Review of
Draft Report:
INVESTIGATION OF VOC REACTIVITY EFFECTS USING EXISTING REGIONAL AIR QUALITY MODELS
By William P. L. Carter, Gail Tonnesen and Greg Yarwood
Draft Dated 9/30/02
Review submitted 2/7/03
Response Completed 4/17/93

Abstract:
p. ii, first paragraph: "A total of 42 different ozone reactivity scales ..." And "... and on using six different methods to derive reactivity scales ..." Claiming 42 different scales is confusing. Do you meant 7 different methods (preferably, metrics – see general comment on Introduction), 3 different grid sizes, and 2 different averaging periods. (7 x 3 x 2) Is this right? If so, could you make this clearer here?

Changed to "a number of different reactivity scales"
Changed "methods" to "methods or metrics"

Abstract:
p. ii, second paragraph: "… regardless of which region or quantification method are employed." Seems to imply sub-regional work was performed already. Does region really refer to grid size? Also, see general comment on Introduction.

"region" changed to "region of the domain"

Introduction:
General comment: It would be helpful if you could add a paragraph defining metric and scale and quantification, also methods (as in the phrases, "methods to quantify" and "methods to derive"). Adding a few sentences here would help the readers.

A footnote was added to the use of the word "metric" in the "objectives" section in an attempt to clarify this. However, the reader should be able to understand these issues after looking at the text in the "methods" sections that derive the methods and metrics in detail.

p. 1, second paragraph: change "… the EPA has not recommended use of the MIR or any other reactivity scale in regulatory applications, and believes …" to "… the EPA believes ..."

deleted "has not recommended use of the MIR or any other reactivity scale in regulatory applications, and"

Objectives:
p. 1, last sentence: Perhaps add something saying that the ACC contracted this project with the oversight and approval of the RRWG (just so nobody questions whether it is biased towards industry because it was contracted by the ACC).

Added sentence **"**Although funded by the ACC, his work was carried out with the oversight and approval of the RRWG as a whole."

p. 3, Figure 1 and other similar figures: Is it possible to add major urban areas such as New York and Chicago on the map? Specifically, I'd like to know if the areas where MIR-3D and MOIR-3D occur are urban or downwind areas.

Geographical identification information is not readily available with the present data sets. Additional work would be required to generate this information. The reader should be able to infer the locations of the major urban areas as well as the authors.

Airshed Model:
p. 4, A brief model performance summary from previous work done by ENVIRON should be presented here. For instance, since the emissions are perturbed to obtain DDM sensitivities, it would be useful to show, at a minimum, pie charts of the emissions, such as Figure ES-2 from the ENVIRON reference. Similarly, Figures ES-7 and ES-8, comparing model results to observations on the 12 K grid, would be useful to include, perhaps in an appendix referenced in this section. Likewise for Figures ES-4 and ES-5, which show the temporal pattern of ozone. Also, a number of sensitivities runs (non-DDM) were performed by ENVIRON to evaluate model performance. Of particular note is the sensitivity to a doubling of mobile source (MOBILE5) emissions, which for the most part improves model performance. This has relevance to several subsequent comments. Also, a number of sensitivities for meteorology, using different meteorological models, were unable to resolve all the meteorological concerns. And this has relevance to model performance at the 4 K grid resolution.

The ENVIRON report does not provide any information indicating this episode was evaluated for its performance on an 8-hr basis, yet RRWG is using it for sensitivities on an 8-hr basis. I have seen cases where model performance degrades going from 1-hr to 8-hr, and cases where it only changes modestly (some improvement and some degradation). It is not obvious what the 8-hr performance will be like for this case. The observational data reside in the NARSTO archives; 8-hr performance should be evaluated as an additional task.

Text provided by Environ was included to address this. A brief discussion of emissions, modeled maximum O_3 concentrations, and model performance is included.

Chemical Mechanism:
Table 1, p.6: ALD2 represents more than just acetaldehyde
Table 1, p. 6, last sentence: change "string radical inhibiting" to "*strong* radical inhibiting"

Table 1 modified to indicate more clearly the types of compounds represented. Typo corrected.

Base ROG mixture:
p. 7, first paragraph: change "However, since the purpose of this project is assess the effects of policies that encourage VOC substitutions ..." to "However, since the purpose of this project is to

examine the effect of VOC substitutions ..." My point is that you really aren't assessing real policies, while you are looking at the AVOC substitution, it isn't even close to being realistic for a real policy.

Sentence changed as requested

p. 7, 2nd paragraph, 3rd sentence: change "… in terms of base ROG mixture is useful …" to "… in terms of base ROG mixture is *a* useful …"
p. 7, 3rd paragraph, 2nd sentence: "… composition should be generally be very similar …" Eliminate one "be".

Changed

p. 7, 3rd paragraph, last sentence: "For model intercomparison purposes …" This report does not explicitly contain the grid-based model intercomparison. It does contain intercomparison of grid-based model results with EKMA results (where the EKMA results are now relative to the chosen base ROG). Clarify that this facilitates intercomparison with other studies, like Russell et al. and any other studies that may follow. And that only the EKMA intercomparison is contained in this report.

"For model intercomparison purposes" changed to "To facilitate comparisons with other studies"

p. 8, Figure 2: Formaldehyde is expressed in chemical symbol form and all other species are expressed in CBIV abbreviation. Consistency with Table 1 is preferred.

I personally prefer using symbols preferred by chemists rather than abbreviations used by models (which sometimes are arbitrarily shortened because of software limitations) when referring to an actual chemical or a species representing an actual chemical. Of course, the model species name should be used if the model species is distinctly different from any actual chemical (such as "TOL" or "ARO2").

DDM Calculations:
p. 8, last paragraph, 2nd sentence: change "derivates" to "derivatives"

Changed

Reactivity Metrics Used:
General: To avoid confusion with so many metrics, include a brief summary table here listing all 7 different metrics, named the same way they are referred to later on in the results section.

Summary table (table 2) added.

General: Have we explored the sensitivity fields enough to jump right into metrics? For example, can you include some time series plots to demonstrate the temporal patterns of sensitivities (analogous to Russell et al.'s Figures 5 and 6)? These would be most helpful if they were

displayed side-by-side with ozone time series plots.

This could be done as part of a follow-on study if desired

General: It may be instructive to examine some fixed grid cells for the temporal evolution of their sensitivities – an urban region, a downwind urban-impacted region, a rural area. The observational analysis report from Task group 2 has revealed that even urban areas are not always VOC-limited. Some reader might question whether the favorable comparison of grid-derived metrics with EKMA reactivities is somewhat attributable to the EKMA-like analysis of the grid model (i.e., one sensitivity per cell chosen at the time of the daily maximum ozone).

This could be examined as part of a follow-on study if desired

Quantification of Ozone Impact in a Cell:
p. 10, first paragraph: Certainly the daily maximum ozone for 1-hr and 8-hr are important, and they provide for the best comparisons with EKMA, but it seems like an artificial constraint to always limit each cell to a single sensitivity per day. This excludes obtaining potentially useful temporal information, such as persistence of high ozone. In this regard, this "quantification method" may be subject to questioning. Can this quantification method be justified? Consistency with EKMA is one potential justification.

For most cells, a 24-hour day ending at midnight is a logical unit upon which to base a metric because the ozone generally declines at night, so the times of the maxima are during the day. However, alternative methods to examine persistence metrics could be examined as part of a follow-on study if desired

p. 11, 1st paragraph: "The 1-hour ozone impact metrics were derived ..." This is confusing because it is not really a separate "metric" it is an impact that is used in calculating a metric. When it says metric it makes it look like one of the 7 metrics but it isn't. I suggest changing to "The 1-hour ozone impacts were derived ..."

"The 1-hour ozone impact metrics were derived" changed to "The metrics based on 1-hour ozone impacts". Similar change made when referring to 8-hour based metrics.

Computation of Global Reactivity Metrics:
P. 11, first sentence: change "… give thousands or relative …" to "… give thousands *of* relative …"

Changed

Regional Average Ozone Metric:
p. 12, second paragraph: "One problem … low O3 cells equally with the cells where O3 approaches or exceeds …" In addition to using a cutoff to address this problem, one could allow (in a separate quantification method) more than one sensitivity value per cell provided the O3 was above the cutoff for that hour. This suggests a model-based, ozone-driven approach to spatial weighting of cells (i.e., if high ozone persists in a cell, that cell's sensitivity for each hour

of high ozone is included in the metric). This would produce a regional ozone metric weighted by persistence.

> Alternative metrics based on persistence could be examined in a follow-on project if desired.

Regional Average Ozone Over the Standard Metric:
p. 12, fifth paragraph: This is where the persistence weighting of cells (as discussed above) could be applied, taking into account temporal persistence of high ozone in a cell.

> Alternative metrics based on persistence could be examined in a follow-on project if desired.

Minimum Substitution Error Metrics:
p. 13, first sentence: change "One of the applications of reactivity-based regulatory polices would be to use ..." to "One potential application of reactivity-based policies *might* be to use ..."

> Changed

Minimum Substitution Error Method 1 (MSE 1): Base ROG for VOC
p. 13, first sentence for this section: Could this be clarified somehow? Would a mathematical expression help? How about using "VOC species" throughout and not switching to compound?
p. 13, first sentence after equation IV: change "... reactivity of species the global ..." to "... reactivity of species *in this* global ..."

> Changed the text to consistently use the term "Test VOC" to refer to the VOC whose relative reactivity is being assessed. Similar changes made to the text discussing MSE 2.

Minimum Substitution Error Method 2 (MSE 2): VOC for Base ROG:
p. 14, first full paragraph after equation VII, second sentence: change "… instability occurs the species …" to "… instability occurs *when* the species …"

> Changed

EKMA Reactivity Scales:
General: Not every reader was able to appreciate that you apparently redid some EKMA work (obtaining slightly different numerical values), and then the readers had to follow and distinguish the terms EKMA reactivities, EKMA MIR and EKMA MOIR from your previous work. Which terms capture the new numerical values with the intercomparison base ROG, and which terms refer to your prior work?

> The following paragraph was added to the end of this section: "It is important not to confuse the EKMA MIR, MOIR and EKMA base case reactivity scales discussed in this work with the MIR and other scales developed previously by Carter (1994a, 2000a). The EKMA scales in this work were calculated using the CB4 mechanism as implemented in the CAMx model, while the previous Carter (1994a, 2000a) EKMA scales were derived

using versions of the SAPRC mechanism. The CB4-based scales are only considered in this work because the objective is to examine effects of model formulation and not effects of using a different chemical mechanism."

p. 16, last paragraph: The first sentence says that the same ROG is used in the EKMA reactivity scales as in the regional scales, but in the second sentence you say that the mixture used in calculating the EKMA MIR reactivity scales is different. Please clarify, and perhaps be more explicit that you reran (or recalculate) new EKMA derived values for relative reactivity with the chosen base ROG.

The term "Base ROG" used in the EKMA calculation was changed to "mixture used to represent the anthropogenic VOCs in the model" in this discussion in an attempt to clarify this.

Categorization of Reactivity Characteristics Throughout the Modeling Domain:
p. 17, Table 2: A specific way to designate the daily maximum ozone, such as $O_3|_{DM}$ would be less confusing than redefining O_3 for the purpose of this table.

Changed references to O_3 in the body of the table to "daily maximum O_3". Kept the portion of note [a] about excluded cells.

p. 17, note [a] to Table 2: Keep the footnote, but change "Location" to "Locations"

Changed

Results and Discussion:
p. 20, second paragraph: Change July 11 to July 12
p. 20, last paragraph: change "… tend to greater than …" to "… tend to *be* greater than …"

Changed

p. 21 to 28, Figures 3-7 and Tables 4-6: The ozone peak increases going from 36 K to 12 K grid size (as expected). However, the ozone peak decreases going from 12 K to 4 K (unexpected). This is very likely an artifact of the meteorological issues mentioned previously (by reference to the ENVIRON report) for this episode. At issue is the fine scale meteorology, which was not reproduced by the meteorological model. This could impact conclusions drawn from metrics, or comparisons to metrics, at the 4 K grid resolution. Some discussion is warranted.

A discussion of fine scale meteorology is beyond the scope of this report.

p. 24, Figure 6: The 12 K peak ozone for this day is given as 0.000, which cannot be correct.

Corrected. (I assume you mean figure 9)

Ozone Sensitivities to VOC Emission Categories:
p. 30 to 31, and Figures 10 to 12:

General: At various places on these pages the terms total anthropogenic VOC, anthropogenic VOC, as well as AVOC, and all anthropogenic VOC sources are used. The term total VOC is also used. This can be confusing. The term "total" typically distinguishes biogenic + anthropogenic from anthropogenic only emissions. Coupling "total" with anthropogenic VOC seems unnecessary. In any event, better clarification would help.

The "total" was removed from anthropogenic where appropriate.

Base ROG vs. Total Anthropogenic VOC Sensitivities:
p. 30 to 31, both paragraphs: The vintage and development of the emission databases used to develop the base ROG and that used to develop the model input for the episode are very similar, so the result is not surprising. What would you anticipate if you had a future case with future case emissions?

A discussion of this is beyond the scope of this report.

Anthropogenic vs. Total VOC Sentisitivities:
p. 31: "Figure 10 shows the ratios of anthropogenic to total VOC sensitivities …" include percentage "… shows the ratios as a percentage of …"

"Ratios" changed to "percentages".

Relative Contribution to Anthropogenic Source Types:
p. 31, General: The model performance was improved for the most part by doubling the mobile emissions (MOBILE5) as noted in the ENVIRON report. If this is taken as an alternative basecase, one could argue that the approximate contributions could just as likely be area (50/130) = 39%, mobile (60/130) = 48%, and point (20/130) = 15% [simplistically doubling the mobile source sensitivity and renormalized]. This reinforces the report's recommendation for updating emissions (i.e., MOBILE6 in this case). Could this be used to help quantify part of the uncertainty in the analysis?

Mobile source emissions have a strong diurnal pattern, potentially (it seems) making it difficult to capture their contribution to ozone formation using sensitivities at the daily maximum ozone only. Could an integrated sensitivity over some time interval better approximate the source contribution to ozone formation, particularly when a source has a strong diurnal pattern in emission, than the sensitivity at the daily maximum ozone (per your current quantification method)? In that sense, the 8-hr analysis may be more interesting, and while the relative contributions for 8-hr results appear consistent, at least at 36 K, all the anthropogenic source sensitivities decrease.

A discussion of this is beyond the scope of this report.

p. 33, Figure 11: Legend shows results as "AVOC > 10%" Percent of what? Please explain. AVOC/total VOC > 10%?

Figure caption explains that it is percentage of AVOC to total VOC sensitivity

p. 40, 3rd paragraph : The high degree of scatter for the toluene species is not seen in GIT report. Is this solely due to the CB4 mechanism used? Can it be concluded that CB4 is not appropriate for a number of VOCs that have strong NOx sinks?

Text is added to indicate that other mechanisms do not indicate such a large scatter for the TOL species. Reference was already made to previous discussion of problems with the CB4 TOL species in the section where the mechanism is discussed.

p. 40, line 5 of 3rd paragraph: Change "… this is may not be …" to "… this may not be …"
p. 40, line 9 of 4th paragraph: Change "… if it such species …" to "… if such species …"

Changed

Regional Relative Reactivity Scales:
p. 41-55, General: Temporal information is excluded by the quantification method as noted previously. Temporal information may be useful. The reader could benefit from having a feel both for ozone temporal patterns as previously mentioned (by including ES-4 and ES-5 from the ENVIRON report) and for sensitivity temporal patterns (requested).

This could be done as part of a follow-on study if desired

Comparison of Regional Metrics:
p. 41, first paragraph: change "... the distributions of sensitivities a given episode..." to "...the distributions of sensitivities *during* a given episode..."

Changed

p. 42: Table 7: Use names of the metrics that are consistent to those on pages 11-15. Also, maybe put them in the order in which they are introduced on those pages. For example, it takes some paging back and forth to determine the relationship between "Regional MIR to MOIR" (on p. 15) for classifying cells and "Avg MIR-MOIR" (=1/2 [MIR + MOIR])? on this table. This should be made more clear.

It is necessary to use abbreviated names to be able to fit them on the table – I think it is useful to have as many scales as possible across the page. The scales are ordered in a manner to represent a progression from low to high effective ranges, as discussed later.

p. 42, Table 7: Add a footnote for EKMA MIR since this is slightly different from previous MIR values.

It was already stated (see above) that the EKMA scales in this work are different than those used in previous studies. Also, the table caption says these are for the CB4 mechanism, which is not the case for the previous EKMA scales.

p. 42, line 4 of 1st paragraph: Change "… some differences I the various metrics …" to "… some

differences between the various metrics …"

I could not find this text

p. 42, second paragraph, first sentence: "... than metrics based on impacts in a single selected cell." to "... than metrics based on impacts in a single selected cell, *such as the regional maximum ozone and the regional MIR metrics*"
p. 45, first paragraph: change: "In some cases the highest O3 cell found to be …" to "In some cases the highest O3 cell *was* found to be …"

Changed

p. 46, 2nd sentence of 1st paragraph (Figure 23): How does the report conclude that the variability is considerably less from Figure 23 than Figure 22? It is not clear.

The variability in Figure 23 (now Figure 25) is less than in Figure 22 (now 24) with the possible exception of formaldehyde. The phrase "possible exception of formaldehyde" was added.

p. 48, first paragraph: What are "… the three metrics that are considered preferable based on the discussion above"? The previous discussion in this section only talks about the avg O3 over the standard and the avg. MIR-MOIR. I see by looking at the plots that the third is the regional average ozone, but you have not discussed why this is one of the preferable ones.

The word "preferable" was replaced by "selected," with a sentence added stating that the regional maximum O_3 and regional MIR are not included because as discussed above other metrics are preferable to them.

Effect of O_3 Cutoff Level:
p. 51, 2nd to last paragraph, first sentence: change "... ozone metric is that the many very low …" to "… ozone metric is that *it weights* the many very low …". Also, change "... from a policy perspective as those cells …" to "... from a policy perspective *the same* as those cells …". In the same paragraph, change "… in ozone increases in those regions result in O3 levels ..." to "… in ozone increases in those regions, *resulting* in O3 levels ..."

Changed to "One critique of the average ozone metric is that includes the many very low O_3 cells where ozone levels are not of concern from a policy perspective in addition to those cells that are of greater concern." Note that the low O_3 cells are not necessarily weighted equally because the weighting is actually by IR, which is usually higher in higher O_3 cells.

p. 51, line 4 of 3rd paragraph: Change "… the36K domain …" to "… the 36K domain …".

Changed

p. 51, last line: change "… regional average ozone metrics." to "… regional average ozone

metrics with the cutoff set above 0.12 ppm."

Correction: "…the values on the far right are the regional ~~average~~ maximum ozone metrics"

p. 54, Figure 33: Use a legend on the plot or a note explaining the 4 lines are 4 episode days.
p. 55: Comparison of Regional a EKMA Relative Reactivities: change to "Regional *and* EKMA"

Changed

Comparison of Regional and EKMA Relative Reactivities:
General Question: How are the units for EKMA reactivities (mass change in ozone per mass change in VOC) and for grid modeling reactivities (concentration of ozone per fractional emission change) compared on a consistent basis? Are there any assumptions needed on EKMA box volume? Would an example calculation in an appendix help?

Following sentence added to the end of the first paragraph in this section: "Note that the EKMA as well as the regional relative reactivities are all given on a carbon basis, so the units are consistent."

p. 55, first paragraph: Clearly, the quantification method adopted here for reactivity analysis using a grid model allows easy comparison to Carter's reactivity scales derived from 1-day EKMA scenarios. This comparison justifies its selection as one quantification method, but how robust is this quantification method?

I do not understand this question.

p. 56, 1st full paragraph: change "… separate sells …" to "… separate *cells* …"

Changed

Results of Large Scale Substitution Calculations:
p. 59, first paragraph: change "... given on Table 9 through Table 11." to "… given on Table *8* through Table 11."
p. 59, second paragraph: change "... completely inert calculations." to "… completely inert *emissions*."
p. 59, second to last paragraph: change "... can be compared distributions …" to "… can be compared *with* distributions ..."

Changed

p. 60 to 63, General: This is a place where it would be convenient to better harmonize with MCNC's method of reporting on their substitutions regarding the use of PAVE plots and tables and metrics. If these reports support policy, it would be convenient to have consensus on the preferred method of reporting for substitution scenario runs.

This could be done as part of a follow-on study if desired

p. 63, Table 11: change heading from percentabes to percentages
p. 69, 2nd full paragraph: change "Figure 42 …" to "Figure 43 ..."
p. 76, paragraph beginning with "Figure 46 shows ..." Change "... the regional average ozone concentration ..." to "... the regional average *(daily maximum)* ozone concentration …". Also in the same paragraph change "… appears to be strongly biases ..." to "... appears to be strongly *biased* ..."

Changed

Classification of Reactivity Scales by Effective Range:
p. 76 to 81, general: The extensive discussion of the "effective range" factor may create a misleading impression about the significance of this factor in selecting an appropriate reactivity-metric (i.e., that a metric should be selected based on effective range considerations). The most pertinent and important science-related factor to consider in metric selection should be the "ambient conditions"-related, reactivity-variability factor. A metric associated with minimum reactivity variability can also be associated with a moderate effective range, but if not, we cannot justify changing the metric for that reason. The other contractors need to comment on and discuss reactivity metrics, offering their views on the relevance or significance of the "effective range" factor in selecting a reactivity metric.

The purpose of this discussion is to point out that this is one of the ways to classify reactivity scales that may have some potential utility to policymakers. A discussion of its relative importance compared to other policy considerations is beyond the scope of this report.

p. 77, middle of last paragraph: change "This also apparent …" to "This is also apparent …".

Changed

Conclusions
p. 82, end of first paragraph: I think the concern extends to the more dynamic and more realistic representation of physicochemical processes in grid modeling versus EKMA, in addition to incorporating meteorology.

The sentence "There is also the concern that the EKMA models cannot represent meteorology, transport and other dynamic processes as realistically as does 3-D regional models." Was added to the end of the first paragraph.

p. 84, last paragraph: The question lingers that the quantification method for analyzing the grid model output closely resembles the EKMA quantification method, but it does not represent the only quantification method. Could this be explored further before suggesting a return to EKMA scenarios, and updating and expanding them?

I am not sure I understand this point. In any case, this could be explored in a follow-on

study.

p. 85, last paragraph: Clarify that DDM sensitivities predict reasonably well for large AVOC changes here because AVOC is << than total VOC.

The sentence "(The reasonably good performance of the linear approximation may be due in part to the fact that the contributions of AVOCs are relatively small compared to biogenic VOCs in most of the domain.)" was added in the middle of the paragraph after the sentence saying the linear approximation performed well.

p. 86, carryover paragraph: change "… if very large substations are employed …" to "… if very large substitutions are employed …"

Changed

Recommendations:
p. 86: 2nd paragraph: Is any update in the text warranted based on Hakami's latest calculations?

That paragraph was modified and shortened so it no longer refers to "unexpected results". Essentially it says, "However, the analysis approach they employed does not permit comparison with all the metrics examined in this work, and it may be useful to independently evaluate these results with a different model", and goes on to say that process analysis may be useful "if unexpected results are obtained".

p. 86, line 1 of paragraph: Change "Russell et and co-workers" to "Russell and co-workers." And (e.g., Hamaki, 2002) to (e.g., Hamaki et al., 2002).
p. 87, last paragraph, change "Therefore, is should be possible ..." to "Therefore, *it* should be possible ..."
p. 88, 2nd to last sentence: change "... the need to reduce these uncertainties become more evident …" to "… the need to reduce these uncertainties *becomes* more evident …."

Changed

References
p. 90, references of Carter et al. (1997a) and (1997b) are repeated from p. 90.

Changed

Please include the "Responses to 13 Key Questions" as an appendix.

I'm not sure it is appropriate to include opinions on policy matters in a technical report such as this, and it does not readily relate to any of the discussion. I would prefer to keep this in a separate document. If what you already have is not satisfactory for this purpose, please let me know, and indicate the format that you would like me touse.

General Comment: There seems to be agreement that use of **relative** reactivity data reduces

substantially all problems arising from the variability of **absolute** reactivity. However, aren't there reactivity applications requiring use of absolute reactivity data? It would be useful if this question would be examined and answered.

> There may indeed be regulatory applications where absolute reactivity may be useful. An example is assessing the relative benefits of the O_3 – removal catalysts compared to VOC controls. However, a discussion of this is beyond the scope of this report.

General Comment: It would be useful – and seems feasible – to derive and offer estimates, even rough ones, of the **total** uncertainty associated with the reactivity data the contractor derived/used, and a discussion of important policy implications, namely, identify alternative ways and the scientifically best-justified way a reactivity policy could/should treat the uncertainty factor.

> Treatment of uncertainty (as opposed to variability – which is discussed) is a major research area that is well beyond the scope of the present project. Analyses of various sources of uncertainties, their estimated magnitudes, and policy implications may be an appropriate subject of follow-on work.