Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-296-RC2, 2019

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Local Fractions – a method for the calculation of local source contributions to air pollution, illustrated by examples using the EMEP MSC-W model (rv4_33)" *by* Peter Wind et al.

Anonymous Referee #2

Received and published: 16 December 2019

Wind et al. describe a method for source attribution of primary particulate matter in the EMEP MSC-W model which they call "local fractions". Source attribution of air pollution is a useful model diagnostic and has important policy applications, for example in the calculation of source/receptor relationships. Several different methods for source attribution of air pollution exist already, but in each case different trade-offs must be made in order to keep the required computational resources within reasonable bounds. The method of local fractions appears to make a different set of trade-offs compared with other source attribution methods, and so therefore appears to make a unique

contribution to the air pollution modelling literature. So far it only seems that the method is useful for source attribution of primary PM, but nevertheless I think the manuscript falls within the scope of GMD.

I have a number of major concerns with the manuscript in its current form, which must be addressed before it could be reconsidered for publication. Of primary concern is the use of sometimes extremely vague language throughout the manuscript, which makes it difficult to follow exactly what the authors are describing. I also think the authors can do a much better job of explaining how their work fits in the context of other source attribution methodologies. The authors are quick to point out the advantages of their method, but any reader not already familiar with the existing literature will not understand the significance of the advantages and disadvantages of the local fractions method. Some additional context would help a lot here.

Specific comments

Page 1, line 4: "distinguish a large" should probably be "distinguish the contribution of a large".

Page 2, line 17: The references given here are not a good representation of studies which trace pollutant concentration back to emitted species by tagging (with or without consideration of nonlinearities). For PM, the study of Kranenburg et al. (2013) should be mentioned here. It is mentioned later in the manuscriupt, but should also be discussed here because it is a tagging method. For ozone, both Wang et al. (1998) and Wu et al. (2011) actually avoid the nonlinearities by tagging ozone based on its geographical region of formation, rather than the geographical region in which precursors are emitted. The reference to Grewe et al. (2013) is more appropriate here, since it is capable of attributing ozone to its emissions, but the authors could consider instead referencing the most up-to-date version of this method, as described by Grewe et al. (2017) and an actual application of this method for ozone attribution by Mertens et al. (2018). Alternative approaches also exist, which make different trade-offs. The approaches shared by Dunker et al. (2002) and Kwok et al. (2015), which tag ozone based on the chemical regime are both well-established and should be referenced in any discussion of modelled source attribution. Yet another approach was described recently by Butler et al. (2018) and has been applied by Lupascu et al. (2019).

The authors are right that large number of tracers can rapidly make tagging computationally expensive, but they should also point out that techniques exist to keep these problems in check. For example Butler et al. (2018) Restrict the number of tracers to a carefully chosen set of representative source sectors; Grewe et al. (2017) make use of the concept of "chemical families" to keep the number of tagged species within reasonable limits; and Lupascu et al. (2019) restrict the length of their simulation period to focus on a pollution episode of interest. Each of these approaches brings different trade-offs, but in each case also significant advantages: the ability to perform source attribution for secondary pollutants; and the ability to perform source attribution for long-range transport. These trade-offs are especially interesting in the context of the present manuscript, since one of the major ways in which the computational complexity of the local fractions method is kept computationally simple is by restricting the size of the "local region" for which the source attribution is performed.

Page 2, line 33: "can be built ... relatively easily" is a vague statement. More detail is needed here.

Page 3, line 5: An important detail missing here is that origins of the pollutants being tracked must be restricted to emissions within a "local region". This should be made clear up front, rather than making the reader wait 2 more sub-sections to find this out.

Page 3, line 8: "source regions" is very vague here. It would help the reader to know already at this stage that the present implementation considers each grid cell as a separate source region, but that in principle the method can be expanded to work with larger source regions.

Page 4, lines 7-8: "reasonable" cost and "preset" numbers of grid cells are used very

vaguely here. These terms are discussed later in the manuscript, but most readers would benefit from forward references to the relevant sections here.

Page 4, lines 12-13: "usually not necessary..." is very vague here. The authors show later that in fact extending vertical resolution to at least the height of the PBL is useful. The authors should also note that this also applies only to the pollutants they assess in their manuscript. Transport in the free troposphere is important for some pollutants such as ozone.

Page 4, line 17: Keeping the size of the LF array down to a reasonable size appears to be the main trade-off associated with this method, and this should be acknowledged here.

Page 4, line 18: The concept of the "local region" is first used here, and only implicitly defined by its context. It would help most readers tremendously if this concept could be introduced a lot earlier, with the explanation that setting the size of the local region represents the major trade-off with using this method.

Page 5, line 20: This is a good point, and could perhaps be mentioned earlier in the manuscript where the authors make the claim that their method is easy to implement in different CTMs.

Page 6, lines 8-9: Generally throughout the manuscript it would also be nice to have some discussion of the limitations of the method.

Page 6, line 27: "given distance" is very vague here. This is why it would be good to already have a well defined and discussed concept of what the "local region" is and why it is needed in this method.

Page 8, line 7: There is no justification given here for choosing 8 levels.

Page 8, lines 11-12: More detail is needed here. Why exactly is this "not a problem" and "actually an advantage" compared with the direct method?

Figures 5 and 6: Labels are missing for the x-axes. It would also be better to reverse the vertical ordering of the line color legends so that they correspond with the ordering of the lines in the plots.

Page 11, line 11: Does this mean that more vertical levels would be required when simulating summer months?

Figure 7: Here the size of the local region is referred to by its "distance", whereas everywhere else in the paper the actual size of the local region in grid cells is given. I presume that a distance of 20 is the same as a local region of 41x41, but this is not at all clear. Please use a consistent way of describing the size of the local region.

Page 14, lines 18-19: This seems like speculation (sub-optimal use of cache memory). Another possible explanation is that the extra memory requirements could be leading to increased communication overhead.

Page 15, line 1: Is this "substantial amount of time" already included in Table 1, or is this additional time? Can the authors quantify this?

Page 15, lines 2-3: The example given (local region of size 21x21x1) is not used anywhere else in the manuscript. It would be more useful to know about the extra storage requirements of the configurations which are actually evaluated in the manuscript. The authors should expand Table 1 to also include the extra storage requirements of the configurations given in this Table.

Page 15, line 5: "origin of pollutants" should acknowledge the limitations of the method. Based on the evaluation presented by the authors, it seems that this method can currently only analyse the origin of some kinds of primary pollutants.

Page 15, lines 11-13: The authors have not provided any other details about interactive graphical user interfaces. Is this something for future work? Or can the authors already provide a reference for this?

Page 15, line 17: This is an important point, and not necessarily a disadvantage of

the method. For some applications it may be acceptable to simply know that a certain amount of pollution originates from outside the local region. This provides some justification for other trade-offs which are made when using this method.

Page 15, line 19: Can the authors go into more detail about the "double counting" problem and how their approach solves it?

Page 15, lines 23-24: Which of the "several" problems are avoided and how? This text is way too vague.

Page 16, line 14: It seems to me that the local fractions deliver information about contributions, not sensitivities.

Page 16, line 14: Why wouldn't the local fractions add up to 100%, and why isn't this a problem? It seems that the final sentence of the manuscript creates all sorts of problems for the interested reader. The authors could consider simply deleting this sentence, and merging the previous sentence with the previous paragraph.

References

Emmons et al. (2012) is in the reference list, but not cited in the text.

The reference list is also quite short for a paper on source attribution of modeled air pollution, which is a well-established field. Discussion of the suggested additional literature would help to place the current work in context, as described in more detail above.

Additional literature

Butler, T., Lupascu, A., Coates, J., and Zhu, S.: TOAST 1.0: Tropospheric Ozone Attribution of Sources with Tagging for CESM 1.2.2, Geosci. Model Dev., 11, 2825-2840, https://doi.org/10.5194/gmd-11-2825-2018, 2018.

Dunker, A., Yarwood, G., Ortmann, J., and Wilson, G.: Comparison of source apportionment and source sensitivity of ozone in a three-dimensional air quality model, Environ. Sci. Technol., 36, 2953-2964, https://doi.org/10.1021/es011418f, 2002.

Grewe, V., Tsati, E., Mertens, M., Frömming, C., and Jöckel, P.: Contribution of emissions to concentrations: the TAGGING 1.0 submodel based on the Modular Earth Submodel System (MESSy 2.52), Geosci. Model Dev., 10, 2615-2633, https://doi.org/10.5194/gmd-10-2615-2017, 2017.

Kwok, R. H. F., Baker, K. R., Napelenok, S. L., and Tonnesen, G. S.: Photochemical grid model implementation and application of VOC, NOx, and O3 source apportionment, Geosci. Model Dev., 8, 99-114, https://doi.org/10.5194/gmd-8-99-2015, 2015.

Lupascu, A. and Butler, T.: Source attribution of European surface O3 using a tagged O3 mechanism, Atmos. Chem. Phys., 19, 14535-14558, https://doi.org/10.5194/acp-19-14535-2019, 2019.

Mertens, M., Grewe, V., Rieger, V. S., and Jöckel, P.: Revisiting the contribution of land transport and shipping emissions to tropospheric ozone, Atmos. Chem. Phys., 18, 5567-5588, https://doi.org/10.5194/acp-18-5567-2018, 2018.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-296, 2019.