 and #2

We would first like to thank both referees for their useful and constructive comments. In order to stimulate discussion between the referees and the authors we have combined the response to both Anonymous Referees. The responses have been grouped by topic to improve readability.

Approach used

Ref#1: Referring to this technique as 'modeling the sub grid scale variability' is misleading because there is nothing in the formulation of the CTM that has been changed here to actually model the unresolved variability. In my view, what they develop is a post-processing, downscaling technique to map regional scale simulations on finer resolution emission data using some parametrization to account for dispersion of these emissions.

Ref#1: The parametrization of emission dispersion does not include any sub grid process such as fast chemistry or deposition. What is more, the time scale of their application is much coarser than the CTM's (annual vs. several minutes though the time step of the EMEP model is not explicitly mentioned). Those choices make it clear, in my opinion, that this effort is not meant to solve the subgrid variability. To do so, the effort would rather focus on high resolution meteorology and emissions on line with the regional CTM to capture the unresolved features of atmospheric chemistry and dispersion.

Authors' response: *We appreciate the referee's opinion on this matter. It is a question of definitions. The referee's interpretation of the term "sub-grid model" appears to be that of a parameterisation within a coarse resolution model (e.g. a CTM) that explicitly takes into account sub-grid variability during the simulation. We use the term "sub-grid model" to mean a parameterisation that can refine the outputs of a coarse resolution model, by providing an estimate of the spatial distribution of the model output (e.g. atmospheric concentrations) within each grid cell of the coarser model. In this respect we a using the term "sub-grid model" in the same sense as that used by various publications cited in the manuscript (Ching et al., 2006; Denby et al., 2011; Dragosits et al., 2002 and Isakov et al., 2007). The terms 'sub-grid model' and 'sub-grid variability' will be defined in the revised manuscript to avoid confusion.*

Ref#2: I have difficulties understanding why NO2 photochemistry is not needed here. By using the NOx fine-scale emission pattern and distribution kernel for redistributing NO2 concentrations this approach assumes a constant NO2/NOx concentration ratio in the whole EMEP 50km grid cell, which is hard for me to believe. Typically this ratio should show a dip close to NOx sources due to the high NO ratio in primary emissions. Also this fraction of primary NO2 in NOx emissions varies strongly between different sources and this should play a role at the local scale. I thought that at least ADMS includes a NOx photochemistry formulation, why is this not used? Ideally such a chemistry scheme should be used, or at least the potential errors discussed. Along similar lines, also the formation of NH4 can influence local NH3 concentrations – it would be interesting to know what the potential errors are when these processes are ignored.

Authors' response: *The assumption of a constant NO_2/NO_x concentration ratio is made to provide a simple parameterisation that is universally applicable. It would be difficult to include a variable ratio in a simple model since it would depend on the local pollution climate (e.g. ozone concentrations), photochemical reactions of NO and NO_2 emitted from the different source types and complex variations in diurnal emission patterns and meteorological drivers. Both AERMOD and ADMS include optional*

simple photochemistry schemes but they depend on background ozone concentrations, which are not homogeneous across the domains, and so these model options cannot be used for developing a simple generic model, which was the aim of the study.

Of course this simplification will introduce uncertainty into the model. This uncertainty can be estimated by analysing the variability of the NO_2/NO_x ratio in the measured data. For the Scottish domain, the coefficient of variation of the ratio is 20% whereas for the Dutch domain it is only 7.5%. Estimating the NO_2 concentrations from the measured NO_x concentrations, assuming a constant ratio, gives mean errors (NMGE) of 18% and 6.3%, for the Scottish and Dutch domains, respectively. Extending this analysis to the annual mean concentrations for the sites in the European air pollution database "airbase" that simultaneously measure NO and NO_2 (1478 sites), gives a coefficient of variation of 17.5% and a mean error of 14.6%.

Similarly the formation of NH_4^+ from NH_3 depends on the concentrations of other pollutants, as well as the complex variations in diurnal emission patterns and meteorological drivers, the inclusion of which is beyond the scope of a simple sub-grid model.

A discussion of these uncertainties will be included in the revised manuscript

Ref#1: How would results look like if no dispersion was taken into account and the same process was done only by using the 1km (or 7km or both) emission proxy?

Authors' response: *This is an interesting question and should be fairly easy to test. We will include this additional analysis in the revised manuscript*

Meteorological data

Ref#2: I find it hard to understand why meteorological data from just one station was used here, instead of grid-specific meteorology consistent with the fields driving the CTM. Even the two "domain-specific" meteorological data sets come only from one station each. I think this needs to be better justified and compared to the effects a grid specific met data set would have.

Ref#2: Since the authors mention this as a possibility I would suggest that they add a sensitivity analysis using grid-specific meteorology consistent with the CTM

Ref#1: Wouldn't it make more sense to use the same meteorology as in the EMEP model at least for this sensitivity test?

Authors' response: *The model was developed using a single meteorological dataset to provide a simple parameterisation that could be of benefit to the wider air quality modelling community. The model was tested with meteorology specific to each modelling domain to show that the model results were not that sensitive to the meteorological data set used. The use of grid-specific met data would require nearly 800 ADMS and AERMOD simulations (110 grid cells for 7 emission heights) and 110 LADD simulations (one for each grid cell). However, before revising the manuscript we will assess the feasibility of carrying out this additional analysis.*

Ref#2: Why is it justified to assume rotationally symmetric dispersion kernels, should not the local predominant wind direction have a significant influence? How big is the error introduced by this rotationally symmetric formulation?

Authors' response: *In order to keep the model parameterisation simple and universally applicable, we decided to remove the influence of local wind direction distributions (which are not valid for the whole domain) by assuming rotationally symmetric dispersion kernels. It can anyway be noted that in the real world local wind-directions will differ from that of a larger scale model. For example, cities are frequently located in locations subject to topographic (e.g. valley) winds, or sea-breezes.*

Although this rotation symmetry obviously does introduce some error, the results shown in Figs. 4 and 5 demonstrate that even with this assumption the sub-grid model provides a valuable improvement over that of the larger scale CTM. This assumption will be discussed in the revised manuscript.

Emissions

Ref#1: I am wondering what do emission sources as large as 1km2 could possibly represent. In my understanding, dispersion models are conceived to represent emission from point sources such as industrial stacks. Is this the right model to represent large area sources such as crops or residential emissions? Is this type of modeling adequate to represent dispersion around busy roads? Don't dispersion patterns depend on the emission sector?

Authors' response: *The use of 1 km^2 emission sources is a simplification due to the fact that we do not have emission inventories for the study domains at a higher spatial resolution. So in reality we may have a large point source within a specific 1 x 1 km^2 grid cell but since we do not know exactly where, we distribute this emission over the whole grid cell. Of course this is a simplification which is going to introduce uncertainty in the model but it is the best we can do with the available emission data. The model, therefore, would not be expected to be suitable for estimating concentrations close to busy roads but the model performs well for the Dutch traffic sites. This will be made clearer in the revised manuscript.*

Averaging period

Ref#1: It would be interesting to look at the effect of the meteorological dataset at a finer time-scale. Especially since the authors claim in their conclusions that this method is easily applicable at finer time scales.

Authors' response: *As stated in the manuscript, the annual mean limit values for NO_2 are generally more stringent than the hourly ones and impacts of NH_3 are only assessed using annual mean concentrations. We agree that it would be interesting to develop a model with a higher temporal resolution, but such a study would raise other issues (e.g. as discussed above, or with the availability and accuracy of time-resolved emissions). Thus, we did not consider this to be 'easily applicable' at finer time-scales, but rather, as we state in the discussion, this is a potential future improvement and is out of the scope of the presented work.*

Model evaluation

Ref#1: I don't think it is appropriate to say that "the sub grid model performed better than the EMEP model". It would be more fair to say that the downscaled version of the
EMEP model compares better with observations.

Authors' response: *This is a fair point and the manuscript will be modified to reflect this.*

Ref#2: Does this mean the downscaling decreased agreement to NO2 non-traffic stations, which are the ones we would actually expect it to improve? As mentioned above, I would not expect the 1x1km model to be representative of traffic stations.

Authors' response: *You would expect the sub-grid model to perform best for the non-traffic stations but for the Dutch domain it had the lowest bias and error for the traffic sites. This suggests that the model may be suitable for simulating concentrations at traffic stations, providing they are located several metres from the roadside (as is the case for most of the Dutch sites), and the good agreement is likely the result of the fact that the majority of roadside sites are not situated along isolated motorways, but rather embedded in urban areas with dense road networks. A discussion of this will be included in the revised manuscript.*

Manuscript structure

Ref#2: I would find it helpful if the downscaling equations were written out as equations and not only described verbally

Authors' response: *This is a good suggestion and the downscaling equations will be included in the manuscript.*

Miscellaneous comments

Ref#1: The inevitable question arises of whether a direct EMEP run at 7km resolution with its corresponding meteorology would bring about the same improvement as the downscaling developed in the present study. And in this case the data would be directly at hourly resolution.

Authors' response: *Although it is currently possible to run the EMEP model at a 7 km resolution for specific projects, they are far too CPU intensive for routine use, and especially where CTMs need to be run 10s-100s of time for emission control assessments. However even at a spatial resolution of 7 km, the EMEP model would not be able to resolve the large horizontal concentration gradients found close to sources and a sub-grid treatment may still be necessary. A discussion of this will be included in the revised manuscript.*

Ref#2: One issue that is not addressed here nor anywhere else in the manuscript is that (to my understanding) the lowest vertical layer of the EMEP model extends from the surface to about 90 meters. In this context, it would be important to explain what is meant by 'the mean atmospheric concentration in each grid square': Is this the estimated surface concentration calculated by applying some standard vertical distribution? But if so, is it then justified to assume that the modelled mean concentration on the 50km grid is 'correct' and just needs to be re-distributed spatially within the grid cell?

Authors' response: *The EMEP model estimates the near surface concentrations by extrapolating the concentration at the mid height of the first vertical layer (45 m) assuming an approximately constant vertical deposition flux. Of course, there are uncertainties in this procedure, especially for NH_3 where deposition velocities can be relatively high, but where bi-directional exchange complicates even the sign of atmosphere-biosphere exchange in real situations. However, we have to work with the EMEP values as provided by the model, and our working assumption is that the errors in the vertical distribution are less important than the errors we seek to minimise – namely those of sub-grid horizontal variability. This will be clarified in the revised manuscript.*