Reviewer #3

Thank you for the opportunity to review this manuscript. The major concern I have on the manuscript is centered on the analysis of the data. These emissions data are multivariate, strictly positive and relative which means they are compositional data and should be analyzed as such. The data can be transformed into log-ratios which places them on the real number line thus enabling application of many familiar statistical tools. References for this approach to the analysis of emissions data are presented in the comments below. Failure to use this approach can result in spurious correlations between the emissions and errors in interpretation of results. Another commonly used technique in emissions analysis and source apportionment is positive matrix factorization (Sekimoto, K., Koss, A. R., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., Warneke, C., Yokelson, R. J., Roberts, J. M., and de Gouw, J.: High- and low-temperature pyrolysis profiles describe volatile organic compound emissions from western US wildfire fuels, Atmospheric Chemistry and Physics, 18, 9263–9281, https://doi.org/10.5194/acp-18-9263-2018, 2018). This multivariate technique does not consider the relative nature of emissions data composition.

Response: We would like to thank the reviewer for the comments and suggestions to improve the current work. We will have the reviewer comments in black, address the comments in blue, and modified sentences in red.

These emissions data are multivariate, strictly positive and relative which means they are compositional data and should be analyzed as such.

Response: We do not believe that compositional data analysis provides the correct basis with which to analyze the data because the data in our work presented here is not strictly positive (i.e., there are zero values), which is a requirement of this data analysis (Greenacre, 2021). The results are not all real, positive, values because there are some species are that observed above our limit of detection in some emissions that are not observed in other types. Consequently, we have ruled out this approach.

Another commonly used technique in emissions analysis and source apportionment is positive matrix factorization (Sekimoto, K., Koss, A. R., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., Warneke, C., Yokelson, R. J., Roberts, J. M., and de Gouw, J.: High- and low-temperature pyrolysis profiles describe volatile organic compound emissions from western US wildfire fuels, Atmospheric Chemistry and Physics, 18, 9263–9281, https://doi.org/10.5194/acp-18-9263-2018, 2018). This multivariate technique does not consider the relative nature of emissions data composition.

Response: As a laboratory we are very familiar with positive matrix factorization (PMF), while PMF could be a useful addition when discussing a time series (similar to the citation shown) (Tong et al., 2021; Qi et al., 2020; Qi et al., 2019; Stefenelli et al., 2019; Wang et al., 2020; Wang et al., 2021; Crippa et al., 2013; Tobler et al.,

2021; Mohr et al., 2012). Though, we disagree in its implementation when using average measured composition across many discreet fuel types.

1. 81-83 This is a 1 sentence paragraph. Either expand the text or include it in the preceding or subsequent paragraph.

Response: We have already merged this single sentence with the preceding paragraph as suggested.

2. 99 Does the VOCUS identify the characteristic compound or does post-sampling analysis by an investigator identify a "characteristic" compound? What is meant by the term "characteristic" compound for a fuel type?

Response: To avoid ambiguity, we have removed all instances of the term "characteristic compounds" from the manuscript. Instead, we now refer to the selected substances as "potential markers" based on the statistical methods used, similar to those employed by Zhang et al. (2023). This change helps in providing a more precise scientific explanation.

3. 111 Are you measuring the solid fuel combustion emissions (which would be from char) or the emissions produced from the combustion of gaseous products produced by the pyrolysis of solid fuels?

Response: The emissions being measured are combination of both: the solid fuel combustion emissions, which would include emissions from the char, and the emissions produced from the combustion of gaseous products generated by the pyrolysis of the solid fuels. The specific contributions of each depend on factors such as the combustion temperature and conditions. Within this study a full cycle of burning commenced, where the VOCUS measured the composition.

For clarification, we have added the following text (Line 124-138):

With those six different fuels, we categorized six burning types for this experiment. 1) beech logs stove, 2) spruce/pine logs stove, 3) spruce/pine branches and needles open, 4) dry straw open, 5) cow dung open and 6) coal stove. We selected these six solid fuels and conducted emissions tests to simulate certain types of burning found in the atmosphere. Among the list above, 1) beech logs stove and 2) spruce/pine logs stove are representative of residential wood burning, which are burned separately in a stove, consistent with the materials used in two previous articles (Bertrand et al., 2017; Bhattu et al., 2019). To represent forest fires or wildfire and agricultural field combustion, 3) a mixture of fresh spruce/pine branches and needles and 4) straw were combusted in an open stainless-steel cylinder (65 cm in diameter and 35 cm in height). Traditional cooking and heating practices in regions like India are represented by 5) cow dung cakes open burning by using half-open stoves (Loebel Roson et al., 2021). Finally, traditional cooking and heating practices in rural regions of developing countries are represented by 6) coal stove burning in a portable cast iron stove purchased from the local market (Liu et al., 2017). Of course, these conditions do not fully accurately represent the conditions found in actual fires, which consistent of a variety of burning species (e.g., trees, underbrush, peat soils, etc…), but represent laboratory burning conditions.

4. 117 Was either proximate or ultimate analysis performed on the fuels? I would expect a significant difference in N in the cow dung compared to the other fuels. If such a difference exists in the unburnt fuel, it would seem that it would translate through the combustion and into the emissions and identification of the characteristic compounds. The fuel composition is also compositional data and should be analyzed accordingly.

Response: We agree that the fuel composition would be of interest, but unfortunately proximate / ultimate analysis was not performed on the fuels and lied outside of the scope of these studies. Proximate / ultimate analysis is not routinely measured in accompanied emission measurements, though it is certainly useful when available. Similar to other studies that do not have this type of analysis, the composition of the fuels should be reflected in the emissions observed from the fuels.

5. 125 What was used to represent agricultural waste? Was agricultural was straw only? Please clarify the difference between the agricultural waste and the fuels used to simulate "forest fires"? Was there a difference between the fuel arrangement or the burning conditions? I recommend that you don't use these fuels to characterize "forest fires" as there is a wide range of fuels which burn in forest and bush fires ranging from peat soils to coniferous and hardwood forest fuels to grasses to various shrub fuels.

Response: We selected these six solid fuels and conducted emissions tests with different combustion methods to simulate certain types of biomass burning found in the atmosphere. Our goal was not to comprehensively characterize any specific type of combustion, such as forest fires. As the reviewer points out, we did not attempt to replicate the diverse fuel types and conditions present in actual forest fires, which indeed can vary significantly. Instead, we focused on the direct emissions from the selected fuels under controlled conditions. Therefore, we chose a representative fuel, e.g., straw, to test a specific type of agricultural waste. In our subsequent analysis, we focused only on individual fuels like straw, rather than analyzing agricultural waste as a broader category. Accordingly, we have revised the sentences and added explanations to clarify this point (Line 124-138):

With those six different fuels, we categorized six burning types for this experiment. 1) beech logs stove, 2) spruce/pine logs stove, 3) spruce/pine branches and needles open, 4) dry straw open, 5) cow dung open and 6) coal stove. We selected these six solid fuels and conducted emissions tests to simulate certain types of burning found in the atmosphere. Among the list above, 1) beech logs stove and 2) spruce/pine logs stove are representative of residential wood burning, which are burned separately in a stove, consistent with the materials used in two previous articles (Bertrand et al., 2017; Bhattu et al., 2019). To represent forest fires or wildfire and agricultural field combustion, 3) a mixture of fresh spruce/pine branches and needles and 4) straw were combusted in an open stainless-steel cylinder (65 cm in diameter and 35 cm in height). Traditional cooking and heating practices in regions like India are represented by 5) cow dung cakes open burning by using half-open stoves (Loebel Roson et al., 2021). Finally, traditional cooking and heating practices in rural regions of developing countries are represented by 6) coal stove burning in a portable cast iron stove purchased from the local market (Liu et al., 2017). Of course, these conditions do not fully accurately represent the conditions found in actual fires, which consistent of a variety of burning species (e.g., trees, underbrush,

peat soils, etc…), but represent laboratory burning conditions.

6. 133-143 The burning of the logs is described. How did this differ from the straw burning? Straw will ignite and burn more quickly than wooden logs. What was the moisture content of the various fuels? Was a constant heating rate used? These pyrolysis and combustion characteristics will affect time to ignition as well as the composition of the emissions.

Response: The method of burning wood (logs) and straw differs, as described in Section 2.1 "Fuel and Burning Types." Straw was combusted in an open stainless-steel cylinder (65 cm in diameter and 35 cm in height), while wood (logs) was burned in a stove. We did measure the moisture content of the woods, where the water content for dried logs was 10-12%, and the water content for the wet (open burning) logs was 30-40%.

We did not use a constant heating rate. Instead, we initiated the burning and then allowed the combustion to proceed according to the properties of the fuels. As expected, we observed that the straw burned faster (fully being consumed within ~3-5 min.) than the logs (burning for ~30-45 min.). In this study, we did not use a heating device to sustain combustion; rather, we aimed to simulate real-world burning conditions, where the fuel burns on its own after ignition. We specifically described the burning of spruce/pine logs in this section because we altered the oxygen content in the stove during the combustion process to explore the changes in emission factors and chemical compositions under different combustion phases (flaming and smoldering) as discussed in Section 3.2. This is particularly relevant since both combustion states are commonly present in household wood burning. Based on your comments, we have revised the sentences and added explanations to clarify this point.

Line 118-119: Six solid fuels were studied (coal briquettes and biomass fuels: beech logs, spruce/pine logs, fresh spruce/pine branches and needles, dry straw, cow dung) with three to six replicate burns.

Line 140-156: The experimental design is shown in Figure S1. In summary, it is made up of a burner and a set of diluters with heated lines. The zero air was provided by a zero air generator (737-250 series, AADCO Instruments, Inc., USA) for cleaning and dilution (Heringa et al., 2011; Bruns et al., 2015). The zero air generator takes ambient air and scrubs particulates and volatile organic compounds from the air leaving a mixture that is largely made up of N_2, O_2, and Ar at ambient concentrations. Other trace gases are scrubbed to lower than atmospheric concentrations including CO_2 (< 80 ppb) and CH_4 (< 40 ppb). Before each burn, a continuous stream of zero air was passed through the gas lines overnight to avoid cross-contamination between burns and to ensure a low background of VOCs. Once a burn is initiated from the various combustibles, emissions are sampled from the chimney through a heated line (473 K). The emissions (both gas and particle phases) are then diluted by two Dekati diluters (DI-1000, Dekati Ltd.) which dilutes the emissions by a factor of ~ 100 (473 K, DI-1000, Dekati Ltd.). Note that beech logs combustion cycles consist of a first cycle referred to as the 'first load' and subsequent cycles, referred to as 'reloads'. The first load consisted of a cold start, flaming, smoldering, and burn-out phase, and the reloads were comprised of a warm start, flaming, smoldering, and burn-out phase. Organic vapor emissions of solid fuel combustion are released within 10-30 min after loading according to the properties of the fuels. We define the time until full ignition duration for burning encompasses 80% of the entire process, starting from loading the fuels to burnout.

Line 357-361: Also, we note a specific difference in the oxygenated aromatic compounds and those with C > 6 for

open wood burning conditions, compared to the stove. This difference may be driven by the difference in the water content of the wood, which is significantly higher for open wood burning (30-40%) compared to stove burning (10-12%). The increase in these oxygenated components comes at the expense of species containing carbonyl and furan functionalities.

7. 180-181 While Andreae and Merlet used the carbon mass balance approach in 2001, it was first used as early as 1969 (Boubel, R. W., Darley, E. F., and Schuck, E. A.: Emissions from burning grass stubble and straw, Journal of the Air Pollution Control Association, 19, 497–500, https://doi.org/10.1080/00022470.1969.10466517, 1969) and was well-established by the mid-1980s (Nelson, R. M., Jr.: An evaluation of the carbon balance technique for estimating emission factors and fuel consumption in forest fire, USDA Forest Service, Southeastern Forest Experiment Station, Asheville, NC, 1982.)

Response: Based on the comments from Meinrat O. Andreae and your suggestions, I have updated the references accordingly.

8. 193 Please make sure that the subscripts for C, O, and N are consistently italicized (or not). Make sure that the subscript for oxygen is O and not zero.

Response: Thank you for pointing that out. We will ensure that the subscripts for C, O, and N are consistently italicized throughout the text, and we will verify that the subscript for oxygen is correctly represented as "O" and not zero.

$$log_{10}{}_C^* = \left(n_C^0 - \mathrm{n}_C^{\mathrm{i}}\right)\mathrm{b}_C - n_O^i b_O - 2\frac{n_C^{\mathrm{i}} n_O^{\mathrm{i}}}{n_C^{\mathrm{i}} + n_O^{\mathrm{i}}} b_{CO} - n_N^{\mathrm{i}} b_N \qquad \text{Equation (3)}$$

9. 202 It has been recently shown that smoke emissions data are multivariate, not independent and are relative values that are dependent on the compounds present in the mixture (Gibergans-Baguena, J., Hervada-Sala, C., and Jarauta-Bragulat, E.: The quality of urban air in Barcelona: a new approach applying compositional data analysis methods, Emerg Sci J, 4, 113–121, https://doi.org/10.28991/esj-2020-01215, 2020; Jarauta-Bragulat, E., Hervada-Sala, C., and Egozcue, J. J.: Air Quality Index revisited from a compositional point of view, Math Geosci, 48, 581–593, https://doi.org/10.1007/s11004-015-9599-5, 2016; Weise, D. R., Palarea-Albaladejo, J., Johnson, T. J., and Jung, H.: Analyzing wildland fire smoke emissions data using compositional data techniques, J. Geophys. Res. Atmos., 125, e2019JD032128, https://doi.org/10.1029/2019JD032128, 2020). These characteristics of the data apply whether the emissions data are expressed as emission factors, emission ratios, mole ratios or mass ratios (van den Boogaart, K. G. and Tolosana-Delgado, R.: Analyzing compositional data with R, Springer, Heidelberg, 258 pp., 2013.).

It has also been shown that MCE as an index describing the completeness of combustion is not independent of

the quantities of other emissions and that it should not be used as a predictor for the other gases in the composition.

Response: This study presents the emissions from many different fuel sources and demonstrates that the composition of the gaseous emissions is indeed dependent upon the compounds present within the fuel mixture itself (assuming the composition of the emissions are similar to those in the combustion source). We demonstrate that changing the MCE impacts the emission factors and the concentrations of the emitted gases within the measurement itself and specifically changes the composition of the emissions as well. This work also demonstrates there is a specific change in the chemical composition of the emissions when the MCE changes.

10. How is the Mann-Whitney test and other techniques used in this manuscript affected by the statistical characteristics of your data? The Mann-Whitney test is a univariate test. You should consider using the generalized multivariate version if it has been applied to compositional data (https://doi.org/10.1016/j.jmva.2022.104946). As compositional data analysis has been used more extensively in Europe, recommend reaching out to the statisticians listed in the various publications above. You should also consider a global test (instead of pairwise comparisons) that controls the experiment-wise probability of committing a Type 1 error (such as false discovery rate-Benjamini, Y. and Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing, Journal of the Royal Statistical Society. Series B (Methodological), 57, 289–300, 1995).

Response: Although it is possible to use a multi-variate approach, we have chosen to use Mann-Whitney to identify outliers (potential markers) to be consistent with previous work and align with identification of markers used in various other mass spectrometric approaches in analytical chemistry (White et al., 2019; Chen et al., 2012; Teunissen et al., 2011; Chmaj-Wierzchowska et al., 2015; Nomura et al., 2004; Jasperse et al., 2007; Nagai et al., 2020; Sun et al., 2019; Tritten et al., 2013). Further the data that is being used is not always positive (there are 0 values), which means the use of compositional data analysis is not a valid approach for our data (Greenacre, 2021). We have chosen the univariate approach to identify molecular formula that are specific statistical outliers relative to the other emissions, which differs from other studies that have focused on emission factors from specific sources. Only species that are confirmed outliers between each and every pair-wise comparison is chosen, we believe this is a cautious approach that lowers the probability of committing a Type 1 error, which is already low in the case for a single pair-wise comparison. If the probability of committing a Type 1 error is 5% and 4 groups are used, then the probability of committing a Type 1 error across 4 different comparisons is 0.00062%.

11. 232 What does 0.99 +/- 0.02 mean? Is this arithmetic mean and standard error or standard deviation? Since MCE is a proportion that can not exceed 1, the correct formula for the confidence interval of this proportion should not exceed 1. The geometric mean is the appropriate measure of central tendency for relative (proportional

data). EFs are expressed as gm pollutant/gm fuel burned which is a rate (and a relative value so a geometric mean should be used as in Butler, B. M., Palarea-Albaladejo, J., Shepherd, K. D., Nyambura, K. M., Towett, E. K., Sila, A. M., and Hillier, S.: Mineral–nutrient relationships in African soils assessed using cluster analysis of X-ray powder diffraction patterns and compositional methods, Geoderma, 375, 114474, https://doi.org/10.1016/j.geoderma.2020.114474, 2020.

Response: We apologize for the error in reporting 0.99 ± 0.02. Based on the data from Table S1, the standard deviation is actually less than 0.001, so we have corrected this and omitted the unnecessary margin. In this study, the emission factors (EFs) are calculated as arithmetic mean ± standard deviation. We chose this method to maintain consistency with previous studies that calculated VOC, CO and CO_2 emission factors using the same approach. This allows for direct comparison with their results (Andreae, 2019; Janhäll et al., 2010; Koss et al., 2018).

12. 257 Correlation is not an appropriate measure for compositional data as the value of the correlation coefficient is dependent upon the other compounds in the composition. Dropping a gas from the composition changes the pairwise correlations (Aitchison, J.: A concise guide to compositional data analysis, 2003. Available at http://ima.udg.edu/activitats/codawork03/; Weise, D. R., Fletcher, T. H., Safdari, M.-S., Amini, E., and Palarea-Albaladejo, J.: Application of compositional data analysis to determine the effects of heating mode, moisture status and plant species on pyrolysates, Int. J. Wildland Fire, 31, 24–45, https://doi.org/10.1071/WF20126, 2022). Proportionality has been suggested as an appropriate measure of the association between two components of a composition (Lovell, D., Pawlowsky-Glahn, V., Egozcue, J. J., Marguerat, S., and Bähler, J.: Proportionality: a valid alternative to correlation for relative data, PLoS Comput Biol, 11, e1004075, https://doi.org/10.1371/journal.pcbi.1004075, 2015). Drawing conclusions based on the measure of association between two variables without determining the significance of the measure by a statistical test of some sort is not recommended.

Response: We agree with the reviewer that a simple correlation matrix is likely not the most appropriate measure for compositional data if we were trying to demonstrate anything quantitative. The reviewer should also consider the point of using such a simplistic approach as a tool to discuss the results. The use of a correlation matrix as shown in Figure 2 highlights the similarity between all biomass combustion sources investigated, and that there is a specific difference for coal. The correlation matrix is not used in absolute terms regarding the emissions but is used in qualitative terms as a discussion tool in order to discuss the reproducibility from burn to burn and the similarity (or differences) between the emissions which are shown in detail in Figures 4 and 5, where a specific difference for coal comes from the emissions of non-oxygenated aromatic compounds. It would be just as feasible to put 6 mass spectra representing the different emission types.

Note the use of the correlation matrix on lines 316 - 324:

To assess the feasibility of distinguishing differences between combustion solid fuel types based on the measured species, we evaluated the similarity of the mass spectra obtained from each experiment using the correlation coefficient (*r*), as shown in Figure 2. Organic vapors from the same burning fuel are strongly correlated (0.82-0.99), indicating the general repeatability of the experiments. Furthermore, we observed a weak intra-fuel correlation between coal and other biomass sources (0.44-0.78), suggesting significant differences in chemical composition. By contrast, the separation between different solid fuel type is not stark and all possess a correlation between 0.6-0.98. Overall, the correlation coefficient highlights similarities between all biomass-based emissions, which will now be discussed in detail.

13. 290 In compositional data analysis, the effect of fuel type on the log-ratios between different groups of compounds can be tested in an analysis of variance context to determine differences. These log-ratios are known as balances (Egozcue, J. J. and Pawlowsky-Glahn, V.: Groups of parts and their balances in compositional data analysis, Mathematical Geology, 37, 795–828, https://doi.org/10.1007/s11004-005-7381-9, 2005; Weise, D. R., Fletcher, T. H., Safdari, M.-S., Amini, E., and Palarea-Albaladejo, J.: Application of compositional data analysis to determine the effects of heating mode, moisture status and plant species on pyrolysates, Int. J. Wildland Fire, 31, 24–45, https://doi.org/10.1071/WF20126, 2022).

Response: As stated above, we do not believe that compositional data analysis provides the correct basis with which to analyze the data because the data in our work presented here is not strictly positive (i.e., there are zero values), which is a requirement of this data analysis (Greenacre, 2021). The results are not all real, positive, values because there are some species are that observed above our limit of detection in some emissions that are not observed in other types. Consequently, we have ruled out this approach.

14. 308 see comment below for figures S4, S5 regarding error bars.

Response: We have double-checked the data and corrected that the error bar represents 1 standard deviation, not ½. We have corrected them in the SI.

15. 348 You are discussing differences in relative terms which is appropriate. The statistics used to describe and test hypotheses should also recognize the relative nature of the data.

Response: I agree that addressing the relative nature of the data is essential, especially when making comparisons. I'll ensure that the statistical methods used align with this approach and clearly reflect the relative differences within the data. We have added the standard deviation in the manuscript (Line 294-296).

On average, EFs for organic vapors in the flaming stage are approximately four times lower (31.4 ± 7.1 g kg^{-1}) than those in the smoldering stage fires (121.9 ± 24 g kg^{-1}).

16. 370 See the earlier comment regarding the Mann-Whitney test, the multivariate nature of the data and the probability of committing a Type 1 error.

Response: We have chosen the univariate approach to identify molecular formula that are specific statistical outliers relative to the other emissions, which differs from other studies that have focused on emission factors from specific sources. Only species that are confirmed outliers between each and every pair-wise comparison is chosen, we believe this is a cautious approach that lowers the probability of committing a Type 1 error, which is already low in the case for a single pair-wise comparison. If the probability of committing a Type 1 error is 5% and 4 groups are used, then the probability of committing a Type 1 error across 4 different comparisons is 0.00062%.

17. 398 Which supplementary table contains the characteristics compounds?

Response: We realized that we forgot to upload the Excel file. The file has now been uploaded.

18. 423 Where is the chemical composition of unburnt cow dung presented?

Table 1 Are the values arithmetic mean +/- standard deviation? Please provide more information on values. Should use geometric mean and present a confidence interval (or the standard error of the mean). Consider including complete fuel composition (CHNSO). Also, proximate analysis because cows have ability to digest cellulose which make affect the burning characteristics or the relative amounts of cellulose, hemicellulose and lignin in the fuels which will affect both the pyrolysis and combustion processes as well as emissions production.

Response: As mentioned above, unburnt chemical analysis of the cow dung was not performed.

19. Figure 2-5 Recommend making the axes and other information larger fonts (relative to titles).

Response: Thank you for your suggestion. We have updated figures to increase the font size of the axes and other information relative to the titles to improve readability.

20. Figure 5 What do the different sized circles labeled 0.1, 0.5, 1 and 2 indicate?

Response: In the caption, we mentioned that the markers are scaled according to the square root of the fractional contribution (%). Therefore, the sizes of the circles labeled 0.1, 0.5, 1, and 2 represent the square roots of the corresponding fractional contributions.

21. Figure S4, S5 Why is ½ of 1 standard deviation used as an error bar? Is this is based on the assumption made for normally distributed data that 1 standard deviation captures about 68 percent of the data and 2 standard deviations capture about 95% of the data? This is based on the population and not the sample. A confidence

interval should be calculated. Also, these data are not normally-distributed. They are proportions which are constrained between 0 and 1 (or 0 and 100).

Response: We have double-checked the data and corrected that the error bar represents 1 standard deviation, not ½. While the data from a single experiment may not be normally distributed, the average data and standard deviation were calculated from multiple repeated experiments. The error bars reflect the standard deviation of the relative contributions across these repeated experiments.

22. Figure S13 Caption needs to be fixed.

Response: We have revised the caption.

Reference:

Andreae, M. O.: Emission of trace gases and aerosols from biomass burning–an updated assessment, Atmos. Chem. Phys., 19, 8523-8546, 2019.

Bertrand, A., Stefenelli, G., Bruns, E. A., Pieber, S. M., Temime-Roussel, B., Slowik, J. G., Prévôt, A. S. H., Wortham, H., El Haddad, I., and Marchand, N.: Primary emissions and secondary aerosol production potential from woodstoves for residential heating: Influence of the stove technology and combustion efficiency, Atmos. Environ., 169, 65-79, 10.1016/j.atmosenv.2017.09.005, 2017.

Bhattu, D., Zotter, P., Zhou, J., Stefenelli, G., Klein, F., Bertrand, A., Temime-Roussel, B., Marchand, N., Slowik, J. G., Baltensperger, U., Prevot, A. S. H., Nussbaumer, T., El Haddad, I., and Dommen, J.: Effect of Stove Technology and Combustion Conditions on Gas and Particulate Emissions from Residential Biomass Combustion, Environ. Sci. Technol., 53, 2209-2219, 10.1021/acs.est.8b05020, 2019.

Bruns, E. A., El Haddad, I., Keller, A., Klein, F., Kumar, N. K., Pieber, S. M., Corbin, J. C., Slowik, J. G., Brune, W. H., Baltensperger, U., and Prévôt, A. S. H.: Inter-comparison of laboratory smog chamber and flow reactor systems on organic aerosol yield and composition, Atmos. Meas. Tech., 8, 2315-2332, 10.5194/amt-8-2315-2015, 2015.

Chen, Y.-T., Chen, H.-W., Domanski, D., Smith, D. S., Liang, K.-H., Wu, C.-C., Chen, C.-L., Chung, T., Chen, M.-C., and Chang, Y.-S.: Multiplexed quantification of 63 proteins in human urine by multiple reaction monitoring-based mass spectrometry for discovery of potential bladder cancer biomarkers, Journal of proteomics, 75, 3529-3545, 2012.

Chmaj-Wierzchowska, K., Kampioni, M., Wilczak, M., Sajdak, S., and Opala, T.: Novel markers in the diagnostics of endometriomas: Urocortin, ghrelin, and leptin or leukocytes, fibrinogen, and CA-125?, Taiwanese Journal of Obstetrics and Gynecology, 54, 126-130, 2015.

Crippa, M., El Haddad, I., Slowik, J. G., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Sciare,

J., and Baltensperger, U.: Identification of marine and continental aerosol sources in Paris using high resolution aerosol mass spectrometry, Journal of Geophysical Research: Atmospheres, 118, 1950-1963, 2013.

Greenacre, M.: Compositional data analysis, Annual Review of Statistics and its Application, 8, 271-299, 2021.

Heringa, M. F., DeCarlo, P. F., Chirico, R., Tritscher, T., Dommen, J., Weingartner, E., Richter, R., Wehrle, G., Prévôt, A. S. H., and Baltensperger, U.: Investigations of primary and secondary particulate matter of different wood combustion appliances with a high-resolution time-of-flight aerosol mass spectrometer, Atmos. Chem. Phys., 11, 5945-5957, 10.5194/acp-11-5945-2011, 2011.

Janhäll, S., Andreae, M. O., and Pöschl, U.: Biomass burning aerosol emissions from vegetation fires: particle number and mass emission factors and size distributions, Atmos. Chem. Phys., 10, 1427-1439, 2010.

Jasperse, B., Jakobs, C., Eikelenboom, M. J., Dijkstra, C. D., Uitdehaag, B. M., Barkhof, F., Polman, C. H., and Teunissen, C. E.: N-acetylaspartic acid in cerebrospinal fluid of multiple sclerosis patients determined by gas-chromatography-mass spectrometry, Journal of neurology, 254, 631-637, 2007.

Koss, A. R., Sekimoto, K., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., and Jimenez, J. L.: Non-methane organic gas emissions from biomass burning: identification, quantification, and emission factors from PTR-ToF during the FIREX 2016 laboratory experiment, Atmos. Chem. Phys., 18, 3299-3319, 2018.

Liu, C., Zhang, C., Mu, Y., Liu, J., and Zhang, Y.: Emission of volatile organic compounds from domestic coal stove with the actual alternation of flaming and smoldering combustion processes, Environ. Pollut., 221, 385-391, 10.1016/j.envpol.2016.11.089, 2017.

Loebel Roson, M., Duruisseau-Kuntz, R., Wang, M., Klimchuk, K., Abel, R. J., Harynuk, J. J., and Zhao, R.: Chemical Characterization of Emissions Arising from Solid Fuel Combustion—Contrasting Wood and Cow Dung Burning, ACS Earth and Space Chemistry, 5, 2925-2937, 10.1021/acsearthspacechem.1c00268, 2021.

Mohr, C., DeCarlo, P., Heringa, M., Chirico, R., Slowik, J., Richter, R., Reche, C., Alastuey, A., Querol, X., and Seco, R.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, Atmos. Chem. Phys., 12, 1649-1665, 2012.

Nagai, K., Uranbileg, B., Chen, Z., Fujioka, A., Yamazaki, T., Matsumoto, Y., Tsukamoto, H., Ikeda, H., Yatomi, Y., and Chiba, H.: Identification of novel biomarkers of hepatocellular carcinoma by high-definition mass spectrometry: Ultrahigh-performance liquid chromatography quadrupole time-of-flight mass spectrometry and desorption electrospray ionization mass spectrometry imaging, Rapid Commun. Mass Spectrom., 34, e8551, 2020.

Nomura, F., Tomonaga, T., Sogawa, K., Ohashi, T., Nezu, M., Sunaga, M., Kondo, N., Iyo, M., Shimada, H., and Ochiai, T.: Identification of novel and downregulated biomarkers for alcoholism by surface enhanced laser desorption/ionization-mass spectrometry, Proteomics, 4, 1187-1194, 2004.

Qi, L., Chen, M., Stefenelli, G., Pospisilova, V., Tong, Y., Bertrand, A., Hueglin, C., Ge, X., Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G.: Organic aerosol source apportionment in Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) – Part 2: Biomass burning influences in winter, Atmos. Chem. Phys., 19, 8037-8062, 10.5194/acp-19-8037-2019, 2019.

Qi, L., Vogel, A. L., Esmaeilirad, S., Cao, L., Zheng, J., Jaffrezo, J.-L., Fermo, P., Kasper-Giebl, A., Daellenbach, K. R., Chen, M., Ge, X., Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G.: A 1-year characterization of organic aerosol composition and sources using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF), Atmos. Chem. Phys., 20, 7875-7893, 10.5194/acp-20-7875-2020, 2020.

Stefenelli, G., Pospisilova, V., Lopez-Hilfiker, F. D., Daellenbach, K. R., Hüglin, C., Tong, Y., Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G.: Organic aerosol source apportionment in Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) – Part 1: Biogenic influences and day–night chemistry in summer, Atmos. Chem. Phys., 19, 14825-14848, 10.5194/acp-19-14825-2019, 2019.

Sun, Y., Chen, Y., Sun, C., Liu, H., Wang, Y., and Jiang, X.: Analysis of volatile organic compounds from patients and cell lines for the validation of lung cancer biomarkers by proton-transfer-reaction mass spectrometry, Analytical Methods, 11, 3188-3197, 2019.

Teunissen, C., Koel-Simmelink, M., Pham, T., Knol, J., Khalil, M., Trentini, A., Killestein, J., Nielsen, J., Vrenken, H., and Popescu, V.: Identification of biomarkers for diagnosis and progression of MS by MALDI-TOF mass spectrometry, Multiple Sclerosis Journal, 17, 838-850, 2011.

Tobler, A. K., Skiba, A., Canonaco, F., Močnik, G., Rai, P., Chen, G., Bartyzel, J., Zimnoch, M., Styszko, K., and Nęcki, J.: Characterization of non-refractory (NR) PM 1 and source apportionment of organic aerosol in Kraków, Poland, Atmos. Chem. Phys., 21, 14893-14906, 2021.

Tong, Y., Pospisilova, V., Qi, L., Duan, J., Gu, Y., Kumar, V., Rai, P., Stefenelli, G., Wang, L., Wang, Y., Zhong, H., Baltensperger, U., Cao, J., Huang, R.-J., Prévôt, A. S. H., and Slowik, J. G.: Quantification of solid fuel combustion and aqueous chemistry contributions to secondary organic aerosol during wintertime haze events in Beijing, Atmos. Chem. Phys., 21, 9859-9886, 10.5194/acp-21-9859-2021, 2021.

Tritten, L., Keiser, J., Godejohann, M., Utzinger, J., Vargas, M., Beckonert, O., Holmes, E., and Saric, J.: Metabolic profiling framework for discovery of candidate diagnostic markers of malaria, Scientific reports, 3, 2769, 2013.

Wang, L., Slowik, J. G., Tong, Y., Duan, J., Gu, Y., Rai, P., Qi, L., Stefenelli, G., Baltensperger, U., Huang, R.-J., Cao, J., and Prévôt, A. S. H.: Characteristics of wintertime VOCs in urban Beijing: Composition and source apportionment, Atmos. Environ. , 9, 100100, 10.1016/j.aeaoa.2020.100100, 2021.

Wang, L., Slowik, J. G., Tripathi, N., Bhattu, D., Rai, P., Kumar, V., Vats, P., Satish, R., Baltensperger, U., Ganguly, D., Rastogi, N., Sahu, L. K., Tripathi, S. N., and Prévôt, A. S. H.: Source characterization of volatile organic compounds measured by proton-transfer-reaction time-of-flight mass spectrometers in Delhi, India, Atmos.

Chem. Phys., 20, 9753-9770, 10.5194/acp-20-9753-2020, 2020.

White, R., Pulford, E., Elliot, D. J., Thurgood, L. A., and Klebe, S.: Quantitative mass spectrometry to identify protein markers for diagnosis of malignant pleural mesothelioma, Journal of proteomics, 192, 374-382, 2019.

Zhang, J., Li, K., Wang, T., Gammelsæter, E., Cheung, R. K., Surdu, M., Bogler, S., Bhattu, D., Wang, D. S., and Cui, T.: Bulk and molecular-level composition of primary organic aerosol from wood, straw, cow dung, and plastic burning, Atmos. Chem. Phys., 23, 14561-14576, 2023.