

Replies to Reviewer 1

Reviewer's comments are in standard font.

Responses and changes to the manuscript are in italics.

General Comments

Source attribution and source sensitivity techniques are valuable tools for air quality planners to understand air quality model results and to design effective emissions control strategies. The author presents an innovative approach that adds to the extensive existing literature on these methods. However, as described below, it is not clear in the manuscript how this method differs from existing sensitivity methods and source attribution approaches. It is also unclear how this method would be used in an air quality planning context, so I recommend revisions to the manuscript to more clearly explain this path-integral method (PIM) and to illustrate its use in the air quality planning context.

As described in the manuscript, a variety of approaches have been used to identify emissions source categories that are important contributors to ozone and other secondary pollutants. These methods can be broadly grouped as either model sensitivity methods or source attribution approaches. Sensitivity methods include forward sensitivities ("brute force" sensitivities and the Decoupled Direct Method (DDM)) and backward sensitivities (adjoint methods). Sensitivity approaches evaluate effects of changes in emissions on ozone or other pollutants relative to a base case model simulation. Source attribution methods rely on tracer species and/or evaluation of mass budgets and are used to evaluate sources that contribute to ozone or other pollutants in a particular model simulation.

Both sensitivity methods and source attribution methods have limitations, and neither fully addresses the needs of air quality planners who are tasked with identifying the most effective combination of emissions controls that demonstrate progress in reducing air pollutants (while avoiding possible dis-benefits of NO_x control) and that ultimately attain national ambient air quality standards. The key limitation of sensitivity methods is that ozone can have either positive or negative sensitivity to changes in emissions, and the magnitude and sign of the sensitivity depends on both the size of the emissions reduction and the sequential order in which different sources are controlled. Source attribution studies address this limitation of sensitivity methods by evaluating the contribution of each emissions source in a particular scenario, typically one that represents current conditions or an historical pollution episode. Thus, source apportionment methods can identify the largest contributors to ozone under current conditions, and this is useful for identifying and prioritizing sources to control. However, source apportionment methods do not predict the sensitivity of ozone to emissions controls in a future scenario because source attribution in the base case does not account for non-linear chemistry effects of emissions changes in the future case. Typically, air quality planners use source apportionment methods to identify potential emissions sources for control and then perform additional model sensitivity simulations to evaluate specific control scenarios

This manuscript describes a novel approach for using sensitivity simulations to assess how ozone and other pollutants respond to changes in emissions. The PIM method is designed to use a finite set of model sensitivity simulations to systematically represent the range of model response to control of each emissions category ranging from 0 to 100% control, while also representing the variation in model response depending on the order in which source categories are controlled. The paper illustrates the PIM method using a highly simplified box model scenario with two layers. The author notes that a limitation of the method is that a large number of model sensitivity simulations are required and that this would result in large computational cost. Another limitation not identified in

the paper is that presentation of the PIM results could also be challenging for a realistic model scenario. It would be more useful to illustrate the application of this method using a realistic model simulation. My initial reaction is that this method would be challenging to use in an air quality planning context, and that it does not provide insights that are not already available from conventional sensitivity and source apportionment methods. Therefore, I recommend that the manuscript be revised to illustrate the PIM method using a more realistic model scenario using a 3-dimensional photochemical model for an historical ozone episode used in an air quality planning context.

Response: The PIM is a new method that introduces new concepts and is consequently more difficult to explain and understand than a modification of an existing method. This paper is intended to explain the method, indicate good options for integrating the sensitivities numerically, and illustrate the method for a chemically realistic model. The model used here contains the full CB6 chemical mechanism used in the 3-D model and therefore all the nonlinearities in the 3-D model. Thus, the work provides a valid test of whether the PIM can accurately account for the effects of the nonlinear chemistry. The PIM has also been used with the full, 3-D version of CAMx in an application to the eastern U.S. This application is reported in another paper that has been submitted to a journal. Combining the two papers would result in a long paper that would likely be less, not more easily understood.

Manuscript change: A reference to the paper describing use of the PIM in a 3-D model has been added at the end of the Introduction and at the end of the Conclusions.

Response: In the 3-D application, the PIM required between 2.5 and 3.0 times the effort of the brute-force (zero-out) method. This is larger than the effort for some other methods, but not so large as to be unmanageable.

Response: If only one emission-control scenario is studied, presenting the PIM results is no more challenging than presenting the results from other source apportionment methods, namely providing the source contributions for specific receptor locations or geographic areas. If one chooses to evaluate multiple emission control strategies, then there will be a different set of source contributions for each control strategy of interest. This adds complexity but also provides new, useful information on how the importance of sources varies depending on the control strategy.

Response: The PIM does provide insights not available from other methods. It provides contributions for anthropogenic sources that sum up to the difference (Δc_i) between a simulation with and a simulation without the anthropogenic sources. This is not a requirement for other methods and is generally not true for results from other methods. If the sum of the anthropogenic contributions is greater than (less than) Δc_i , the results indicate that reducing anthropogenic emissions will produce greater (lesser) pollutant reductions than will actually occur (based on the model predictions). This can obviously be misleading for air quality planning purposes. The PIM also provides source contributions that effectively are averaged over a range of emissions and chemical regimes. Other source apportionment methods and sensitivity methods only use the emissions and chemistry in the base case, without any emission reductions. Lastly, the PIM provides source contributions for all species in the chemical mechanism simultaneously, e.g., O_3 , NO_2 , and air toxics, without additional assumptions exterior to the model, e.g., when the chemistry is VOC-limited vs. NO_x -limited. This is not true for source apportionment methods based on tracers.

Manuscript change: The following paragraph has been added to Section 2.1:

“Because the sensitivities are integrated over the path P in Eq. (2), the PIM considers a range of chemical conditions in calculating the source contributions, from zero to the full anthropogenic emissions in the base case. Methods based on tracers or a Taylor series expansion (e.g., with first- and second-order sensitivities) use only the emissions and the chemical conditions of the base case. Thus, the PIM provides source contributions that are averaged over the emission-control scenario, not specific to the base case.”

It would also be helpful to compare PIM with the high-order DDM method which accounts for some nonlinearity in the photochemical reactions. Can the high-order DDM provide the same information that PIM calculates using multiple sensitivity simulations?

Response: The high-order DDM (HDDM) does not provide the same information as the PIM. The HDDM involves constructing a Taylor series expansion through 2nd order for the emissions (and the chemical regime) in the base case, and the expansion is then used to extrapolate to lower emissions. However, the extrapolation is generally accurate for no more than a 50% reduction in anthropogenic emissions. At larger reductions in anthropogenic NO_x emissions, the chemical regime changes to a high VOC/NO_x ratio due to the large biogenic VOC emissions, the chemistry becomes progressively more sensitive to NO_x emissions, and the sensitivities calculated by HDDM for the base-case emissions are no longer accurate. Therefore, the sum of the source apportionments from the HDDM does not agree well with the anthropogenic increment of the O₃ concentration. The PIM avoids this problem by using sensitivities over the full range of anthropogenic emissions (0% to 100%). Furthermore, the HDDM approach calculates 2nd order cross sensitivities between different sources, but, as discussed in Section 2.2, the cross terms are not assigned to individual sources in the source apportionment. The PIM effectively assigns these cross terms to sources based on the emission-control strategy. Section 2.2 and the Supplementary Information show how this could be done if a Taylor series expansion of 1st order sensitivities is accurate (though, as indicated above, such an expansion actually has a limited range of applicability).

As a general comment, the description of the method seems to be overly abstract and it would be helpful to explain in simple language the physical significance of terms such as the path variable and hypercube. Also see comments on page 5 below.

Manuscript changes:

The sentences in Section 2.1 where u is introduced are revised to: “The path P can be described via a path variable u that describes position along the path. Each λ_m is a function of u , such that as u varies from 0 to 1, each $\lambda_m(u)$ also varies from 0 to 1, and the path P defining the changes in anthropogenic emissions is traced from the background case to the base case in the M -dimensional space of the scaling parameters λ_m . “

The second sentence in Section 2.2 is revised and a new sentence added: “This is a path along the edges of a hypercube in Λ -space. (The hypercube defines all possible emission-control strategies, contains M axes, one axis for each λ_m , and includes all values of λ_m from 0 to 1.)”

Specific Comments

Page 3, lines 10-17: The text in this paragraph is difficult to follow, see comments below: “If the anthropogenic increment is allocated to sources, the PIM requires that the base-case concentration minus the sum of the anthropogenic source contributions equals the background concentration. Other methods do not have this requirement, and thus may over- or under-allocate the anthropogenic increment to the anthropogenic sources and also allocate a concentration to the

background sources that is not the concentration from a simulation with only background sources included.”

Are there cases in PIM in which the anthropogenic increment is not allocated to sources? How does the PIM treat apportionment if the anthropogenic increment is negative? Does this method assume that increments are always positive? Also, it is not necessarily correct to state that other approaches “over- or under-allocate”. Given the constraint of accounting for negative sensitivity of O₃ to precursors and accounting for the contribution of both VOC and NO_x to O₃, each method adopts a unique strategy for mass attribution. Allocation can be internally consistent with the adopted strategy and therefore technically accurate, and yet provide estimates that differ from other apportionment approaches. Instead, the author might argue that the strategies adopted in other apportionment approaches are poorly understood and result in incorrect interpretation of the results, or are incorrectly implemented and therefore produce inaccurate results. However, more description and analysis of results from other apportionment methods is needed to support such a conclusion.

Response: As discussed in the 3rd paragraph of Section 2.1, the PIM can also allocate the total concentration, not just the anthropogenic increment. However, to allocate the total concentration, the initial concentrations, boundary concentrations and emissions from all sources (biogenic and anthropogenic) must be included in the analysis.

Response: If the anthropogenic increment is negative, then some anthropogenic source contributions may be positive, but the sum of all the anthropogenic contributions will be negative (Eq. (1)).

Response: The method does not assume that anthropogenic increments are always positive. The O₃ increment is negative from hours 0 to 12 in the simulation (Figure 2) and the NO_x source contributions are negative during these hours (Figure 4).

Response: To my knowledge, none of the other methods requires that the base-case concentration minus the sum of the anthropogenic source contributions (call this difference δ) equals the background concentration. If δ is less than the background concentration, then the method assigns too much importance to the anthropogenic sources and will give the impression that reducing anthropogenic emissions will lower the pollutant concentration more than will actually occur (over-allocation of the anthropogenic increment to the anthropogenic sources). Similarly, if δ is greater than the background concentration, the method is assigning too little importance to the anthropogenic sources (under-allocation of the anthropogenic increment). I think that the text is a fair statement of this important difference between the PIM and other methods, given the definition of over- and under-allocation used here. However, to explain the issue more clearly, changes have been made to the Introduction and Conclusions.

Manuscript changes:

Section 1. Introduction, p. 3, lines 10-17 are revised to: “An important advantage of the PIM is its ability to allocate to sources a concentration increment, i.e., the difference between two simulations (base and background cases). If the anthropogenic increment is allocated to sources, the PIM requires that the base-case concentration minus the sum of the anthropogenic source contributions equals the background concentration. Other methods do not have this requirement, and thus may ascribe too much or too little importance to the anthropogenic sources.”

Section 5. Conclusions, first paragraph is revised to: “As shown in Sect. 4, the PIM can allocate the difference in concentration between two simulations to emission sources. Consequently, the PIM requires that the base-case concentration minus the sum of the anthropogenic source contributions

(difference δ) equals the background concentration (within the accuracy of the numerical integration). Other methods do not have this constraint. If δ is less than the background concentration, then the method assigns too much importance to the anthropogenic sources and will give the impression that reducing anthropogenic emissions will reduce the pollutant concentration more than will actually occur (over-allocation of the anthropogenic increment to the anthropogenic sources). Similarly, if δ is greater than the background concentration, the method assigns too little importance to the anthropogenic sources (under-allocation of the anthropogenic increment). The PIM ensures that the anthropogenic increments to O_3 and the other species are neither over- nor under-allocated to the anthropogenic sources."

Page 3: Equation 1 includes only first-order sensitivities of c_i with respect to the scaling parameters. A term is also needed to represent higher order sensitivities.

Response: Eq. (1) is exact as written. No higher-order sensitivities are involved because the integrals are over the variables (λ_m) with respect to which the derivatives are taken. Eq. (1) is a generalization to multiple dimensions of the familiar relationship: $f(b) - f(a) = \int_a^b \frac{df}{dx} dx$. In one dimension, there is only one path from a to b . In multiple dimensions, there is an infinite number of possible paths from the starting to ending point of the integration

Manuscript change: The first paragraph of Section 2.1 is revised to: "The PIM is based on an exact mathematical equation that is in itself not new. In particular, the equation is routinely used in thermodynamics (Sect. 2.3). However, the application of the equation to atmospheric modeling is new. The equation is the generalization to multiple variables of a familiar relationship for a single variable, namely that the integral of the derivative of a function ($\int_a^b \frac{df}{dx} dx$) is equal to the difference in the value of the function at the ends of the integration interval ($f(b) - f(a)$)."'

How does this result differ from higher order DDM?

Response: See the comments above about HDDM.

Page 4, lines 20-22: "However, if all the source contributions and Δc_i are calculated, then Eq. (1) can be used to check the accuracy of the integration procedure. The integration procedure can be modified then, if necessary, so that the sum of the source contributions equals Δc_i within the desired error tolerance."

Because models are not strictly mass conservative and are subject to numerical error, an approach is needed to avoid accumulation of error. Thus, a method is needed to prevent accumulation of error in the case where all source contributions are not calculated.

Response: The only difference between calculating all source contributions S_{im} and calculating a subset of them is simply that, for the omitted S_{im} , the sensitivities aren't calculated and the integration in Eq. (2) is not done. There are no changes to the sensitivities that are calculated, and there are no additional errors in the numerical integration procedure for the S_{im} that are calculated compared to the case where all S_{im} are calculated. (The DDM gives the same value for a particular sensitivity coefficient regardless of whether only that sensitivity or multiple sensitivities are calculated.) Therefore, there are no new errors and no accumulation of error in the case where only a subset of the source contributions are calculated.

Page 5, lines 3-7: It seems very problematic that the source apportionment result depends on the order in which emissions sources sensitivities are calculated. Also, the definitions of path (P), path

variable (u) and normalized difference (s) are not clear. What is the physical significance of normalized distance or absolute distance along P ?

Response: The source apportionments depend on the path P but not on the order in which emissions sensitivities are calculated. The apportionments depend on the path because the chemical environment changes depending on how emissions are controlled. The simplest example is the special case of successive zero-out of sources, Section 2.2. The apportionments depend on which source is removed first, which second, etc.

Response: As indicated above, the sentences in Section 2.1 where u is introduced are revised to clarify the definition of u and P .

Response: There is no special physical significance to the normalized distance s . Use of s rather than u is just a change of variable that can make it easier to understand where the Gauss-Legendre integration points will be, ensure a more uniform distribution of the points, and thus may make the integration more accurate.

Manuscript change: The sentence prior to Eq. (4) is revised to: "Changing the integration variable from u to s , the source contribution becomes"

One of the key limitations of source sensitivity methods is that sensitivities are not additive, and that that O₃ can have negative sensitivity to precursors in some cases. A key motivation for source apportionment methods is to estimate the actual mass contribution of a source to O₃ rather than the sensitivity of O₃ to that source. It would seem that the PIM methods suffers from the limitation of sensitivity approaches and does not provide a mass attribution estimate that is unaffected by nonlinear sensitivities.

Response: The PIM provides an allocation of pollutant concentrations to emissions by integrating first-order sensitivities over emission levels. Because it integrates over the emission levels, it includes all nonlinear effects, and the integral of each sensitivity is a mass amount, not a sensitivity. The source contributions are additive, as shown by Eq. (1).

Page 5, lines 23-24: "The simplest and shortest integration path, termed the diagonal path, is defined by $\lambda_m = u$, all m . This is a straight line from $\Lambda = 0$ to $\Lambda = 1$ along which the emissions from all sources are reduced or grown by the common factor u ."

Previously " u " was defined to be the path variable, which was unclear. If u is a factor by which emissions are adjusted, this could be explained more clearly. It is confusing that an emissions change factor is described as a distance.

Response: As indicated above, the sentences in Section 2.1 where u is introduced are revised to clarify the definition of u . The λ_m scale the emissions, u describes position along the emission-control path, λ_m is a function of u , and for the diagonal path $\lambda_m(u)$ equals u . For other paths, $\lambda_m(u)$ does not equal u , e.g., Eqs. (6, 7).

Manuscript change: The two sentences are revised to: "The simplest and shortest integration path, termed the diagonal path, is defined by $\lambda_m(u) = u$, all m . This is a straight line from $\Lambda = 0$ to $\Lambda = 1$ along which the emissions from all sources are reduced or grown by the common factor $\lambda_m(u) = u$."

Page 8. "Analogy in thermodynamics". Suggest deleting this section as it does not seem relevant and is not helpful for illustrating the PIM method.

Response: Reviewer 2 asked for examples of other problem domains where Eq. (1) has been used, so it seems important to keep this example.

Page 14, lines 7-10: “The PIM allows source contributions to be either positive or negative. If the secondary pollutant formation is inhibited by emissions of some species, source, or geographic area, the sensitivity to these emissions will be negative for at least some values of the scaling parameter m , and the integral in Eq. (2) may be negative.”

The above statement highlights the difference between sensitivity and source apportionment methods. Negative “contribution” indicates that this is a sensitivity method, not a mass attribution method. NO_x emissions can contribute to ozone production even when ozone has a negative sensitivity to changes in NO_x emissions, and source attribution methods such as OSAT in CAMx are designed to quantify the mass contribution of NO_x to O₃. Thus, the PIM method is not quantifying the mass contribution to ozone production (in the sense that it evaluated in a source apportionment approach), rather, it is characterizing the negative sensitivity of ozone to NO_x.

Response: The anthropogenic increment can be negative. (See O₃ in Figure 2, hours 0 to 12). This means that removing the anthropogenic emissions actually increases the O₃ concentration, which is due to the titration of O₃ by NO emissions. If the anthropogenic source contributions are all positive, then the implication is that eliminating the anthropogenic sources will reduce O₃, which is incorrect. The PIM integrates the sensitivities over the anthropogenic emissions and accumulates the effects of both positive and negative sensitivities to the emissions. The integral of the sensitivity is no longer a sensitivity, but a source contribution. If the anthropogenic increment is negative, then the sum of the anthropogenic source contributions will be negative, which is a consistent and correct result.

Page 14, lines 23-24: “The concentrations in the background simulation can be determined by an actual simulation or by subtracting the source contributions from the base-case concentrations.”

The solution for source contributions is non-unique, i.e., the solution depends on assumptions made in the order in which sources are evaluate. Therefore, subtracting source contributions from the base case does not provide a unique estimate of background concentrations. The only reliable modeling approach to estimate background concentrations is to perform a model simulation that does not include anthropogenic emissions

Response: The source contributions are not unique, but the sum of the source contributions is always the difference Δc_i in Eq. (1).

Manuscript change: The sentence is revised to: “The concentration in the background simulation can be determined by an actual simulation or by subtracting the sum of all the source contributions from the base-case concentration.”

Reviewer 2

Reviewer's comments are in standard font.

Responses and changes to the manuscript are in italics.

The manuscript by A. M. Dunker describes a method for the source apportionment of photochemical pollutants using a novel method termed PIM (Path Integral Method). An interesting feature of this method is that it allows not only the apportionment of the total concentration of any given pollutant in a simulation to particular sources, but also the differences in concentration between two simulations with nonzero sources. The manuscript is organised logically and written clearly. The method is described well, and applied to a very simple case study involving a two-box photochemical simulation. Without being shown to work in a three dimensional air quality model simulation, the method has not yet been fully proven, but this initial proof of concept study is still clearly within the scope of GMD. This method represents a potentially interesting addition to the toolbox of photochemical modellers interested in source apportionment. I recommend publication in GMD subject to minor revisions.

The author could be slightly more balanced in their summary of the previous literature. Each method of source apportionment has its own particular strengths and weaknesses, including the PIM. While the strengths of the PIM are clearly described, it would help if some of the drawbacks of the PIM were also already mentioned in the introduction section. The PIM requires that the model being used be modified to calculate first-order sensitivities (for example using the Decoupled Direct Method), requires extra computation time to do multiple simulations along the emission control path, and potentially provides an infinite number of possible source apportionments.

Response: A sentence has been added to the Introduction stating that the PIM requires more computational effort than some other source apportionment methods. Calculation of first-order sensitivities by the decoupled direct method or the adjoint method has already been implemented in many models, so I don't think that the requirement to calculate these sensitivities is a significant limitation of the PIM. Also, while there is an infinite number of source apportionments in principle, the connection of the source apportionments to emission-control strategies limits the number of apportionments of interest in practice. Emission-control strategies often focus on similar reductions from all sources, which could be represented well by the diagonal path described in the manuscript.

Manuscript change: Addition of the following sentence to the 4th paragraph of the Introduction: "The PIM does require more computational effort than some other source apportionment methods because first-order sensitivities must be calculated at several levels of anthropogenic emissions."

Furthermore, several recent source apportionment schemes have been missed in the introduction to this manuscript. For example, Emmons et al. recently published a method for attributing ozone production to NO_x emissions (doi:10.5194/gmd-5-1531-2012), Butler et al. published a method for attributing ozone and VOC degradation products to emitted CO and VOC (doi:10.1016/j.atmosenv.2011.03.040), and Grewe et al. published a general approach for tagging both NO_x and VOC (doi:10.5194/gmd-3-487-2010).

Response and manuscript changes: The papers by Emmons et al., Butler et al., and Grewe et al. have been added to the reference list and cited in the Introduction as tracer methods.

On lines 16-18 of page 9093, it is mistakenly claimed that other source apportionment methods assign ozone produced from CO emissions to VOC, and ozone produced from HONO emissions to

NO_x. At least in the case of Butler et al. (2011), it would be possible to tag CO emissions separately from VOC, and in the case of Emmons et al. (2012), it would be possible to tag HONO emissions separately from NO_x, avoiding this problem.

Response and manuscript change: The sentence has been deleted.

It would also be useful to know more about the background of the PIM. On lines 3-4 of page 9083, it is mentioned that the mathematical equation behind the PIM is not itself new, but that the application to source apportionment is new. Here it would be interesting to know the other problem domains to which the method has been applied.

Response and manuscript change: The beginning of Section 2.1 has been revised to:

“The PIM is based on an exact mathematical equation that is in itself not new. In particular, the equation is routinely used in thermodynamics (Sect. 2.3). However, the application of the equation to atmospheric modeling is new. The equation is the generalization to multiple variables of a familiar relationship for a single variable, namely that the integral of the derivative of a function ($\int_a^b \frac{df}{dx} dx$) is equal to the difference in the value of the function at the ends of the integration interval ($f(b) - f(a)$).”

There may be additional applications of the equation beyond thermodynamics, but I am not aware of them.

1 Path-integral method for the source apportionment of 2 photochemical pollutants

3

4 **A. M. Dunker¹**

5 [1]{A. M. Dunker, LLC, Bloomfield Hills, Michigan, USA}

6 Correspondence to: A. M. Dunker (amdunker@gmail.com)

7

8 **Abstract**

9 A new, path-integral method is presented for apportioning the concentrations of pollutants
10 predicted by a photochemical model to emissions from different sources. A novel feature of
11 the method is that it can apportion the difference in a species concentration between two
12 simulations. For example, the anthropogenic ozone increment, which is the difference between
13 a simulation with all emissions present and another simulation with only the background (e.g.,
14 biogenic) emissions included, can be allocated to the anthropogenic emission sources. The
15 method is based on an existing, exact mathematical equation. This equation is applied to relate
16 the concentration difference between simulations to line or path integrals of first-order
17 sensitivity coefficients. The sensitivities describe the effects of changing the emissions and are
18 accurately calculated by the decoupled direct method. The path represents a continuous
19 variation of emissions between the two simulations, and each path can be viewed as a separate
20 emission-control strategy. The method does not require auxiliary assumptions, e.g., whether
21 ozone formation is limited by the availability of volatile organic compounds (VOC's) or
22 nitrogen oxides (NO_x), and can be used for all the species predicted by the model. A simplified
23 configuration of the Comprehensive Air Quality Model with Extensions is used to evaluate the
24 accuracy of different numerical integration procedures and the dependence of the source
25 contributions on the path. A Gauss-Legendre formula using 3 or 4 points along the path gives
26 good accuracy for apportioning the anthropogenic increments of ozone, nitrogen dioxide,
27 formaldehyde, and nitric acid. Source contributions to these increments were obtained for paths
28 representing proportional control of all anthropogenic emissions together, control of NO_x
29 emissions before VOC emissions, and control of VOC emissions before NO_x emissions. There

1 are similarities in the source contributions from the three paths but also differences due to the
2 different chemical regimes resulting from the emission-control strategies.

3

4 **1 Introduction**

5 The goal of source apportionment is to determine, quantitatively, how much different emission
6 sources contribute to a given pollutant concentration. Source apportionment is thus a useful
7 tool in developing efficient strategies to meet air quality standards by identifying the most
8 important sources. If emissions are involved in only linear processes between where they are
9 emitted and where they impact a receptor location, the concentration of the pollutant at the
10 receptor is the sum of independent contributions from the individual emission sources. For
11 example, one can define a tracer for each source of primary, unreactive particulate matter (PM)
12 in an air quality model such that the sum of the tracer concentrations is the total primary PM
13 concentration and the tracer concentrations form the source apportionment. However, if a
14 secondary pollutant is formed by nonlinear chemical reactions, source apportionment is more
15 complicated and, indeed, there is no unique apportionment.

16 Reflecting this non-uniqueness, a number of approaches have been developed for source
17 apportionment of secondary pollutants. The simplest approach is source removal or the brute
18 force method. Simulations with and without a particular source are compared, and the changes
19 in predicted concentrations are assigned to emissions from that source (Marmur et al., 2006;
20 Tong and Mauzerall, 2008; Wang et al., 2009; Zhang et al., 2014). A related approach is the
21 factor-separation method, which for M sources involves analysis of a set of 2^M simulations
22 (Stein and Alpert, 1993; Tao et al., 2005). Each simulation includes emissions from a different
23 source or a different combination of sources. Pollutant concentrations are assigned not just to
24 sources but to interactions among sources.

25 Another approach involves the use of reactive tracers for individual chemical species, sources,
26 and/or geographic regions (Yarwood et al., 1996; Dunker et al., 2002b; Mysliwiec and
27 Kleeman, 2002; Wagstrom et al., 2008; Wang et al., 2009; [Grewe et al., 2010](#); [Butler et al.,](#)
28 [2011](#); [Emmons et al., 2012](#); [-Kwok et al., 2013](#)). However, various chemical assumptions
29 (beyond those in the chemical mechanism) are needed to track production of the secondary
30 pollutant in nonlinear reactions. In addition, the source contributions are often restricted to be
31 positive even if some primary pollutants can inhibit formation of the secondary pollutant.
32 Assignment methods trace through all the reaction pathways from products back to parent

1 reactants (Bowman and Seinfeld, 1994; Bowman, 2005). These methods also require extra
2 chemical assumptions for reactions in which a product results from multiple reactants. Lastly,
3 local sensitivity coefficients have been used to apportion ozone (O₃) and PM (Dunker et al.
4 2002b; Cohan et al., 2005; Koo et al., 2009). This approach involves constructing a Taylor
5 series representation of the concentration as a function of source emissions and extrapolating
6 the representation to zero emissions.

7 This work presents a new approach for source apportionment called the Path-Integral Method
8 (PIM). The PIM provides a new, direct mathematical connection between sensitivity analysis
9 and source apportionment and a connection between source apportionment and emission-
10 control strategies. In contrast to reactive-tracer and assignment methods, the PIM does not
11 require additional chemical assumptions. An important advantage of the PIM is its ability to
12 allocate to sources a concentration increment, i.e., the difference between two simulations (base
13 and background cases). If the anthropogenic increment is allocated to sources, the PIM requires
14 that the base-case concentration minus the sum of the anthropogenic source contributions
15 equals the background concentration. Other methods do not have this requirement, and thus
16 may ascribe too much or too little importance to the anthropogenic sources. over- or under-
17 allocate the anthropogenic increment to the anthropogenic sources and also allocate a
18 concentration to the background sources that is not the concentration from a simulation with
19 only background sources included. The PIM does require more computational effort than some
20 other source apportionment methods because first-order sensitivities must be calculated at
21 several levels of anthropogenic emissions.

22 The PIM is applied here to allocate the anthropogenic increments of O₃ and other species using
23 a 2-cell configuration of the Comprehensive Air Quality Model with Extensions (CAMx)
24 (ENVIRON, 2013). Another application of the PIM using a detailed, 3-D CAMx configuration
25 for the eastern U.S. will be reported elsewhere ([Dunker et al., 2015](#)).

26

27 **2 Description of the PIM**

28 **2.1 Equations**

29 The PIM is based on an exact mathematical equation that is in itself not new. In particular, the
30 equation is routinely used in thermodynamics (Sect. 2.3). However, the application of the
31 equation to source apportionment atmospheric modeling is new. The equation is the

1 generalization to multiple variables of a familiar relationship for a single variable, namely that
 2 the integral of the derivative of a function ($\int_a^b (df/dx) dx$) is equal to the difference in the
 3 value of the function at the ends of the integration interval ($f(b) - f(a)$).

4 For this work, the equation (Kaplan, 1959) takes the form

$$6 \Delta c_i(\mathbf{x}, t) = c_i^1(\mathbf{x}, t; \mathbf{\Lambda} = 1) - c_i^0(\mathbf{x}, t; \mathbf{\Lambda} = 0) = \sum_{m=1}^M \int_P \frac{\partial c_i(\mathbf{x}, t; \mathbf{\Lambda})}{\partial \lambda_m} d\lambda_m$$

5 (1)

7 The c_i^1 is the concentration of species i in the base case, with all emissions present, and c_i^0 is
 8 the concentration in the background case, with M emission sources removed. $\mathbf{\Lambda}$ is the array of
 9 the parameters λ_m that scale the emissions of the M sources. If all $\lambda_m = 0$ ($\mathbf{\Lambda} = 0$), the emissions
 10 are those of the background case, and if all $\lambda_m = 1$ ($\mathbf{\Lambda} = 1$), the emissions are those of the base
 11 case. The $\partial c_i / \partial \lambda_m$ are the first-order sensitivities of c_i with respect to the scaling parameters.
 12 The integrals on the right side of Eq. (1) are taken over a curve or path P in M -dimensional
 13 space leading from the emissions in the background case to those in the base case. The Δc_i is
 14 the difference between the concentrations in the base and background cases at the same spatial
 15 location \mathbf{x} and time t .

16 Although the focus here is on emissions, Eq. (1) can also include parameters that scale the initial
 17 and boundary concentrations. Furthermore, if the background case has all emissions and initial
 18 and boundary concentrations set to zero, then $c_i^0 = 0$ and Δc_i is the total concentration. Thus,
 19 the PIM can allocate the total concentration in a simulation as well as concentration differences
 20 between simulations.

21 The contribution of source m to Δc_i , S_{im} , is defined to be

$$23 S_{im}(\mathbf{x}, t; P) = \int_P \frac{\partial c_i(\mathbf{x}, t; \mathbf{\Lambda})}{\partial \lambda_m} d\lambda_m$$

22 (2)

24 The PIM does not strictly require that the source contributions be calculated for all M sources
 25 or that Δc_i be calculated. The sensitivities can be determined for a subset of the sources and
 26 integrated to obtain the S_{im} only for the sources of interest. However, if all the source
 27 contributions and Δc_i are calculated, then Eq. (1) can be used to check the accuracy of the

1 integration procedure. The integration procedure can be modified then, if necessary, so that the
2 sum of the source contributions equals Δc_i within the desired error tolerance.

3 The source contributions depend on the path P from the point $\Lambda = 0$ to the point $\Lambda = 1$. Because
4 there are an infinite number of paths between these two points, there are an infinite number of
5 sets of source contributions, one set corresponding to each path. Viewed in the direction of
6 integration, from $\Lambda = 0$ to $\Lambda = 1$, emissions are added into the background case until the base
7 case is reached. Viewed in the opposite direction, emissions are controlled from the base case
8 until the background case is reached. Thus, each path P represents a possible emission-control
9 scenario, and the contribution of a given source to the change in concentration Δc_i depends on
10 the control scenario.

11 Because the sensitivities are integrated over the path P in Eq. (2), the PIM considers a range of
12 chemical conditions in calculating the source contributions, from zero to the full anthropogenic
13 emissions in the base case. Methods based on tracers or a Taylor series expansion (e.g., with
14 first- and second-order sensitivities) use only the emissions and the chemical conditions of the
15 base case. Thus, the PIM provides source contributions that are averaged over the emission-
16 control scenario, not specific to the base case.

17 The path P can be described via a path variable u that describes position along the path. Each
18 λ_m is a function of u . Then, $\Lambda = \Lambda(u)$, such that as u varies from 0 to 1, each $\lambda_m(u)$ also varies
19 along P from 0 to 1, and the path P defining the changes in anthropogenic emissions is traced
20 from the background case to the base case in the M -dimensional space of the scaling parameters
21 λ_m . However, u may not equal the normalized distance along P , denoted by s , and s can be
22 useful in designing the numerical integration procedure because it is easier to understand the
23 distribution of the integration points using s . The absolute distance D is related to u by

$$25 \quad D(u) = \int_0^u \left[\sum_{m=1}^M \left(\frac{d\lambda_m}{du} \right)^2 \right]^{1/2} du$$

24

(3)

26 Then, $s(u) = D(u)/D(1)$. Using ~~Changing the integration variable from u to s~~ , the source
27 contribution becomes

$$28 \quad S_{im}(\mathbf{x}, t; P) = \int_0^1 \frac{\partial c_i(\mathbf{x}, t; \Lambda)}{\partial \lambda_m} \Big|_{\Lambda=\Lambda(s)} \frac{d\lambda_m}{ds} ds$$

1

2 with

$$4 \quad \frac{d\lambda_m}{ds} = \frac{d\lambda_m}{du} \frac{D(1)}{\left[\sum_{m=1}^M \left(\frac{d\lambda_m}{du} \right)^2 \right]^{1/2}}$$

3

(5)

5 The sensitivity in Eq. (4) is evaluated along the specific path defined by $\Lambda(s)$. Also, though the
6 emissions are reduced along the path and the concentrations are determined in a simulation with
7 the reduced emissions, the sensitivity of c_i is to λ_m , which scales the full emissions in the base
8 case, not the reduced emissions. The decoupled direct method (DDM) provides an accurate,
9 efficient means for calculating the sensitivities (Dunker, 1981, 1984; Yang et al., 1997). The
10 DDM has been implemented in current 3-D models for the formation of O_3 and particulate
11 matter (Dunker et al., 2002a; Cohan et al., 2005; Napelenok et al., 2006; Koo et al., 2007).

12 The simplest and shortest integration path, termed the diagonal path, is defined by $\lambda_m(u) = u$,
13 all m . This is a straight line from $\Lambda = 0$ to $\Lambda = 1$ along which the emissions from all sources
14 are reduced or grown by the common factor $\lambda_m(u) \equiv u$. If there are two sources, Fig. 1 displays
15 the diagonal path, Path 1, and two other possible paths. Path 2 is defined by the equations:

$$16 \quad \lambda_1(u) = u^3 \tag{6}$$

$$17 \quad \lambda_2(u) = \sin\left(\pi \frac{u}{2}\right) \tag{7}$$

18 Beginning at the base case, point B, emissions from source 1 are reduced much more rapidly
19 than those from source 2 along Path 2. As the first 80% of the emissions from source 1 are
20 reduced, only 20% of the emissions from source 2 are reduced. Then the remaining 80% of the
21 emissions from source 2 are reduced as the remaining 20% of the emissions from source 1 are
22 reduced, down to the background case, point b. Path 3 is the opposite of Path 2, obtained by
23 interchanging the definitions of λ_1 and λ_2 in Eqs. (6, 7). For the diagonal path, the normalized
24 distance and path variable are identical, $s(u) = u$, and $d\lambda_m/ds$ in Eq. (4) is identically 1. For
25 Paths 2 and 3, $s(u) \neq u$, and $d\lambda_m/ds$ must be determined from Eq. (5).

26 The Gaussian numerical integration formulas have maximum precision (Isaacson and Keller,
27 1966). This means that for a given number of points at which the integrand is evaluated, n , the
28 formulas give an exact integration of all polynomials of degree 0 up to $2n-1$, the maximum

1 degree possible using n points. Thus, the Gaussian formulas should minimize the number of
 2 points at which the integrand in Eq. (4) must be evaluated to achieve a given accuracy. This is
 3 useful because the major computational effort in the PIM is determining the sensitivities at
 4 multiple points along the path P . The Gauss-Legendre formula is one version of Gaussian
 5 integration suited to integration of a function $f(z)$ over a finite interval:

$$7 \int_a^b f(z) dz \cong \frac{b-a}{2} \sum_{k=1}^n w(\xi_k) f\left(\frac{b-a}{2} \xi_k + \frac{b+a}{2}\right)$$

6 (8)

$$9 z = \frac{b-a}{2} \xi + \frac{b+a}{2}$$

8 (9)

10 The ξ_k are the zeroes of the Legendre polynomials, and the $w(\xi_k)$ are weights determined to give
 11 the formula the maximum precision. The ξ_k and $w(\xi_k)$ are readily available (efunda, 2014).

12 2.2 Special cases

13 One special case is successive zero-out (SZO) of the sources. In SZO, the emissions from one
 14 source are reduced to zero while leaving all other emissions unchanged, then the emissions
 15 from a second source are reduced to zero, etc. until the background case is reached. This is a
 16 path along the edges of ~~the a~~ hypercube in Λ -space. (The hypercube defines all possible
 17 emission-control strategies, contains M axes, one axis for each λ_m , and includes all values of λ_m
 18 from 0 to 1.) In Fig. 1, one SZO path would be B - b_2 - b and the other, B - b_1 - b . Along the
 19 segment B to b_2 of the former path, the sensitivities are nonzero, but $d\lambda_2 = 0$. Therefore, the
 20 only contribution to Δc_i in Eq. (1) is that for source 1, and this contribution equals $c_i^B - c_i^{b_2}$.
 21 Similarly, along the segment from b_2 to b , $d\lambda_1 = 0$, the only contribution to Δc_i is that for source
 22 2, and the contribution equals $c_i^{b_2} - c_i^b$. Thus, a SZO path is a special case of PIM in which no
 23 calculation or integration of sensitivities is required, only a series of simulations to obtain the
 24 concentrations at the corners of the hypercube. Calculation and integration of the sensitivities
 25 is necessary if two or more sources are controlled simultaneously, and the path is then interior
 26 to the hypercube.

1 Another special case involves expanding the sensitivities in a Taylor series in the λ_m at $\Lambda = 1$
 2 (base case). If there are two sources and the Taylor series through first order in λ_m is integrated
 3 along the diagonal path, then (see Supplementary Information (SI))

$$5 \quad S_{i1}(diag) = \left. \frac{\partial c_i}{\partial \lambda_1} \right|_{\Lambda=1} - \frac{1}{2} \left. \frac{\partial^2 c_i}{\partial \lambda_1^2} \right|_{\Lambda=1} - \frac{1}{2} \left. \frac{\partial^2 c_i}{\partial \lambda_1 \partial \lambda_2} \right|_{\Lambda=1}$$

4

(10)

$$7 \quad S_{i2}(diag) = \left. \frac{\partial c_i}{\partial \lambda_2} \right|_{\Lambda=1} - \frac{1}{2} \left. \frac{\partial^2 c_i}{\partial \lambda_2^2} \right|_{\Lambda=1} - \frac{1}{2} \left. \frac{\partial^2 c_i}{\partial \lambda_1 \partial \lambda_2} \right|_{\Lambda=1}$$

6

(11)

8 The cross term ($-\partial^2 c_i / \partial \lambda_1 \partial \lambda_2$) is split evenly between S_{i1} and S_{i2} . If the integration is done
 9 instead on the path $B-b1-b$ in Fig. 1, the full cross term is assigned to S_{i1} and is absent entirely
 10 from S_{i2} . Similarly, if the integration is along the path $B-b2-b$, the full cross term is assigned to
 11 S_{i2} and is absent from S_{i1} . Thus, the source contributions are the same for these 3 paths except
 12 for the location of the cross term. Cohan et al. (2005) expanded c_i in a second-order Taylor
 13 series about $\Lambda = 1$ and used it to develop source apportionments that are the same as Eqs. (10,
 14 11) except that they did not assign the cross term to the individual sources. The PIM shows
 15 that the cross term can be assigned to sources based on the emission-control path.

16 **2.3 Analogy in thermodynamics**

17 The dependence of the source contributions on path has an analogy in thermodynamics. For
 18 example, in the case of a single-component gas, the energy E is a function of the state variables:
 19 temperature T , and volume V . The change in E between two states of the system, ΔE , depends
 20 only on the initial and final values of T and V . However, when ΔE is split into contributions
 21 from the heat exchange with the surroundings ($\int_p dq$) and the pressure (p)-related work
 22 ($\int_p p dV$) in the equation, $\Delta E = \int_p dq - \int_p p dV$, the heat exchange and work depend on
 23 the path P from the initial to final states of the system. Thus, the concentrations c_i from an air
 24 quality simulation may be regarded as analogous to E and the emissions, initial and boundary
 25 concentrations, meteorology and chemical mechanism as analogous to T and V . The Δc_i
 26 between two simulations differing only in emissions can be allocated to sources, but this
 27 allocation is analogous to heat exchange and work and depends on the path along which the
 28 emissions are changed.

1

2 **3 Model and inputs**

3 Time-dependent inputs were developed for CAMx, v6.00, configured with 2 cells in a vertical
4 column. The lower cell varied diurnally in height from 100 →300 →100 m and the upper cell
5 varied in height such that the top of the column was 1500 m. Diurnally varying emissions were
6 introduced at the bottom boundary. The simulations were run for 3 days, June 20-22, beginning
7 with clean initial concentrations in both cells. There was no transport into the cells via the
8 lateral or top boundaries. The latitude was that of Los Angeles and Atlanta. The Carbon Bond
9 6 (CB6) chemical mechanism represented the gas-phase chemistry (Yarwood et al., 2012). The
10 effect of the inputs is that cleaner air from the upper cell is entrained into the lower cell during
11 the morning as the lower cell grows in height. Then, in the evening, the lower cell shrinks in
12 height and leaves pollutants aloft in the upper cell. Consequently, there is carry-over of
13 pollutants from day to day affecting the chemistry in the lower cell. Additional details of the
14 simulations are in Table S1 (SI).

15 The emissions were developed from the national totals in the 2008 U.S. National Emission
16 Inventory, version 3 (U.S. EPA, 2013b) with several adjustments. Emissions from wildfires
17 and prescribed fires were excluded because these vary greatly from year to year and were
18 unusually high in 2008. Also, to represent summer conditions, emissions from residential wood
19 combustion were excluded. Further, emissions of NO from lightning were added (Koo et al.,
20 2010). The emissions were segregated into biogenic (plus lightning) emissions and 5 major
21 source categories of anthropogenic emissions: fuel combustion, industrial sources, on-road
22 vehicles, non-road vehicles, and other emissions. Vegetation and soil emissions and their
23 speciation are from BEIS3.14 (Pierce et al., 1998). Anthropogenic emissions of volatile organic
24 compounds (VOC's) from a major source category were allocated to CB6 species using
25 speciation profiles from SPECIATE 4.3 for 1 or 2 sub-categories of sources comprising a
26 significant fraction of the VOC emissions (Simon et al., 2010; U.S. EPA, 2013a). The annual
27 emissions of VOC species, NO_x (=NO + NO₂), CO, and HONO for each source category were
28 allocated to hours of a Wednesday in June using temporal profiles (U.S. EPA, 2013c). On a
29 national scale, the biogenic VOC emissions are large compared to the anthropogenic VOC
30 emissions, but this is not the case in urban areas. To represent better an urban area the
31 anthropogenic emissions were weighted by a factor of 5 and the biogenic emissions by a factor

1 of 1. A summary of the resulting daily emission rates for all source categories is given in Table
2 1, and the complete set of emission rates is in Table S2.

3 The model and inputs are not intended to be a detailed representation of a specific urban area
4 but rather to provide a useful platform for testing the PIM, specifically different integration
5 formulas and the dependence of the source contributions on paths.

6

7 **4 Results**

8 The concentrations of O₃ and formaldehyde (FORM) in the background simulation (biogenic
9 emissions only), the base simulation (both the biogenic and anthropogenic emissions) and the
10 difference between the simulations (anthropogenic increment) are shown in Fig. 2. Similar
11 plots for NO₂ and HNO₃ are in Fig. S1. The peak O₃ concentration remains relatively constant
12 over the 3 days in the background simulation (47-52 ppb) but increases steadily in the base
13 simulation (from 75 ppb on day 1 to 151 ppb on day 3) due to the additional anthropogenic
14 emissions on days 2 and 3 and the carryover of pollutants in the upper cell. Both O₃ and FORM
15 have sizeable concentrations in the background case whereas NO₂ and HNO₃ have very low
16 concentrations due to the low biogenic NO_x emissions. The O₃ increment is negative at the
17 beginning of day 1 due to the titration of O₃ by the anthropogenic NO emissions. The VOC/NO_x
18 ratio in the base case increases from 5-7 on day 1 to 9-20 ppbC/ppb⁻¹ on day 3. Overall, the
19 simulations provide a wide range of conditions for testing the PIM.

20 **4.1 Accuracy of the numerical integration**

21 The O₃, FORM, NO₂, and HNO₃ increments were allocated to the 5 anthropogenic source
22 categories and to the 4 species or groups of species emitted by each source category: VOC, CO,
23 NO_x, and HONO. Thus, a total of $M = 20$ sensitivities were calculated and integrated in the
24 PIM. Source apportionments were determined for 3 emission-control paths: diagonal (Diag);
25 VOC first (VOCF); NO_x first (NOxF). Along the Diag path, the scaling parameters $\lambda_m^{VOC} =$
26 $\lambda_m^{CO} = \lambda_m^{NOx} = \lambda_m^{HONO} = u$, for each source category $m = 1, \dots, 5$. Thus, the sources and
27 emission species are treated equivalently. The VOCF path emphasizes initial control of VOC
28 and CO emissions followed by later control of NO_x and HONO emissions, as defined by λ_m^{VOC}
29 $= \lambda_m^{CO} = u^3$ and $\lambda_m^{NOx} = \lambda_m^{HONO} = \sin(\pi u/2)$, $m = 1, \dots, 5$. The NOxF path has the reverse
30 assignments of u^3 and $\sin(\pi u/2)$. Viewing λ_m^{VOC} , λ_m^{CO} as analogous to λ_l in Fig. 1 and λ_m^{NOx} ,

1 λ_m^{HONO} as analogous to λ_2 , , then the VOcF path in 20-dimensional space is analogous to Path
 2 2 in Fig. 1 and the NOx F path is analogous to Path 3.

3 The Gauss-Legendre formula was tested for accuracy using different numbers of integration
 4 points and different integration variables. One set of tests, labeled GLns, used the distance s as
 5 the integration variable and n integration points. Another set of tests, labeled GLnr, used a
 6 transformation of the variable s to $r = s^{1/2}$. Equation (4) then becomes

$$8 \quad S_{im}(\mathbf{x}, t; P) = 2 \int_0^1 \frac{\partial c_i(\mathbf{x}, t; \Lambda)}{\partial \lambda_m} \Big|_{\Lambda=\Lambda(s[r])} \frac{d\lambda_m}{ds} \Big|_{s(r)} r dr$$

7 (12)

9 Because the background case contains no anthropogenic emissions, O₃ formation is strongly
 10 limited by the availability of NO_x. As a consequence, the sensitivity of O₃ with respect to any
 11 λ_m that scales NO_x emissions is very large near $\Lambda = 0$, but the sensitivity decreases very rapidly
 12 as NO_x emissions are added. The transformation to r has two potentially beneficial effects for
 13 the source apportionment of O₃. First, the points for the numerical integration are chosen for
 14 the variable r . When transformed back to the variable s , the points for s are closer to $\Lambda = 0$
 15 than if s were the integration variable, giving more resolution where the sensitivity is changing
 16 most rapidly. Second, the factor r in Eq. (12) reduces the magnitude of the integrand near $r =$
 17 $s = \lambda_m = 0$, and makes the integrand identically 0 at $r = 0$. This can yield an integrand that is
 18 easier to integrate. Finally, as a simple alternative, the source contributions were calculated by
 19 the trapezoidal rule using the 2 points at $\Lambda = 0$ and 1 (labeled TR2).

20 The sum of the source contributions on the 3 paths was compared to the anthropogenic
 21 concentration increment (right- vs. left-hand sides of Eq. (1)) to determine the accuracy of the
 22 formulas. Table 2 gives the mean absolute error and mean bias of the formulas for O₃ and
 23 FORM, and Table S3 gives the error and bias for NO₂ and HNO₃. For comparison, the mean
 24 absolute values of the increments ΔO_3 , $\Delta FORM$, ΔNO_2 , and ΔHNO_3 are 34.9 ppb, 1.52 ppb,
 25 7.67 ppb, and 16.0 ppb, respectively. Though they use the same number of points, there is a
 26 large reduction in error and bias from TR2 to GL2s or GL2r, indicating the significant
 27 advantage of the GL formulas. As the number of points included in the GLns or GLnr formulas
 28 increases, the error decreases for O₃, FORM, and NO₂ and generally the bias as well. There are
 29 some exceptions to this trend for HNO₃, but these occur for cases where the error and bias are
 30 already quite low (average error < 4% of the average increment). For O₃ and the Diag path, the

1 GLnr formula gives more accurate results than the GLns formula for 2 or 3 points and
2 essentially the same accuracy for 4 points. For FORM, the GLnr formula is always more
3 accurate than the GLns formula. The GLnr formula is usually less accurate than the GLns
4 formula for NO₂ and HNO₃ and for O₃ with the NO_xF and VO_{CF} paths.

5 Table 2 also shows that the accuracy of a formula is lower for the VO_{CF} path than the other
6 paths when using the same number of points. This difference can be understood by examining
7 the integrand in Eq. (4). Figure 3 displays the integrands for allocating ΔO₃ to sources at the
8 time of peak O₃ on day 3, when it is most difficult to obtain good agreement between the sum
9 of the source contributions and ΔO₃. Along the Diag and NO_xF paths, the integrands have a
10 constant curvature, either positive (Diag) or negative (NO_xF), and the integrands are mainly
11 positive, with only small negative values near $s = 1$. However, along the VO_{CF} path, 4 of the
12 integrands have positive curvature from $s = 0$ to $s = \sim 0.5$ and then negative curvature for the
13 remainder of the path. Also, the integrands vary over a wider range along the VO_{CF} path than
14 the other paths. Further, the integrands for on-road vehicles and fuel combustion are both
15 positive and negative, resulting in the cancellation of contributions to the integrals from
16 different sections of the path. The change in curvature, wider range of variation and especially
17 the cancellation of contributions require more points on the VO_{CF} path to obtain an accurate
18 integration.

19 Overall, the GL3r formula for the Diag path and the GL4s formula for the other paths give quite
20 accurate results and were used to calculate the source apportionments in Sect. 4.2. Figure S2
21 gives a comparison of the sum of the source contributions vs. ΔO₃, ΔFORM, ΔNO₂, and ΔHNO₃
22 at each hour of the simulation. The plots show again that the largest errors occur for the VO_{CF}
23 path.

24 **4.2 Source apportionments**

25 Figure 4 presents the apportionment of ΔO₃ to the 5 source categories and 4 emission species
26 using the Diag path. The VOC contributions are always positive, and the largest contributions
27 are from industrial sources and on-road and non-road vehicles. The NO_x contributions are small
28 and primarily negative on day 1, when the atmospheric VOC/NO_x < 7.5 ppbC/ppb⁻¹ in the base
29 case. Under these conditions, NO_x emissions tend to inhibit O₃ formation, and hence the
30 contributions are negative. On day 2, however, the NO_x contributions become positive and then
31 increase from day 2 to day 3. The total of the NO_x contributions from all sources at 42 h is

1 essentially the same as the total VOC contribution, and at 66 h, the total NO_x contribution is
2 twice the total VOC contribution. The increasing importance of the NO_x contributions is due
3 to the increasing VOC/NO_x, which is 10-20 ppbC/ppb⁻¹ after 36 h, resulting in NO_x-limited O₃
4 formation.

5 ~~Other methods for source apportionment assign O₃ only to VOC and NO_x emissions and thereby~~
6 ~~likely include any contributions of CO and HONO emissions with the VOC and NO_x emissions,~~
7 ~~respectively. However, the~~ The PIM can separate the contributions of all emission species.

8 Figure 4 shows that the CO contributions from on-road and non-road vehicles are not negligible
9 compared to the VOC contributions of these sources. For on-road vehicles, the CO
10 contributions are generally 20-45% of the VOC contributions, and for non-road vehicles, 10-
11 30%. HONO emissions are assigned only to on-road and non-road vehicles and are small (0.8%
12 of NO_x, Table 1). For both of these sources, their HONO emissions contribute < 0.35 ppb to
13 the ΔO₃.

14 Figure 5 displays the source contributions to ΔO₃ obtained with the 3 paths. (The contributions
15 of all emission species from a source are combined together.) Results for the Diag and NO_xF
16 path are similar. For these paths, on-road vehicles have the largest and non-road vehicles the
17 second-largest contributions during most of the simulation, and the “other” category contributes
18 <3 ppb to ΔO₃. However, industrial sources are more important than fuel combustion for the
19 Diag path and the reverse is true for the NO_xF path. The source contributions for the VO_xF
20 path are distinctly different. Over most of the simulation, the ranking of the contributions is
21 industrial sources > non-road vehicles > on-road vehicles, the opposite of the Diag path. Also,
22 fuel combustion has a negative contribution over the entire simulation and the other category
23 has a larger contribution (up to 6.5 ppb) than for the Diag and NO_xF paths.

24 The different results for the VO_xF path can be explained by the fact that the NO_x emissions are
25 controlled last on this path or, in terms of the integration, essentially only NO_x emissions are
26 added near $s = 0$. The sensitivity of O₃ to these emissions is large and positive near $s = 0$ (Fig.
27 3) because the VOC/NO_x ratio is high in the background case. However, the VOC/NO_x ratio
28 decreases rapidly as s increases along the VO_xF path, the sensitivity to NO_x emissions becomes
29 negative, and O₃ formation becomes VOC-limited for most of the path. Thus, fuel combustion
30 has a negative source contribution because its emissions are mostly NO_x, and industrial sources
31 have the largest positive contribution because they have the largest VOC emissions. Also, non-
32 road vehicles have a larger contribution than on-road vehicles because both sources have a

1 similar magnitude of VOC emissions but on-road vehicles have 82% more NO_x emissions,
2 which suppress O₃ formation on the VOC-limited section of the path.

3 The source contributions to ΔFORM for the 3 paths are also in Fig. 5. For the Diag path, the
4 relative importance of the sources on days 2 and 3 is the same for ΔFORM as for ΔO₃, and this
5 is also true for the NO_xF path. For the VO_{CF} path, the on-road and non-road vehicles
6 contribute more to ΔFORM than the industrial sources, but the reverse is true for the
7 contributions of these sources to ΔO₃. The on-road and non-road vehicles have the largest
8 contributions to ΔFORM on each path because these sources have the largest primary FORM
9 emissions and the largest emissions of olefins, which are important precursors to secondary
10 FORM from oxidation reactions (Table S2).

11 Figure S3 contains the apportionment of ΔNO₂ and ΔHNO₃ to sources. The source
12 contributions to ΔNO₂ for the Diag and NO_xF paths are quite similar; those for the VO_{CF} path
13 differ in that the contributions of the industrial sources and other category are primarily negative
14 after 18 h. The source contributions to ΔHNO₃ for the Diag and NO_xF paths are again quite
15 similar, and the ranking of the sources in importance is the same as the ranking of their NO_x
16 emissions. The source contributions to ΔHNO₃ for the VO_{CF} path are similar to those for the
17 other paths except that the contributions of non-road vehicles and fuel combustion are reversed
18 in importance. The reversal is likely due to the much larger VOC emissions from non-road
19 vehicles, which would enhance the oxidation of NO_x on the VOC-limited part of the path.

20

21 **5 Conclusions**

22 As shown in Sect. 4, the PIM can allocate the difference in concentration between two
23 simulations to emission sources. Consequently, the PIM ~~ensures~~ requires that the base-case
24 concentration minus the sum of the anthropogenic source contributions (difference δ) equals
25 the background concentration (within the accuracy of the numerical integration). Other
26 methods do not have this constraint. If δ is less than the background concentration, then the
27 method assigns too much importance to the anthropogenic sources and will give the impression
28 that reducing anthropogenic emissions will reduce the pollutant concentration more than will
29 actually occur (over-allocation of the anthropogenic increment to the anthropogenic sources).
30 Similarly, if δ is greater than the background concentration, the method assigns too little
31 importance to the anthropogenic sources (under-allocation of the anthropogenic increment).

1 The PIM ensures that the anthropogenic increments to O₃ and the other species are neither over-
2 nor under-allocated to the anthropogenic sources ~~(within the accuracy of the numerical~~
3 ~~integration).~~

4 Another advantage is that the PIM is based on an exact mathematical relationship that is
5 independent of the chemistry or model and does not require added relationships or
6 approximations. The PIM allows source contributions to be either positive or negative. If the
7 secondary pollutant formation is inhibited by emissions of some species, source, or geographic
8 area, the sensitivity to these emissions will be negative for at least some values of the scaling
9 parameter λ_m , and the integral in Eq. (2) may be negative.

10 Once a model has been modified to calculate the first-order sensitivities, the PIM requires only
11 very simple post-processing of model results, specifically, calculating a linear combination of
12 sensitivities from different simulations. This can be readily done with existing post-processing
13 packages such as the Package for Visualization of Environmental data (PAVE) or the
14 Visualization Environment for Rich Data Interpretation (VERDI) (Univ. of North Carolina,
15 2004, 2014). The PIM is not focused on just one species, e.g., O₃. The calculations needed to
16 allocate Δc_i for species i also generate all the information needed to allocate Δc_j for any other
17 species j predicted by the model, and there is minimal additional effort needed to allocate Δc_j
18 for the second and subsequent species. Finally, the PIM highlights the importance of the
19 background simulation. For a simulation with anthropogenic emissions included to be useful
20 in designing emission controls, there is an implicit assumption that a simulation without the
21 anthropogenic emissions gives concentrations consistent with estimates for clean air. The
22 concentrations in the background simulation can be determined by an actual simulation or by
23 subtracting the sum of all the source contributions from the base-case concentrations.

24 In principle, there is an infinite number of source apportionments available from the PIM.
25 However, each source apportionment is linked to an emission-control strategy. If a control
26 strategy is defined along with the timing of the controls, the number of source apportionments
27 is reduced to just one.

28 The major disadvantage of the PIM is that it requires more computational effort than other
29 methods because the sensitivities must be determined at several emission levels between the
30 base and background simulations. This disadvantage is mitigated, to some degree, because the
31 additional simulations provide information on how concentrations and sensitivities will change
32 along the emission-control path.

1 The PIM has been applied in this work to a simplified configuration of CAMx that includes the
2 nonlinear chemistry but not transport or dispersion. However, transport and dispersion do not
3 involve nonlinear interactions among the species. Because the nonlinear dependence of the
4 sensitivities on the integration variable (Fig. 3) is driven by the nonlinear chemistry and a full
5 3-D configuration should not have any other sources of nonlinearity, the number of integration
6 points required for PIM for a 3-D configuration should be similar to the number required for
7 the simplified configuration (3 or 4) ([Dunker et al., 2015](#)).

8

9 **Supplementary information**

10 Application of the PIM to the special case involving the Taylor series expansion, input data and
11 emissions for the model simulations, accuracy in allocating ΔNO_2 and ΔHNO_3 to sources using
12 different integration formulas, comparison of the sum of the source contributions to the
13 anthropogenic increment at each hour, and source contributions to ΔNO_2 and ΔHNO_3 .

14

1 **References**

- 2 Bowman F. M.: A multi-parent assignment method for analyzing atmospheric chemistry
3 mechanisms, *Atmos. Environ.*, 39, 2519-2533, 2005.
- 4 Bowman, F. M. and Seinfeld, J. H.: Ozone productivity of atmospheric organics. *J. Geophys.*
5 *Res.*, 99, 5309-5324, 1994.
- 6 [Butler, T. M., Lawrence, M. G., Taraborrelli, D., and Lelieveld, J.: Multi-day ozone production](#)
7 [potential of volatile organic compounds calculated with a tagging approach, *Atmos. Environ.*,](#)
8 [45, 4082-4090, 2011.](#)
- 9 Cohan D. S., Hakami A., Hu, Y., and Russell A. G.: Nonlinear response of ozone to emissions:
10 source apportionment and sensitivity analysis, *Environ. Sci. Technol.*, 39, 6739-6748, 2005.
- 11 Dunker, A.M.: Efficient calculation of sensitivity coefficients for complex atmospheric models.
12 *Atmos. Environ.*, 15, 1155-1161, 1981.
- 13 Dunker, A. M.: The decoupled direct method for calculating sensitivity coefficients in chemical
14 kinetics, *J. Chem. Phys.*, 81, 2385-2393, 1984.
- 15 Dunker, A. M., Yarwood, G., Ortmann, J. P., and Wilson, G. M.: The decoupled direct method
16 for sensitivity analysis in a three-dimensional air quality model- implementation, accuracy, and
17 efficiency, *Environ. Sci. Technol.*, 36, 2965-2976, 2002a.
- 18 Dunker, A. M., Yarwood, G., Ortmann, J. P., and Wilson, G. M.: Comparison of source
19 apportionment and source sensitivity of ozone in a three-dimensional air quality model,
20 *Environ. Sci. Technol.*, 36, 2593-2964, 2002b.
- 21 [Dunker, A. M., Koo, B., and Yarwood, G.: Source apportionment of the anthropogenic](#)
22 [increment to ozone, formaldehyde, and nitrogen dioxide by the path-integral method in a 3-D](#)
23 [model, submitted for publication, 2015.](#)
- 24 efunda: available at http://www.efunda.com/math/num_integration/findgausslegendre.cfm
25 (last access: 29 January 2014), 2014.
- 26 [Emmons, L. K., Hess, P. G., Lamarque, J.-F., and Pfister, G. G.: Tagged ozone mechanism for](#)
27 [MOZART-4, CAM-chem and other chemical transport models, *Geosci. Model Dev.*, 5, 1531-](#)
28 [1542, 2012.](#)

1 ENVIRON: Comprehensive Air Quality Model with Extensions, available at;
2 <http://www.CAMx.com> (last access: 15 May 2013), 2013.

3 [Grewe, V., Tsati, E., and Hoor, P.: On the attribution of contributions of atmospheric trace gases](#)
4 [to emissions in atmospheric model applications, *Geosci. Model Dev.*, 3, 487-499, 2010.](#)

5 Isaacson, E. and Keller, H. B.: Analysis of Numerical Methods, John Wiley, New York, 1966.

6 Kaplan, W.: Advanced Calculus, Addison-Wesley, Reading, Massachusetts, 1959.

7 Koo, B., Dunker, A. M., and Yarwood, G.: Implementing the decoupled direct method for
8 sensitivity analysis in a particulate matter air quality model, *Environ. Sci. Technol.*, 41, 2847-
9 2854, 2007.

10 Koo, B., Wilson, G. M., Morris, R. E., Dunker, A. M., and Yarwood, G.: Comparison of source
11 apportionment and sensitivity analysis in a particulate matter air quality model, *Environ. Sci.*
12 *Technol.*, 43, 6669-6675, 2009.

13 Koo, B., Chien, C.-J., Tonnesen, G., Morris, R., Johnson, J., Sakulyanontvittaya, T.,
14 Piyachaturawat, P., and Yarwood, G.: Natural emissions for regional modeling of background
15 ozone and particulate matter and impacts on emissions control strategies, *Atmos. Environ.* 44,
16 2372-2382, 2010.

17 Kwok, R. H. F., Napelenok, S. L., and Baker, K. R.: Implementation and evaluation of PM2.5
18 source contribution analysis in a photochemical model, *Atmos. Environ.*, 80, 398-407, 2013.

19 Marmur, A., Park, S.-K., Mulholland, J. A., Tolbert, P. E., and Russell, A. G.: Source
20 apportionment of PM2.5 in the southeastern United States using receptor and emissions-based
21 models: conceptual differences and implications for time-series health studies, *Atmos.*
22 *Environ.*, 40, 2533-2551, 2006.

23 Mysliwiec, M. J. and Kleeman M. J.: Source apportionment of secondary airborne particulate
24 matter in a polluted atmosphere, *Environ. Sci. Technol.*, 36, 5376-5384, 2002.

25 Napelenok, S. L., Cohan D. S., Hu, Y., and Russell A. G.: Decoupled direct 3D sensitivity
26 analysis for particulate matter (DDM-3D/PM), *Atmos. Environ.*, 40, 6112-6121, 2006.

27 Pierce, T., Geron, C., Bender, L., Dennis, R., Tonnesen, G., and Guenther, A.: Influence of
28 increased isoprene emissions on regional ozone modeling, *J. Geophys. Res.*, 103, 25611-25629,
29 1998.

1 Simon, H., Beck, L., Bhawe, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A.,
2 Reff, A., Sarwar, G., and Strum, M.: The development and uses of EPA's SPECIATE database,
3 Atmospheric Pollution Research, 1, 196-206, 2010.

4 Stein, U. and Alpert, P.: Factor separation in numerical simulations, J. Atmos. Sci., 50, 2107-
5 2115, 1993.

6 Tao, Z., Larson S. M., Williams A., Caughey, M. and Wuebbles D. J.: Area, mobile, and point
7 source contributions to ground level ozone: a summer simulation across the continental USA,
8 Atmos. Environ., 39, 1869-1877, 2005.

9 Tong, D. Q. and Mauzerall, D. L.: Summertime state-level source-receptor relationships
10 between nitrogen oxides emissions and surface ozone concentrations over the continental
11 United States, Environ. Sci. Technol., 42, 7976-7984, 2008.

12 University of North Carolina, Package for Analysis and Visualization of Environmental Data,
13 version 2.3, available at:
14 http://www.ie.unc.edu/cempd/EDSS/pave_doc/EntirePaveManual.html (last access: 17
15 November 2014), 2004.

16 University of North Carolina, Visualization Environment for Rich Data Interpretation, version
17 1.5, available at: <https://www.cmascenter.org/verdi/> (last access: 17 November 2014), 2014.

18 U.S. EPA: Carbon Bond and SAPRC Speciation Profiles, available at:
19 <http://www.cmascenter.org/download/data.cfm> (last access: 19 November 2013), 2013a.

20 U.S. EPA: 2008 National Emissions Inventory, Version 3, Technical Support Document,
21 September 2013-Draft, available at: <http://www.epa.gov/ttn/chief/net/2008inventory.html> (last
22 access: 20 November 2013), 2013b.

23 U.S. EPA: CAIR Platform Data, available at:
24 <http://www.epa.gov/ttn/chief/emch/temporal/index.html> (last access: 24 December 2013),
25 2013c.

26 Wagstrom, K. M., Pandis, S. N., Yarwood, G., Wilson, G. M. and Morris, R. E.: Development
27 and application of a computationally efficient particulate matter apportionment algorithm in a
28 three-dimensional chemical transport model, Atmos. Environ., 42, 5650-5659, 2008.

1 Wang, H., Jacob, D. J., LeSager, P., Streets, D. G., Park, R. J., Gilliland, A. B., and van
2 Donkelaar, A.: Surface ozone background in the United States: Canadian and Mexican
3 pollution influences, *Atmos. Environ.*, 43, 1310-1319, 2009.

4 Wang, Z. S., Chien, C. J. and Tonnesen, G. S.: Development of a tagged species source
5 apportionment algorithm to characterize three-dimensional transport and transformation of
6 precursors and secondary pollutants, *J. Geophys. Res.*, 114, D21206, doi:
7 10.1029/2008JD010846, 2009.

8 Yang, Y-J., Wilkinson, J. G. and Russell, A. G.: Fast, direct sensitivity analysis of
9 multidimensional photochemical models, *Environ. Sci. Technol.*, 31, 2859-2868, 1997.

10 Yarwood, G., Morris, R. E., Yocke, M. A., Hogo, H. and Chico, T. Development of a
11 methodology for source apportionment of ozone concentration estimates from a photochemical
12 grid model, in: *Proceedings of the 89th Annual Meeting of the Air & Waste Management*
13 *Association*, Air and Waste Management Association, Pittsburgh, PA, Paper 96-TA23A.06,
14 1996.

15 Yarwood, G., Gookyoung H., Carter, W.P.L., and Whitten, G.Z.: *Environmental Chamber*
16 *Experiments to Evaluate NO_x Sinks and Recycling in Atmospheric Chemical Mechanisms*,
17 *Texas Air Quality Research Program Project 10-042*, University of Texas, Austin, 2012.

18 Zhang, Y., Wang, W., Wu, S.-Y., Wang, K., Minoura, H. and Wang, Z. Impacts of updated
19 emission inventories on source apportionment of fine particle and ozone over the southeastern
20 U.S., *Atmos. Environ.*, 88, 133-154, 2014.

21

1 Table 1. Summary of daily emission rates used in the base-case simulation.

Species	Emission Rate (mol day ⁻¹ km ⁻²)					
	Biogenic Sources ^a	Fuel Combustion	Industrial Sources	On-road Vehicles	Non-road Vehicles	Other Sources
NO	13.5	77.4	19.7	132.9	73.2	1.9
NO ₂	0.00	8.60	2.19	13.59	7.48	0.21
HONO	0.00	0.00	0.00	1.18	0.65	0.00
CO	35.9	51.8	58.2	1158.4	683.0	57.0
VOC	166.8	6.1	244.3	129.9	115.1	59.3
VOC/NO _x ^b	29.8	0.09	16.6	1.4	2.4	31.8

2 ^a Includes lightning

3 ^b NO_x = NO + NO₂. VOC/NO_x units are mole C (mole NO_x)⁻¹

4

1 Table 2. Average error and bias for different numerical integration formulas. The sum of the
 2 source contributions calculated using the formula is compared to the anthropogenic increment
 3 of O₃ or FORM.

Path	Formula ^a	Mean Absolute Error ^b (ppb)	Mean Bias ^b (ppb)
O ₃ Increment			
Diag	TR2	65.93	65.93
Diag	GL2s	7.38	-7.36
Diag	GL2r	5.95	5.71
Diag	GL3s	3.32	-3.30
Diag	GL3r	1.64	-1.49
Diag	GL4s	1.51	-1.50
Diag	GL4r	1.54	-1.49
NOxF	GL3s	2.20	2.15
NOxF	GL3r	7.73	-7.67
NOxF	GL4s	1.57	-1.54
VOCF	GL3s	7.56	-7.32
VOCF	GL3r	10.46	9.62
VOCF	GL4s	4.68	-4.63
FORM Increment			
Diag	TR2	2.45	2.45
Diag	GL2s	0.21	-0.20
Diag	GL2r	0.19	0.19
Diag	GL3s	0.12	-0.12

Diag	GL3r	0.04	0.02
Diag	GL4s	0.05	-0.04
Diag	GL4r	0.03	-0.02
NOxF	GL3s	0.11	-0.10
NOxF	GL3r	0.08	-0.01
NOxF	GL4s	0.08	0.08
VOCF	GL3s	0.30	-0.30
VOCF	GL3r	0.17	0.11
VOCF	GL4s	0.09	-0.08

1 ^a TR2 = trapezoidal rule, 2 points. GL_n_x = Gauss-Legendre formula using *n* points and *x* as the
2 integration variable.

3 ^b Hourly average over the 3-day simulation.

4

1 **Figure captions**

2 Figure 1. Three possible integration paths when the concentration difference between the base
3 (point B) and background (point b) cases is allocated to two sources with emissions scaled by
4 λ_1 and λ_2 . Path 1: equal control of emissions from both sources (diagonal path). Path 2:
5 emphasis on control of emissions from source 1 first followed by control of emissions from
6 source 2. Path 3: opposite of Path 2. Points b1 and b2 have the emissions from the background
7 case plus source 1 and source 2, respectively.

8 Figure 2. Results from the 2-cell model simulations. Ozone and formaldehyde concentrations
9 for the base case and the background case and the difference between them (anthropogenic
10 increment).

11 Figure 3. Dependence of the integrands for allocating O₃ to sources on the distance s along the
12 Diag, NO_xF and VO_{CF} paths. The integrand (Eq. (4)) is calculated at the time of peak O₃ on
13 day 3 (66 h).

14 Figure 4. Contributions of sources and VOC, NO_x, CO, and HONO emissions to the
15 anthropogenic O₃ increment. Results are for the Diag path.

16 Figure 5. Apportionment of the anthropogenic O₃ increment (left) and the FORM increment
17 (right) to sources using the Diag, NO_xF, and VO_{CF} emission-control paths.

18