

Interactive comment on “**Estimation of Secondary PM_{2.5} in China and the United States using a Multi-Tracer Approach**” by Haoran Zhang et al.

The manuscript demonstrates the multi-tracer estimation algorithm (MTEA), to identify the primary and secondary components from routine observation of PM_{2.5}, and validates the method by comparing the long-term and short-term measurements of aerosol chemical composition in China and a network from the United States. This method provides a useful and uncomplicated way to estimate primary and secondary PM, using routine observation species and emission inventories. This manuscript aims to address important questions quantifying primary and secondary aerosols and is within the scope of ACP.

However, regarding the method itself, the method should be carefully introduced with more details. The validation part is a bit weak, and should be strengthened in the next version. It is vital because only with good validation can one trust the result from the model. In addition, in the result and discussion part, the discussion is superficial, which needs to be improved in depth, and backed up by more scientific evidence and/or publications.

As a conclusion, the manuscript provides a novel algorithms in primary and secondary particle concentrations, however, the manuscript is not carefully written from the perspective of science and scientific writing, with certain degree of improvement for publication in ACP. Therefore, this manuscript needs a major revision in terms of major context and English language.

Major comments:

- Introduction: the introduction is poorly written and need to be re-write. If I were you, I would write the introduction based on this outline: 1) introduction of atmospheric aerosols, including sources, type, chemical composition and impacts on air quality, human health and climate, 2) summarise other studies, you must state what has been achieved and what is the current challenging, 3) what is your paper about, how this paper can narrow the gap.

In the current version, the point 1) is addressed, but should be introduced in smoother way. The author is trying to address the point 2), but the studies mentioned in the paragraph 3 in page 3 look not very relevant. For example, the author summarises the online and offline studies, which is good, and people can see the drawbacks of field and lab measurement to study the PPM and SPM, so the next paragraph should state to overcome these drawbacks, people use model to study the PPM and SPM, and should also state what these model studies have achieved and/or the drawbacks of these method. Finally, this paragraph can lead the final paragraph in the introduction, namely, introduce this study and how this study advances the model studies on PPM and SPM estimation.

- Methodology: the methodology part is written in a reasonable logic, but the author needs to pay more attention to specify the technical details, e.g., the definition of some terms.
- Model validation: this part straightforwardly delivers the good validation result between model and observation. Good correlation is shown in this part, suggesting good model performance. However, this part also requires more interpretation on the model's over/underestimation behaviour compared to observation, which is now absent. Ideally, the author should focus most on this part, because only when the model is reasonably validated can we trust the result and make further interpretation on the result. Therefore, from my own perspective, the author should strengthen this part.
- Result and discussion: this part also very straightforwardly and logically reports the results. However, the interpretation of results should be more comprehensive and backed up by previous studies and/or solid evidence, which is absent now and needs to be added. In addition, the discussion of the result is very superficial, lacking depth, which should also be improved.
- Conclusion: it summarises the significance of the study, but one or two paragraphs need to be re-written, based on the revised context in Section 4.

Other comments:

Introduction:

-Page 3 Line 4: please check the size of these emission sources. Dust resuspension and brake wear are typically in coarse mode. Also sea salt and sea spray by definition have certain degree of overlap.

-Page 3 Line 6: please specify how the secondary aerosol is formed, e.g., gas-particle partitioning mechanism for new particle formation.

-Page 3 paragraph 1: please restructure this paragraph, by introducing primary and secondary, instead of saying “one source....and the other source....”. Better still, avoid to say the sources of PM_{2.5}, because the sources you listed as primary sources produce more coarse mode particles, rather than accumulation mode and/or Aitken mode, e.g., dust, sea spray.

-Page 3 Line 18: please give an example of the toxicities on human health.

-Page 3 Line 21: please define “individual” and “short-term”, and make it more clear.

-Page 3 Line 26: please give a range of percentage to support the word “dominant”.

-Page 3 Line 28: please also mention the instrument and method using in this study.

-Page 3 Line 29: actually, this sentence describes the general situation. Even if you don't cite this study, you still can make the statement. Please specify what the significance of the study and how it support the main point in this paragraph.

-Page 3 Line 31: please note that the offline filter measurement is typically considered as a cheaper method compared to the online measurement. Using offline filter measurement, we can collect samples from various sampling sites for a long period.

-Page 4 Line 5: to make it consistent, it should be “Section 3” instead of “section 3”.

-Page 4 Line 14: MEE is not defined in the previous text.

Methodology:

-Page 4 Line 24: why specifically CO and why don't select black carbon? Here the author states that he/she uses the PMC to track flying dust, but in the introduction, the author states dust belongs to the size range of PM_{2.5}. Please specify.

-Page 4 Line 31: "fine mode" is not defined. It seems that it refers to PM_{2.5}, but please specify and give the definition of "fine mode" and "coarse mode" in the study.

-Page 5 Line 3: "OC" is not defined in the previous text.

-Page 5 Line 14: good to show the table, but it is better to include 0.01 and 0.05 case, as a sensitivity test.

-Page 5 Line 25: why is this range and why is this range reasonable?

-Page 5 Line 26: "each varying ratio may obtain a series of SPM, along with a coefficient of"? What does the "series" mean? Do you mean time series?

-Page 5 Line 27: what is the R (coefficient of determination)? Is it a Pearson correlation coefficient? How is it defined? How is it calculated? Please also introduce Figure S1 and explain it a bit here or in the SI.

-Page 5 Line 27: please also introduce Figure S1 and explain it a bit here or in the SI. Regarding the y-axis, does it refer to R² or *p*-value? What does the *p*-value mean?

-Page 5 Line 27: how do you know if primary and secondary PM are from the same or different sources? What if they are from similar sources, which side of the green are would you choose?

-Page 6 Line 13: MEE is not defined in the previous text. Which region does MEE network cover, whole China?

-Page 6 Line 16: under what guideline and/or operation standard does MEE do the measurement? Presumably, Ambient air quality standards (GB 3095-2012)?

-Page 6 Line 20: in Table S3, is the PM_{2.5} concentration the annual mean PM_{2.5} concentration? Please specify.

-Page 6 Line 24: "the major gaseous and particle pollutants" sounds awkward. Maybe change it to "the major gaseous pollutants and particle". Here particle pollution is presumable PM, right? You can just say "the major gaseous pollutants and particulate matter".

-Page 6 Line 26: just saying "for exmaple" is fine.

-Page 6 Line 30: awkward here. Please read the sentence again, it is hard to identify the verb in the sentence. Does the author want to say something like: these studies employ offline filter-based measurement? If yes, please change the word “employed” to “employ”.

-Page 7 Line 1: change “compositions” to “composition”, the same applies to Page 7 line 8

-Page 7 Line 1: change “for directly comparing” to “to directly compare”.

-Page 7 Line 16: still, please specify the region covered by this network.

-Page 7 Line 24: why does the author lower the time resolution? Presumably for the computational efficiency? If so, please specify.

Model validation:

-Page 8 Line 23: why is the RMA regression model chosen? Please specify the reason. Has the author tried other regression models?

-Page 8 Line 24: why are the values over “ $30 \mu\text{g}\cdot\text{m}^{-3}$ higher than the city average” removed? Please specify the reason. In addition, why is the number $30 \mu\text{g}\cdot\text{m}^{-3}$ chosen?

-Page 8 Line 7: in the figure, the author specifies the “r”, “s” and “d” in the caption, but not here. Please also explain the meaning here.

-Page 8 Line 9: what does “NMB” mean? How is it calculated?

-Page 8 Line 9: please also try to interpret the reason of over/underestimation.

-Page 8 Line 13: why not report “NMB” here?

-Page 9 Line 6: here the author is trying to interpret the discrepancies between model and observation. However the sentence here looks like he/she is trying to summarise the general reasons for discrepancies by saying “However, we find that there are still a few discrepancies between the estimated and observation-based results, and the main reasons might be”, but the author specifies the reason for for discrepancies for some individual sites. Therefore, the author can state that the he/she observed the discrepancies in some cities, and then specify the reasons for each site.

-Page 9 Line 19: how are the spatial, temporal and spatio-temporal correlation calculated? I guess for spatial, the author gets the average concentration of each site and plot them in the estimation vs observation space? Please specify how they are calculated.

-Page 9 Line 20: awkward, there is one value for estimated concentration and one value for observed concentration, then the author plots them in the estimation vs observation space. So for each pair, there is only one point for one time point or one site.

-Page 9 Line 26: here, please also specify why the Pearson correlation coefficient for SPM is higher than it for PPM.

-Page 9 Line 29: here, please also specify why the slope for PPM is higher than it for SPM.

Result and discussion:

-Page 10 Line 11: In Figure 5, what is the size of the solid circle? The legend is missing here for this information. Please specify this.

-Page 10 Line 18: the high energy consumption cannot directly link to the serious pollution, it also depends on what sort of energy. In the northern part of China, power plants combust coal for electricity and in the winter for heating as well. But we use nuclear power plants for this purpose, we don't have air pollution at all even if we have high energy consumption. This is the point. However, in the south, heating provided by the government and power plants is not very common. That is reason why China has this contrast between the north and the south. Please check the literature to back up this statement.

-Page 10 Line 24: please see how the atmospheric condition in the south is more favourable for SPM, and back up this argument with scientific literature.

-Page 10 Line 31: primary emission can also be regionally transported, This interpretation is not strong.

-Page 11 Line 9: the boundary layer height and temperature inversion is the major reason that particle concentration is higher at night than during the day. The author can simply say the high level of pollution is related to high emission from biomass burning and coal combustion as it is in the paper, and the more stagnant air condition (which means bad mixing condition) in the winter than summer.

-Page 11 Line 13: replace the “is” after PPM and SPM by “of”

-Page 11 Line 16: please interpret why the contribution is lower in autumn, and what this phenomena suggests.

-Page 11 Line 20: please explain why AOC is higher in warm season, and back this up by literature.

-Page 11 Line 22: this is a weak argument, because the primary pollutant can also be transported from the north to the south. The SPM can also be formed during the transport, or the air mass passes through a region which can facilitate the SPM formation. So there are many possibilities. In addition, please also add plots indicating the trajectories of air masses from the north to the south, or add some literature to support your statement of monsoon transport.

-Page 11 Line 27: it is an issue of English language. It should be “dotted blue line” and “solid blue line”.

-Page 12 Line 4: it is another issue of English language. The author uses past tense in Section 3, but the present tense here. Please use consistent tense throughout the paper.

-Page 12 Line 6: please cite studies to justify the interpretation here about ozone pollution.

-Page 12 Line 11: not convinced by the interpretation, please provide more evidence from your analysis and/or other studies.

-Page 12 Line 17: in the figure caption of Figure 7. Line 3 in the caption, “bule” should be “blue”. Please check the grammar and conjugate the verbs correctly for plural and third person singular.

-Page 13 Line 3: “were decreased” should be “decreased”, and also change for “increase” likewise.

-Page 13 Line 5: please define “BTH” in the previous text, or when it first appears in the main text.

-Page 13 Line 7: “its averaged value” here is a bit nebulous. Please specify the length of the period being averaged before COVID.

-Page 13 Line 8: here the present tense is used, whereas in two lines, a past tense verb is used. Please describe your result in a consistent tense.

-Page 13 Line 11: in the Figure 8 caption, “The application of M-TEA in estimating PPM/SPM during the COVID-19 lockdown” here has a typo, and it should be “The application of MTEA in PPM/SPM estimation during the COVID-19 lockdown”. The word “post” in the phrase “post-lockdown” in line 4 means “after”, but it is not what the author means here.

-Page 13 Line 17: I guess the author would like to say the condition is not good for atmospheric mixing. Please check the difference between mixing and diffusion.

-Page 13 Line 18: “was increased” should be “increased”. Please check this throughout the text.

-Page 13 Line 21: “YRD” is not defined in the previous text. Please define it.

-Page 13 Line 26: not accurate. The PM_{2.5} is the particulate matter with aerodynamic diameter smaller than 2.5 μm, which can be solid particle or liquid droplet. The PM_{2.5} can be directly emitted from primary sources, and some species in PM_{2.5} can be formed via chemical reactions. If the particles are directly emitted, then the word “precursor” in the sentence is wrong. Ozone in the troposphere is mainly produced by the reactions of NO_x and VOCs.

-Page 13 Line 32: please check the Figure S3 caption. There is typo “bule”, which should be “blue”

-Page 14 Line 1: please also include the cities which fail in the test in the figure when you reply the referee’s comment.

-Page 14 Line 2: please conjugate the verb correctly. “positive relationship” should be presumably “positive correlation”?

-Page 14 Line 3: the correlation coefficient “ r ” here should be *italic* “ r ”. Please change it throughout the text.

-Page 14 Line 9: please cite these “series of recent studies”.

-Page 14 Line 13: “compare” should be “compared”.

-Page 14 Line 14: the definition of BTH and YRD should be addressed in the early text.

-Page 14 Line 20: Figure 9 caption has some typos and mistakes in grammar. Which method is used specifically for “Least Squares method”? The trust-region Levenberg-Marquardt least orthogonal distance regression?

-Page 14 Line 20: This paragraph should be rewritten. It is hard to accept that the correlation coefficient r smaller than 0.3 shows some meaningful correlation, especially those with r about 0.09 and 0.14. There are very closed to 0, which means not correlated at all. From the subfigures in Figure 9, it is also hard to say there is any correlation.

-Page 15 Line 12: Should “reduces” be “increases”?

Conclusion:

-Page 16 Line 2: Should “offered” be “provided”?

-Page 16 Line 20: Please re-think about this paragraph.

-Page 16 Line 27: This paragraph looks like a repetition of the introduction. It should focus more on the outlook, e.g., how your study advance the current understanding, and what implication it might provide for future studies.