

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2021-1016-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2021-1016

Anonymous Referee #2

Referee comment on "Measurement report: Intra- and interannual variability and source apportionment of volatile organic compounds during 2018–2020 in Zhengzhou, central China" by Shijie Yu et al., Atmos. Chem. Phys. Discuss.,
<https://doi.org/10.5194/acp-2021-1016-RC2>, 2022

The study entitled "Measurement report: Intra-, inter-1 annual variability and source apportionment of VOCs during 2018-2020 in Zhengzhou, Central China" by Yu et al. contains some interesting long-term data and falls within the scope of ACP. Unfortunately, the work is currently not of sufficient quality to be published in ACP. The work may become suitable after major revisions.

Major concerns:

- One of my major concerns is that the authors making big statements about relative VOC abundances, the contribution ozone formation potential and the contribution to the total OH reactivity declaring certain compounds to be the "most abundant" VOCs without having measured oxygenated VOCs. Certain oxygenated compounds such as methanol, acetone or formaldehyde can contribute significantly to the total VOC burden, while formaldehyde and acetaldehyde often contribute significantly to the total OH reactivity. Since OVOC mixing ratios in China can be higher than alkane mixing ratios see e.g. Sun et al. 2019 Environ Sci Pollut Res (2019) 26:27769–27782 a lot of statements need to be qualified to state that they apply only to studies which failed to measure OVOCs. Also comparing mixing ratios in ppb introduces a bit of a bias. It may be better to convert to microgram per m³ (the units used to define air quality standards for compounds). The total reactive carbon mass contributed by something heavier(e.g. toluene) is much larger than the mass contributed by ethane even when the toluene mixing ratio is half that of ethane because the molecular mass is so much higher.

The authors have focused their measurements on compounds that come only from

primary emissions and have no secondary photochemical sources. This can be justified for a PMF study, where the focus is source apportionment. The advantage of that approach is that when OVOCs are excluded the PMF will not form several different factors for photochemically formed compounds. Neither will it form factors for air masses with different photochemical age as was seen in some PMF source apportionment studies from China which included OVOCs. The disadvantage of leaving out OVOCs is that it makes no sense at all to talk about ozone formation potential or OH reactivity when the most reactive ozone precursors are missing. These things can only be discussed when the most important contributors, the OVOCs, are included in the study. Hence with the present data it is better to focus the objective around identifying sources and comparing the pie charts of the total VOCs and of individual compounds with different emission inventories. If the authors focus on pollution identification and mitigation and on the question which emission inventories for this region are most accurate then they can write a scientifically sound paper. Their current work is not yet suitable for ACP.

- My second major concern is with the PMF factor identification which is quite unconvincing for some of the factors and not supported by source samples. Pallavi et al. 2019 (Atmos. Chem. Phys., 19, 15467–15482, 2019 <https://doi.org/10.5194/acp-19-15467-2019>) showed that factor profiles can be contrasted with samples collected near a source or directly at the tailpipe to validate factor profiles and justify factor identification. For factors for which this should be easy such as vehicular exhaust and coal/biomass burning source profiles should be recorded and plotted together with factor profiles to avoid blaming the wrong sources.

Among the factors reported, the following factors are likely to be correctly identified: "Biogenic", "LPG/CNG" and "Vehicular exhaust". Based on the source fingerprints published by Hakkim et al. 2021 ATMOSPHERIC ENVIRONMENT: X 11 (2021) 100118 this "vehicular exhaust" fingerprint is plausible for a fleet comprising of a mixture of LPG, CNG and conventional petrol/diesel vehicles. It may be best evaluated against a traffic junction sample since it doesn't carry the signature of any specific tail pipe exhaust.

The following three factors may have problems with the factor resolution/factor identification Factor 1 "Industrial source", Factor 2 "solvent use", and Factor 5 "Coal burning + Biomass burning". For Factor 1 "Industrial source" and Factor 5 "Coal burning + Biomass burning" the g-space plot for these two factors needs to be shown. It is possible that the same source is getting split into two separate sources already in the 6-factor solution. In fact, the current paper contains no figure that argues why a 6-factor solution is the most plausible in the present case. Hence it cannot be ruled out that the dataset can only separate 5 factors without the g-space plot for these two factors which share almost identical conditional probability plots and diurnal profiles in Figure 7. If the authors can make a case for these two being two separate factors with the g-space plot, then the two sources are still most of the time colinear, and it is hence likely that both are two industrial sources that are very close to each other. Factor 5 cannot be a solid fuel combustion source as solid fuel combustion always produces a set of aromatic hydrocarbons (Benzene, toluene, xylenes) and some other compounds e.g. several alkanes and alkenes at the same time. A source with large quantities of few very specific compounds (ethane, benzene, acetylene in this case) is typically industrial in nature. Benzene and acetylene are the only two compounds that have significantly higher emission factors in any type of flaming combustion with very high combustion efficiency while ethane and the small quantities of other hydrocarbons such as propane in the source profile could be unburned fuel leaking. So, it appears to me that this source profile is from a very specific source that is burning almost pure ethane gas. If there is petrochemical industry around then this could be a gas flare. It is very unlikely that the source is combusting coal or solid biofuel. Solid fuels like coal and biomass burning usually emit a larger set of aromatic hydrocarbons including C8

and C9 aromatics, and larger quantities of alkenes such as ethene and propene. If the authors want to make a case that this is coal burning or solid biofuel burning, they would have to collect source samples to show the source profile. However, many authors have published source profiles for a wide range of combustion in different kind of devices in China. E.g. Yan et al. 2016 Atmospheric Environment Volume 143, 261-269 published source profiles from power plants fired with different types of coal and biomass. Alkenes dominate the emission profiles followed by alkanes and aromatics. Wang et al. 2013 Front. Environ. Sci. Eng. 2013, 7(1): 66-76 showed that alkenes and carbonyls dominated emissions when solid fuels such as coal or biomass briquettes used in residential stoves. Yang et al. 2020 published emission factors for coal and oil boilers Journal of Environmental Sciences Volume 92, June 2020, Pages 245-255. Overall, this factor is incompatible with solid or even liquid fuel and must be caused by some kind of gas burner unless it is the same source as the "industrial source" getting split into two different sources to account for different combustion efficiencies in different parts of the industrial process cycle.

The factor identification of Factor 2 Solvent use may be correct but would have to be supported with source profiles. In fact, among all factor profiles in this study, this factor profile is probably the one most compatible with stack sample of coal / biomass fired power plants from Yan et al. 2016 Atmospheric Environment Volume 143, 261-269 or coal/oil fired boilers Yang et al. 2020 Journal of Environmental Sciences Volume 92, June 2020, Pages 245-255. Hence the authors would have to match the source profile against solvent source profiles e.g. from Lui et al. 2008 Atmospheric Environment Volume 42, Issue 25, August 2008, Pages 6247-6260 and the alternative source profiles (coal/biofuel) and would have to decide based on that analysis as well as the conditional probability plots (in which direction from the observational site is Zhengzhou thermal power plant located) in addition to the correlation with other combustion tracers currently used to affirm the identification.

- The language used is often colloquial and inappropriate. The manuscript will have to go through language editing.
- The information describing the measurements in the methods section is insufficient. What does "continuously monitored" in line 114 mean what is the time resolution of the sampling?
- The authors have conveniently outsourced all their PMF methods section to the supplement. This is not appropriate, certain key points must be described in the main text. Also, the relevant information is incomplete. Information on which species were weak and strong is missing. It appears the authors are running the PMF with mixing ratios in ppb instead of converting everything to microgram per m³ as it should be done. Without converting the input files to mass loadings, it will not be possible to compare the pie charts that can be extracted from the PMF with emission inventories which have the VOC emissions in units of weight.

The justification why their VOC dataset can resolve 6 factors is not sufficient. The authors need to show a plot that depicts how the *i* factors identified by the PMF change from the 3 to the 8-factor solution. Then they need to justify their chosen 6 factor solution (why not 5 factors, why not 7 factors) Can the solution be defended based common sense (are all sources that should be expected there and do their source profiles match with the factor profiles produced by the PMF). Is the *g*-space plot of the solution OK or problematic? When two factors are linearly correlated in their *g*-space plot then the PMF cannot resolve that many factors and a solution with less factors must be used.

Overall, the quality of the text is not up to the mark for example the author's state: "In this study, seven-factors are extracted by PMF model based on: (1) principal component analysis of the VOC data; (2) VOC emission inventory of research region based on field investigation; and (3) Q/Q_{exp} ratio for different factor numbers in the

PMF (Fig. S1).” Yet they actually present a 6-factor solution. I am afraid in the current form the manuscript cannot be considered to be up to the quality benchmark of ACP.

There is no description of the uncertainty analysis performed by the authors. Did they run bootstraps? F-peak analysis? Displacement runs?

- I am not quite comfortable with the OH-reactivity analysis given that OVOCs which often contribute close to 80% of the total OH reactivity has not been measured.

Minor comments:

Line 29: Can you be more specific on which species control strategies should focus. Don't leave it to policy makers to define "key species" after reading your paper. Tell them which compounds to focus on.

Line 39: You probably mean to say that pollution control policies are working

Line 50: Colloquial language like "is booming" should be avoided in scientific manuscripts

Line 55: since you only have references from China it may be better to say "in many regions in China". If you are attempting to make a "global" statement then you need to support it with references from different regions around the globe.

Line 56: The statement about the ozone formation potential must be qualified. "Studies which do not report OVOCs usually identify aromatics and alkenes as the largest contributors of ozone formation potential (OFP) (Li et al., 2019b; Yan et al., 2017), while other studies typically identify OVOCs as the largest contributor to OFP (Li et al. 2014, Atmospheric Environment Volume 99, December 2014, Pages 403-410)".

Line 110: Typical urban environment is an acceptable statement. No large industrial sources not really. Most of the VOCs measured have lifetimes of > 1 hr. The authors may calculate their fetch region from which air can reach the measurement site within 2-3 hours at the average wind speed. Since that radius would likely be more than 10 km Zhengzhou thermal power plant and several industrial areas would likely fall within the fetch region.

Line 185 ff: This comparison of total VOC mixing ratios is a bit problematic. The comparison should be split into studies that reported only alkanes, alkenes, terpenes and aromatic compounds and studies that reported other additional compounds such as OVOCs and/or halocarbons. The studies that report a large set of VOC classes usually report higher mixing ratios. It is not OK to compare apples with pears without saying which ones are the apples and which ones are the pears. It may be easier to compare mixing ratios of specific VOCs between sites without introducing accidental bias.

Line 194 ff: This statement is only valid if the studies cited did measure OVOCs if those studies, just like the present one, did not then it is important to qualify that this applies only to studies which failed to measure OVOCs.

Line 198 they do not account for 90% of the VOCs the account for 90% of the compound classes monitored in the present study.

Line 208: Isoprene emissions have also been reported from biomass burning e.g. from smouldering rice straw fires (Kumar et al. 2021 Science of the Total Environment 789 (2021) 148064).

Line 293ff: The authors flag diagnostic emission ratios (e.g. B/T ratios) for different sources and compare them with ambient observations but not with their PMF source fingerprints. Had they compared with their PMF finger prints they would have realized that the "Solvent" factor has a B/T ratio somewhere in between traffic and coal combustion. It makes no sense to compare diagnostic ratios to ambient mixing ratios which represent a source mixture instead of comparing them to individual PMF factor profiles (which should have emission ratios that match source fingerprints) in a paper presenting a PMF analysis.

Line 304ff: Again diagnostic I-pentane to n-pentane ratios should be compared to source fingerprints not to ambient mixing ratios.

Section 3.3.4: CPF have no meaning if not related to sources around the site. Which side of the monitoring site is the local coal fire power plant located? Refinery? Industrial areas? Forests? This type of analysis makes sense with a local map (e.g. Google Earth) in which relevant features are labelled. Not without.

Section 3.5.1. Since OVOCs which contribute most to the total OH reactivity in most sites around the world were not monitored, I would recommend removing this analysis. If retained it must be qualified that several very reactive compounds that contribute more than half of the total OH reactivity at some sites in China where they were measured (Aldehydes) were not included in this current analysis. Again while comparing with the OH reactivity reported in other studies one has to be careful and must group into studies that reported only the same functional groups that were reported in the present study and

studies that include OVOC and/or halogenated compounds.

Section 3.5.2. has the same problem as section 3.5.1. Since OVOCs were not monitored the relative rating of the contribution to the OFP is very biased. I would recommend removing this analysis.