

Responses to Reviewers, “Turbulent Transport, Emissions, and the Role of Compensating Errors in Chemical Transport Models”

Reviewer 1: William Stockwell.

We thank the reviewer for his time and effort in his very complimentary review.

Reviewer 2: Douw Steyn.

We thank the reviewer for the complimentary review – we hope that the paper will be useful to the regional modelling community. The contrast between the models’ performance and the critical role of both emissions timing and the strength of diffusion suggest that there is room for improvement of these aspects of air-quality models within the broader community, and we hope that this paper helps bring that work about.

With regards to the specific comments made:

- 1) *Pages 5599 and 5600: The term “science processes” is a very strange one. The authors are clearly not referring to the processes of science as a scholarly activity, which is what the term literally can only mean. What they are referring to is physical and chemical processes captured by the computational models.*

We have modified that sentence to read, “The study noted the potential difficulties in the systematic assessment of individual chemical and physical processes on the model results, due to the complexity and interconnected nature of those processes.”

- 2) *Page 5601, line 4: The use of UTC here is odd. All processes referred to in the study are more closely related to local time. Later in the manuscript the authors use LST. I suggest that should be used here.*

The UTC time was used out of habit: the times for which meteorological analyses are created are on a 6 hourly UTC schedule – though we agree that this is confusing with regards to the rest of the paper. We have modified the text to read “... 16:00 local time (00:00 UTC) on each day”.

- 3) *Page 5602, line 6: The “in” in this line is redundant.*

This has been corrected.

- 4) *Page 5606, line 17 and Figure 1: The proper name is Vancouver International Airport. All 5 monitoring stations in the LFV should be named in the caption to Figure 1.*

The name in the text has been changed, and the station names have been added to the caption of Figure 1.

- 5) *Page 5607, line 3: "...concentrations of which were higher in..." appears to be aprt of an incomplete sentence.*

The sentence "In an urban region at night, the dominant ozone chemical process is usually the destruction of ozone through titration by NO, the predicted surface concentrations of which were higher in AURAMS than in CMAQ." has been modified to read, "In an urban region at night, the dominant ozone chemical process is usually the destruction of ozone through titration by NO. The predicted surface concentrations of NO were higher in AURAMS than in CMAQ."

- 6) *Page 5608, lines 16 & 17: The term "higher level(s)" appears to refer to higher concentration levels of ozone, but could easily refer to higher (altitude) levels in the model.*

The words "level" and "levels" have been changed to "concentration" and "concentrations" to make this more clear.

- 7) *Page 5613, line 16: "..to which that diurnal emissions.." appears to be either incomplete, or contain a redundancy.*

The excess "that" has been removed.

- 8) *Page 5614, lines 15 to 17 and Tables 4b and 5: The sentence "Normal font base cases" is redundant at this point in the text, but should be incorporated into the captions to Tables 4b and 5.*

We have removed the redundant text. Note that Table 4(a) has the description of the font meanings: we have added "Font description as in Table 4(a)." to the caption for Tables 4b and 5, for clarity.

- 9) *Page 5620, line 13: "differences" should be "difference".*

Corrected.

- 10) *Table 3: The definitions in this table are standard, and well-known. The table could be deleted without detracting form the paper.*

While we agree with Dr. Steyn on the definitions being well known, we would prefer to leave the explanatory table in the paper. While the definitions are well known to most modellers with experience in the field, they may be less so to new modellers (students, etc). Also, the inclusion of the table makes our definitions absolutely clear (e.g. we are using R, not R^2 , etc.).

Reviewer 3: Jonathan Pleim

We thank the reviewer for his detailed review, particularly his comments and suggestions regarding updates to the treatment of minimum eddy diffusivity in more recent versions of CMAQ. These have led to several additional model simulations, and significant modifications to our original manuscript. We have responded to his comments and concerns as follows:

Re: *“Meteorology models typically have very low eddy diffusivity minimum values (e.g. 0.01 m²/s) for numerical considerations. Using values that are greater than the eddy diffusivity values computed by the PBL schemes during stable conditions is unjustified.”*

We agree with Dr. Pleim with one important caveat – if the PBL schemes used in the meteorological model underestimate the strength of actual turbulence, then the use of a value greater than that set in the meteorological model’s PBL scheme, within the air-quality model, is justified. Harmonizing with the meteorological model makes sense, provided that the result improves or maintains the accuracy of both meteorological and chemical variables. Our concern is that an improved representation of the strength of mixing (whether achieved through a cutoff or an improved PBL scheme) may not be to account for all of the model error, and that the older methodologies for dealing with this issue may have hidden other causes of model error.

“Thus it is reasonable to conclude that if air quality models produce better results with the greater minimum eddy diffusivity values then it is compensating for other errors.”

This is not what we found nor what we concluded; e.g. it could just as easily mean that the meteorological model’s eddy diffusivity is insufficiently high (and perhaps the meteorological variables are less sensitive to eddy diffusivity than the air-quality variables). Once the most accurate picture of turbulence is achieved, other factors must be considered for further model improvement. Our results suggest that the magnitude of emissions at night plays an important role in achieving accuracy for both O₃ and PM_{2.5} predictions, a role that was masked when a simple cutoff in minimum diffusivity was used.

“While it is clear that this is the main thesis of this paper, the point that the greater minimum eddy diffusivity values used in some models, typically older models, is scientifically unjustifiable should be made more strongly from the start. I think this point has been recently recognized in the modeling community resulting in the reduction in these artificially high minimum values. This historical and evolving perspective should be clearly presented.”

Our main thesis from the paper is not (only) that the use of a high minimum diffusivity is inappropriate – rather, that higher turbulence levels (regardless of their origin in either a turbulence scheme or a minimum cutoff) can only explain part of the model error. The use of higher turbulence levels will not improve both primary and secondary chemical outputs at the same time. At the same time, the use of a higher level for the cutoff in vertical diffusivity may be justified, for cases where the meteorological model is known or suspected to provide insufficient turbulence, relative to the ambient atmosphere. For example, recent observations and simulations summarized by Leroyer et al (2013) for the Vancouver area suggest that model resolutions at or better than 250m may be necessary to adequately resolve turbulent processes in urban regions. Parameterizations such as lower limits in diffusivity may be necessary when lower resolutions are employed, as was the case in our study.

In our revised paper, we have included several additional simulations. We found that the more recent versions of CMAQ made use of additional meteorological variables which were not available in the meteorological inputs from the original study, precluding its use here due to the time constraints of the review process. We were, however, able to evaluate the diffusivity lower limit methodology used in the more recent versions of CMAQ, via incorporating that methodology within the AURAMS model and repeating the simulations. We also included three additional simulations, described in the revised manuscript and outlined later in this response.

We have clarified these issues through the following rewording of part of the abstract to read:

“A process analysis of the models revealed that many of the differences between the models’ results could be attributed to the strength of turbulent diffusion, via the choice of an *a priori* lower limit in the magnitude of vertical diffusion coefficients, with AURAMS using $0.1 \text{ m}^2/\text{s}$ and CMAQ4.6 using $1.0 \text{ m}^2/\text{s}$. The use of the larger CMAQ4.6 value for the lower limit of vertical diffusivity within AURAMS resulted in similar performance for the two models (with AURAMS also showing improved $\text{PM}_{2.5}$, yet degraded O_3 , and a similar time series as CMAQ4.6). The differences between model results were most noticeable at night, when the higher minimum turbulent diffusivity resulted in an erroneous secondary peak in predicted night-time O_3 . A spatially invariant and relatively high lower limit in diffusivity could not reduce errors in both O_3 and $\text{PM}_{2.5}$ fields, implying that other factors aside from the strength of turbulence might be responsible for the $\text{PM}_{2.5}$ over-predictions. Further investigation showed that the magnitude, timing and spatial allocation of area-source emissions could result in improvements to $\text{PM}_{2.5}$ performance with minimal O_3 performance degradation. AURAMS was then used to investigate a land-use-dependant lower limit in diffusivity of $1.0 \text{ m}^2\text{s}^{-1}$ in urban regions, linearly scaling to $0.01 \text{ m}^2\text{s}^{-1}$ in rural areas, as employed in CMAQ5.0.1. This strategy was found to significantly improve mean statistics for $\text{PM}_{2.5}$ throughout the day and mean O_3 statistics at night, while significantly degrading $\text{PM}_{2.5}$ correlation coefficients and slope of observed to model simulations. The use of this latter strategy was also shown to reduce the model’s sensitivity to the temporal allocation of its emissions inputs. The modelling scenarios suggest that while turbulence plays a key role in O_3 and $\text{PM}_{2.5}$ formation in urban regions, and in their downwind transport, the spatial and temporal allocation of primary $\text{PM}_{2.5}$ emissions also has a potentially significant impact on $\text{PM}_{2.5}$ concentration levels. The results show the complex nature of the interactions between turbulence and emissions, and the potential of the strength of the former to mask the impact of changes in the latter.”

We have also expanded the CMAQ model description in Section 2 to acknowledge the CMAQ developments post-v4.6, and our additional simulations, with the following additional text:

“More recent versions of CMAQ (5.0.1) use a bilinear interpolation in urban land use area fraction for the lower limit of eddy diffusivity, with $0.01 \text{ m}^2\text{s}^{-1}$ employed for entirely rural areas and $1.0 \text{ m}^2\text{s}^{-1}$ for entirely urban areas. These recent changes to CMAQ were implemented in recognition that the rural minimum should not exceed that of the driving meteorological model, whereas the use of a higher urban minimum may be necessary if the driving meteorological model is not capable of accounting for the turbulence-enhancing effects of the urban environment. Here, the impact of these strategies was investigated in a set of scenario simulations.”

Thus, I don’t think that it’s acceptable to use CMAQv4.6 in this study since a main object of the study, the minimum eddy diffusivity, has been drastically revised since then.

Unfortunately, users are not always able to migrate to the latest version of a model for various reasons, such as the need to maintain consistency with previous simulations (e.g., policy applications), the need to meet challenging protocols for migrating to newer versions, or outside constraints due to the limitations or resolution of the meteorological model employed for the simulations. For example, CMAQ v4.6 is still in use as the operational air quality forecast model of the U.S. National Weather Service (P. Lee, personal communication, January 2014), and makes use of a different meteorological driving model than WRF. Performance evaluations of CMAQ v4.6 were also reported in previous studies such as Appel et al. (2007, 2008) and Eder et al. (2007). The results from the present study with v4.6 can provide insights for retrospective understanding and interpretations of these previously published results. The work has value to those currently using version CMAQ 4.6, in that they may wish to upgrade either the model version, the meteorological driver and/or its components, or the v4.6 strategy for minimum diffusivity, based on these findings.

It is clearly not representative of the comparative performance of the two models to use such an outdated version of CMAQ compared to a more recently updated version of AURAMS. The sensitivity of model results to the magnitude of the minimum eddy diffusivity is sufficiently demonstrated through the comparison of the AURAMS scenarios, particularly AURAMS5 where the minimum eddy diffusivity is set to 0.6 m²/s. The comparison to CMAQ seems to be superfluous to the main objectives of the paper. I suggest the CMAQ part either be removed or the CMAQ results be replaced with runs using CMAQv5.0.1, which is publically available at www.cmascenter.org.

Please see our above response. We investigated the possibility of using CMAQ5.0.1 to repeat simulations during the review process – unfortunately, this would require additional meteorological variables (and, ideally, a different meteorological model for air-quality model input). This was beyond the scope of our work. However, we were able to investigate the impact of the CMAQ5.0.1 strategy for minimum diffusivity in additional AURAMS sensitivity simulations, and our revised paper reports on those findings. We also note that our previous work (tabulated statistics columns for “AURAMS1 Base Case versus AURAMS1b Code Improvements” show that the changes associated with those AURAMS code improvements were relatively minor compared to the effect of the changing the minimum in eddy diffusivity (“AURAMS 5” simulation). Any “gains” from the more recently updated version of the AURAMS model were shown to be minor in the initial submission – indeed, many of the statistics for PM_{2.5} show a degradation of performance with those updates (see Table 4b).

The strategy of an urban land-use-fraction-dependent lower limit in diffusivity was investigated in additional AURAMS simulations, described in our revised manuscript in section 4.2.4, with simulation results described in Section 4.3.5. We also added a time-based evaluation of the model, where the model statistics relative to observations were calculated at common UT hours across the 35 day simulation (section 4.4). These simulations and analysis showed that the use of the land-use-fraction strategy improved PM_{2.5} mean statistics while degrading PM_{2.5} correlation coefficients and slopes, though they also improved O₃ statistics at night, while degrading some O₃ statistics during the day. The vertical extent of over which the lower limit was employed was also examined – imposing a height limit equivalent to typical urban mixed layer heights in large cities and in the region studied had little impact on these results, indicating that the land-use dependent lower limit of diffusivity need only be applied for the urban mixed layer.

As we've noted in our revised manuscript's conclusions, these additional runs show that the relative impacts of turbulence and emissions timing and spatial allocation are complex. Accurately simulating the strength of turbulence in all three dimensions is essential. At higher levels of turbulence, temporal variations in primary PM_{2.5} emissions have a lower impact on model results, while in more stable conditions, these temporal variations in emissions may be critical, and if improved could result to significant improvements in model PM_{2.5} performance without degrading O₃ performance. Conversely, if model turbulence is too strong relative to observations, these emissions effects may be masked.

We do, however, agree that there is a risk that readers of the paper might assume that the performance of more recent versions of CMAQ would be the same as the one version used here, and we definitely don't want that to be the take-home message of the paper. To that effect, we've modified the text to note that the version of CMAQ used here is an older code, and more recent CMAQ versions may perform differently than the one used in our work (including the abstract):

- The version numbers of both CMAQ and AURAMS have been included in the abstract and the conclusions
- The abstract and conclusions have been revised to include the results of the additional simulations (abstract noted above).

The main point of the paper that the nocturnal overpredictions of PM_{2.5} when lower values of eddy diffusivity are used are largely attributable to the diurnal allocation of emissions seems too simplistic. While the models are clearly sensitive to this as shown in this study, there may be many other factors contributing to these concentration errors.

We did not make the claim that the nocturnal overpredictions of PM_{2.5} that occurred when lower values of eddy diffusivity were used are largely attributable to the diurnal allocation of emissions – that is not one of our conclusions. We also acknowledged in the other manuscript that other model parameterizations might make a difference, and in the revised manuscript that urban turbulence parameterizations are also critically important. Quoting from the abstract of the original manuscript (underlining added here):

“...the magnitude, timing and spatial allocation of area-source emissions could result in improvements to PM_{2.5} performance with minimal O₃ performance degradation. While the strength of turbulence plays a key role in O₃ and PM_{2.5} formation, more accurate primary PM_{2.5} temporal emissions data may be needed to explain observed concentrations, particularly in urban regions.”

In the Discussion section of the paper, we said

“(2) The hypothesis that at least some of the PM_{2.5} prediction errors may result from errors in the emissions inputs has some merit”

We also acknowledged in the Discussion section that other factors may be worth examining:

“(3) There may also be other factors which may act to reduce ‘effective’ PM_{2.5} emissions.”

The example given there was correction factors for PM_{2.5} area source emissions which are applied to some area source emissions but not others as a possible means of reducing PM_{2.5} concentrations. In our

revised manuscript, we have also mentioned the particle deposition algorithms as a source of possible differences, at this point. The revised text reads

“(4) There are other factors which may act to reduce PM_{2.5} concentrations aside from temporal and spatial allocation.”

Our points (5, 6, and 7) in the Discussion also provide caveats and note that there are other processes at work.

Most of the emission modifications seem to be corrections that should be made anyway.

We agree with the reviewer that these emissions modifications should be made. However, our point is that that they have not been made or investigated in current model emissions databases, and these issues with emissions timing have not been identified, prior to this study. Hence we feel that they are worth reporting to the broader scientific community.

The final test (AURAMS6) represents an extreme and unrealistic sensitivity test which does show that the model is sensitive to the diurnal emission profiles, reducing the nocturnal high bias in PM_{2.5} but increasing the frequent daytime high bias.

We disagree with the reviewer’s characterization of this sensitivity test as extreme and unrealistic. This particular sensitivity study (AURAMS6) stemmed from the others, wherein we noted that for many of the non-mobile area sources, a subset of the total model emissions, a time-invariant emission temporal profile was assumed, despite many activities which would be expected to have a diurnal signature. A time-invariant profile is often used as a “default” in emissions processing, if no source-specific information is available. Our question here was “how much of a difference would a diurnal profile for these sources make towards the model results?” – not as a means of justifying that use of that specific profile – but to show that obtaining better information on the diurnal behavior of these emissions is worth doing. To reflect this point better, in our statement in point (2) of the discussion - “further improvements to the accuracy of the PM_{2.5} may be achievable through the collection of improved area-source temporal and spatial emissions data.” – we have changed the word “are” from the submitted manuscript to “may be” here. We have also used the words “non-mobile area source emissions” instead of just “emissions” in the first sentence of section 4.3.4, which analyses the results of this scenario, and in our revised manuscript, limited this change only to the non-mobile primary emissions of PM_{2.5}.

With regards to the reviewer’s phrase “reducing the nocturnal high bias in PM_{2.5} but increasing the frequent daytime high bias” – that was not a result of the original work. We did not examine daytime versus nighttime biases, but rather overall hourly statistics. In our revised manuscript, we show hourly statistics across the model grid as time series in a new pair of figures (13 and 14). This analysis shows that the shift of primary PM_{2.5} emissions from night to day actually results in a consistent improvement in mean PM_{2.5} statistics at all hours, without degrading PM_{2.5} correlation coefficients. The same analysis suggests that employing a urban-land-use-fraction dependent lower limit in diffusivity improves nighttime O₃ performance and PM_{2.5} mean statistics, while significantly reducing PM_{2.5} correlation coefficients and slopes.

In the introduction section there is an extensive review of previous model comparison and evaluation studies. The results of the AQMEII study are particularly well described including model performance of not only PM_{2.5} and PM₁₀ but also particular speciated aerosols such as seasalt, organic, and inorganic

and processes like dry deposition and stable boundary layers. In contrast, the present study seems too limited as is admitted in the discussion section. The focus is totally on minimum eddy diffusivity and emission diurnal temporal allocation. Also, the evaluation is exclusively on ozone and PM_{2.5} whereas analysis of model performance for primary versus secondary PM and primary gasses such as CO and NO_x would probably enhance understanding of model and emission errors.

Yes, the AQMEII study investigated these issues. However, the AQMEII study did not investigate the issues associated with emissions magnitude and temporal allocation, or the impact of assumptions on lower limits for diffusivity. The AQMEII study did not harmonize model grids, meteorological inputs, and all aspects of the emissions data (notably biogenic emissions). Here, we followed the path of a very focused look at two aspects of the modeling systems, that were run in a highly harmonized fashion, in order to eliminate some of the causes of differences between the model simulations.

Ozone and PM_{2.5} were chosen since those are the “most important” endpoints of the model predictions from the standpoint of most regulatory action, and also have the best data coverage (i.e. the most stations, the most hourly values).

Specific comments:

P5598 In21: I think “consistently matched” needs some elaboration.

We have replaced the phrase with a direct quote from the referenced paper and regarding the North American statistics: “no model was found to consistently match the observations for all locations throughout the entire year,” with North American correlation coefficients for PM_{2.5} ranging from 0.34 at a 99% confidence level to less than 0.10 at a 10% confidence level.

P5600 Ins19-25: The CMAQ version is quite out of date (7 years old) while the AURAMS has been recently updated. Is the AURAMS1b scenario the updates describe in Kelly et al (2012)? If so, please summarize what these updates entail.

See our note above regarding the impact of the AURAMS improvements. The updates that constitute the AURAMS1b scenario were described in the original manuscript, see section 4.2.5. Kelly et al gives a very comprehensive background description of AURAMS, hence the reference (Kelly et al describes our “AURAMS1”). Most of the model as described in Kelly et al and Smyth et al dates back to 2006. The updates between Smyth et al and Kelly et al (AURAMS1) include:

- The use of specified, rather than zero gradient lateral boundary conditions (thus bringing AURAMS into line with CMAQ 4.6).
- For O₃, the use of tropopause height to set the ozone boundary condition, as described in Makar et al (ACP, 10, 8997-9015, 2010).
- Changes to the model code structure to allow nesting.
- Flags and inputs to allow the model to use inputs from the different convective schemes in the GEM meteorological model.
- Updates to the secondary organic aerosol yields.
- The inclusion of the KPP RODAS3 gas-phase chemical solver as an optional solver within the code (not used here).
- Options for writing output to subdomains within the model.

P5601 In10-11 and In23: The meteorology data is interpolated to a 12 km grid for the air quality modeling but the type of map projection is not mentioned.

Table 1 first line, describes the horizontal projection used for both models (polar stereographic). Page 5601 of the submitted manuscript describes the projection used for the driving meteorology (rotated latitude-longitude).

P5602 In11-14: This short paragraph is the entire description of the dry deposition calculations of the 2 models. Considering the importance of dry deposition to the concentrations of gasses and aerosols as demonstrated in the AQMEII study, I think there should be much more description, comparisons, and sensitivity model runs included. The reference given for the CMAQ dry deposition is actually a description of the implementation and testing of the Pleim-Xiu land surface model (PX LSM) in the MM4 and not about dry deposition. CMAQ is usually run with WRF using the PX LSM so that the stomatal conductance and aerodynamic conductance as well as land use related parameters such as roughness length, LAI, vegetation fraction, etc, can be read directly from the WRF output and used in CMAQ for calculating dry deposition velocities. Since this study uses GEM for meteorology, the dry deposition calculations cannot follow this usual procedure and therefore should be described here.

A more detailed description of the gas and aerosol deposition algorithm is now incorporated – we thank the reviewer for catching that reference (included in an earlier version due to the mention in its abstract of part II which was to have included deposition). Our revised description in Section 2 of the deposition algorithms used in both models is as follows:

“Gas-phase dry deposition velocities for CMAQ4.6 are computed using an electrical resistance analog model (M3Dry, version 1.8, Pleim et al, 2001) in a pre-processing step in CMAQ’s Meteorology-Chemistry Interface Processor (MCIP) (in Otte and Pleim, 2010). A modified version of the MCIP preprocessor (Smyth et al, 2005, 2006a) was used to convert GEM meteorological files for CMAQ4.6 input. AURAMS gas-phase dry deposition, also uses an electrical resistance analog, and is calculated according to Zhang et al (2002). CMAQ uses the particle dry deposition scheme of Giorgi (1986); AURAMS uses the scheme of Zhang et al (2001), which is also based on Giorgi (1986).

CMAQ4.6 incorporates particle deposition as a boundary condition on vertical diffusion (Binkowski and Roselle, 2003), with the species mass in each of three modes being deposited separately, and no impaction term for the coarse mode particles, following Binkowski and Shankar (1995). CMAQ4.6 does not incorporate particle settling. AURAMS1.4.2 calculates particle settling for each of the 12 size bins of the model. In our AURAMS1 (base case) simulations the settling and deposition velocities are used to calculate mass fluxes between layers, with the settling (or deposition) velocity being used to determine the destination layer of the falling particles. In our remaining AURAMS simulations, a 1-dimensional semi-Lagrangian advection approach is taken, with the settling and deposition velocities being used to determine the mass transport and new vertical distribution of the particle mass. The latter modification had a minor impact on model results.

Both models made use of the same land use parameters provided by the GEM model, and both models made use of the same emissions inventories “

In the above change, we noted that the CMAQ literature does not seem to include particle settling as a transport term. Under most circumstances this will have little impact on the net mass, but when relative humidities are high, this term can become significant. If particle settling is included in CMAQ, we would be happy to include references that the reviewer provides.

P5604 In12-13: To make the model evaluations more comparable, the hourly average output from CMAQ (ACONC output file) could be used with the AURAMS hourly averages.

We are glad the reviewer caught this! ACONC was what was actually used in the comparison, and the word “instantaneous” is in error, here. The term “instantaneous hourly” has been changed to “hourly averages” in the revised manuscript.

P5605 In 9: Perhaps this is a good place to explain what updates went into AURAMS1b. Section 4.2.5 could be moved to this location.

We’ve added a line to the text at this point mentioning that the updates are described in more detail in section 4.2.5.

P5606 In1: Calling these results a “marked difference” from the earlier study (Smythe et al. 2009), seems to be an overstatement. Both studies showed that “AURAMS outperformed CMAQ” for most statistical metrics. The bigger differences seem to be with the even earlier studies (Smyth et al., 2006 and Steyn et al., 2013) when CMAQ ozone performed better. Some explanation of these various result would be nice.

Fair enough – we have changed the phrase “marked difference from” to “in contrast to”, since the correlation coefficient was the only O₃ statistical parameter where CMAQ was doing better than AURAMS for Smyth et al (2009). Note that the domains, emissions, time periods and duration of simulations all differ between these references. The Steyn et al (2013) simulations were for shorter time periods (2 to 3 days) than used in our study, and were chosen to be representative of specific synoptic meteorology patterns. Smyth et al (2006) simulated for a 12 day time period using emissions for 2001, and made use of a much smaller model domain than used here. NO_x emissions were much higher in the 2001 inventories than in the inventories for later years, possibly accounting for some of the differences. Different meteorological drivers were used for the different studies. Differences like these highlight the value of trying to harmonize model inputs, as we have done in our study; in order to reduce the number of sources of variability. The end of that paragraph has been modified to read:

“Previous work with CMAQ for simulations in the lower Fraser Valley region for a 12 day period in August of 2001 had significantly better O₃ performance for NMB and NME than found here (Smyth et al, 2006b: 13% and 51%, respectively, versus 75% and 82% in the current work). CMAQ simulations by Steyn et al (2013) for the region for specific short episodes in 2006, 2001, 1995, and 1985 reported NME values ranging from 43% to 79% (compare to 53 to 81% in the different simulations of the current work) and NMB from -12% to 64% (compare to 31 to 75%). Different meteorological drivers, emissions inventories, and domains were used for these studies compared to the more recent work, and these may account for some of the differences in statistics, as well as the shorter time periods used in these earlier studies”

P5608 In4-6: This is not an accurate statement. The use of the higher minimum eddy diffusivity is not to compensate for “inability to resolve turbulence at smaller scales”. Representing the effects of subgrid turbulence is the purpose of eddy diffusion schemes. This statement implies that they fail in their purpose. I would say that higher minimum eddy diffusivity is sometime used to compensate for specific inadequately modeled effects such as urban heat island as is currently done in CMAQ where a higher minimum eddy diffusivity is used in urban areas.

Fair enough, though we would add “complex urban topography” to “urban heat island”. We have modified the sentence to read, “The use of a higher level of diffusion than predicted by meteorological models is intended to compensate for specific inadequately modeled aspects of turbulence and transport, such as subgrid-scale flows through complex urban topography and turbulence induced by urban heat islands.”

P5608 In6-9: I also take issue with this statement concerning the study presented by Pleim and Gilliam (2012). This study used CMAQv5.0.1 where the minimum eddy diffusivity is 0.01 m²/s except in urban areas where it ramps up to 1.0 m²/s. Thus, the values considered in that study over the great majority of grid cells are much lower than are being discussed in this paper. The purpose of that study was to see if the high values in the urban areas could be removed when a simple urban surface parameterization is used. While the results showed marginally worse performance when the high values in the urban areas were removed, there is hope that improved urban schemes will eliminate this need.

The sentence has been changed to “Recent work with CMAQv5.0.1 using a linearly interpolated cutoff (between 0.01 m²s⁻¹ in rural areas and 1.0 m²s⁻¹ in urban areas) showed marginally worse performance when the high cutoff values in urban areas were removed (Pleim and Gilliam, 2012). ” This is in contrast to the simulations using the same strategy with AURAMS, which showed improved performance for PM_{2.5} mean statistics, O₃ nighttime statistics, and reduced performance for PM_{2.5} correlation coefficient and slope.

P5608 In9: Change “...that in the use of a lower-limit cutoff for the model diffusion coefficients ...” to “...that the use of a greater lower-limit cutoff for the model diffusion coefficients in the air quality model than are used in the meteorology model...” Note that all models have some lower limit to eddy diffusivity to avoid numerical errors.

We have changed the line to read “that the use of a lower-limit cutoff that may be higher than may be appropriate to describe subgrid-scale turbulent mixing”. Different meteorological models may employ different lower limits – ultimately the issue is to ensure that the meteorological model simulations what is happening in the ambient atmosphere.

P5621 In11: terminology here and earlier is confusing. Using “the lower cutoff” implies a lower value but that is opposite of what is meant. Clarify by changing to something like: “a greater value of the minimum eddy diffusivity”

We’ve reworded as suggested.

P5622 In1-3: I disagree with this statement. It should be the goal of all air quality models to use the same PBL model and the same eddy diffusivities as the meteorology model to which it is linked. There is no physical reason to model the turbulent mixing of chemical scalars differently from meteorological scalars. These greater values of the minimum eddy diffusivity are simply artifacts which compensate for other errors in the CTM systems. That is not to say that PBL models don’t need improvement; they clearly do. But the improvements should be made identically to both the meteorological and CTM systems.

Fair enough. Generalizing a bit further, we have changed that first sentence to read “These results should not be taken to imply that improvements to the model representations of turbulent mixing and/or other factors should be ruled out as a line of investigation for achieving improved model performance.”

Technical Comments:

P5602 In6: Change "...in relative..." to "...relative..."

Done.

P5607 In12-25: This paragraph seems to be a repetition of an earlier paragraph. It should be eliminated or reduced.

The repeated information has been removed from the second paragraph.

P5610 In7: Please clarify "second to 4th largest"

Depending on the hour of the night – this has been added to the text.

P5612 In 9: remove "altitude"

Done.

P5615 In7: Change "important" to "importance"

We searched for "important" – page 5613 is probably the location noted (the word did not appear in the original document on the page number noted by the reviewer); we have fixed this.