Atmospheric
Chemistry
and Physics
Discussions

Open Access

ACPD

Interactive
comment

Interactive comment on "Long-term measurements (2009–2015) of non-methane hydrocarbons (NMHCs) in a megacity of China: implication for emission validation and source control" *by* Yarong Peng et al.

Anonymous Referee #3

Received and published: 13 April 2021

Title: Long term measurements (2009-2015) of non-methane hydrocarbons (NMHCs) in a megacity of China: implication for emission validation and source control

Authors: Peng et al.

The study of Peng et al. aimed to investigate the characteristics and sources of non-methane hydrocarbons (NMHCs) in Shanghai, China. The study was based on the

Printer-friendly version

Discussion paper

2009-2015 NMHC dataset, volatile organic compounds ratio, and positive matrix factorization (PMF)-derived factor temporal variation and trend analyses. The papers based on temporal variations, trend analysis, PMF, or toluene-to-benzene ratio are present in the literature for many years. The whole manuscript, the applied methods for data analysis, discussion, and conclusions are too generic and too basic. I would expect a more advanced methodological approach for revealing factors governing NMHCs' environmental fate, the evolution of their sources and sinks, or their interrelations with meteorological conditions. Moreover, I feel that the scientific novelty is missing.

Two of the three main objectives of this study (lines 67-69) were to "assess the source evolutions of NMHCs over time" and to "validate the speciation of emissions inventory primarily", but the authors used basic and straightforward methodology I believe is not capable of achieving them. I'll describe my concerns in the following text. Namely, regarding the applied methodology and concept, I have found some major shortcomings:

1. The authors excluded the influence of meteorological parameters: "we performed MLR model with ambient NMHCs and meteorological variables (temperature, wind speed and direction, air pressure and relative humidity) based on stepwise multiple linear regression". The first point I would like to address is the restriction to just five meteorological parameters. If the aim was to assess the impact of meteorology, the meteorological context had to be described broadly by using some of the available modeled data, i.e., Global Data Assimilation System with more than 20 important parameters. 2. The authors used 16 species for the source apportionment. Why the authors didn't use some of the standard procedures to determine the number of PMF-derived factors such as the scree plot (line 260)? Why the authors restrict the study to five? 3. 2009-2015 is a long period. If NMHCs concentration varied, why the authors didn't apply PMF for shorter periods, i.e., for each year separately (the data set is large enough), and try to see if a new source emerged or the composition of a particular source changed over time? This could happen, having in mind the influence of mitigation measures, technology development, or easily excluded meteorological

condition change (line 28) during 2009-2015. The analysis applied in this manuscript covered only PMF-derived factor trend and time variations, but not the variations in their composition which I think must have been included. 4. The authors argue that their primarily PMF-based analysis points out the omissions of emission source inventories (lines 35-36), which is the argument that I cannot agree with. To find the omissions in emission source inventories, I believe, a significantly advanced methodology/research has to be conducted with a disproportionately larger number of environmental factors included than the factors available in this study. There are many methods capable of modeling complex, heterogeneous, noisy, nonlinear, interactive, etc. interrelations between environmental factors such as machine learning (i.e. extreme gradient boosting). Moreover, there are many explainable artificial intelligence methods capable of explaining the derived dependencies in an extremely complex urban environment (i.e. Shapley additive explanations). PMF is not capable of meeting the goals of this study as the authors claim.

Proofreading by a native speaker is mandatory. It will clear some sentences, statements, and grammatical issues. I didn't make the corrections throw-out the manuscript because proofreading could significantly improve it.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-1108, 2020.