

## Response to Referee number 2

25 August 2023

The authors would like to thank Referee no. 2 very much for his/her expert, detailed and valuable comments to further improve and clarify the MS. We have considered all recommendations and made the appropriate alterations. Our specific responses are as follows, while the textual modifications amended to the MS can be traced in its marked-up version, which is available online.

The OP of three regions in and around Budapest; urban, suburban and rural, over all four seasons is investigated in this paper. The DTT and AA assays are used to measure OP, and source apportionment and linear regression employed to determine the sources affecting these assays at the different locations and for different seasons. The method has been widely applied in other European locations by some of the co-authors, so although the methods are not new, the results are since the location is novel. Overall, the results are interesting, if OP can be/is linked to health endpoints, since they increase the knowledge of factors affecting PM<sub>2.5</sub> OP, which are consistently showing the importance of biomass burning and non-tailpipe vehicle emissions. The analysis is very detailed and for the most part the paper is well organized and clearly written. This paper covers an important topic and is suitable for publication in ACP with the following revisions to consider.

Last line of Abstract, does this imply use of OP or just sources in developing regulations?

Response 1: Utilisation of OP for expressing the health endpoints caused by PM would/could be the final goal from several aspects. For the moment, regulations of some source types seem more feasible. The sentence was modified to express these main options.

Line 40, how are ambient particles biologically complex? Is this referring to point 8) in the following lines? Please clarify.

Response 2: By referring to biological complexity of aerosol particles, we meant the presence of active substances of biological origin such as bacteria, viruses, pollens and moulds. This is later expressed in point 8 (section 1 Introduction and objectives).

Lines 66 to 68, this is an incomplete list of current studies linking OP to health effects. Eg, the Weichenthal group has published a number of papers on this topic in the last few years. There may be more, but that could be a place to start an expanded literature search. A better literature review is needed here since this is a critical argument on why these assays are useful, ie, that there is a link to health effects.

Response 3: The MS already contains many citations. We further extended the list of references as requested and emphasized the existing evidence between the OP<sub>AA</sub> and OP<sub>DTT</sub> on the one side and the concrete health effects on the other side.

Lines 128 to 130, how was water on the filters controlled or not controlled as part of the mass measurement (ie, equilibrium reached as some low RH)?

Response 4: The exposed and blank filters were pre-equilibrated before weighing them twice in each case at a room  $T$  of 19–21°C and RH of 45 %–50 % for at least 48 h. The details of the usual analytical methods were described in a previous article, which reference was added now.

Line 139, if metal ions, such as Cu(II), Fe(II), Mn(II) play a role in OP (see line 92), why were total metals analyzed. How does this influence the interpretation of the data, ie, how can the metals data be linked to OP? One might consider the solubility of the metals. Low soluble metals, such as iron may have little relation to the total iron concentration, whereas for copper with higher solubility, the issue may be different.

Response 5: The transition and heavy metals were determined by PIXE analysis, which gives their total concentration. Their dissolution actually happens in human respiratory tract lining fluid, and this is expected to be larger than in water. Furthermore, an acid mediated dissolution of transition metals (Fang et al., Highly acidic ambient particles, soluble metals, and oxidative potential: a link between sulfate and aerosol toxicity, *Environ. Sci. Technol.*, 51, 2611–2620, 2017) may further increase the solubility. The lung bioaccessibility was assessed in the present work by the total amounts of the chemical species as the first approximation. A short sentence was added to specify this.

Line 158, are insoluble species included in the analysis? Specifically, what fraction of insoluble species are expected to be included in the measured OP. Does soot/EC, which may have surface active species, contribute? More quantitative details are needed here.

Response 6: The sample extracts were not filtered. All insoluble chemical species including those with active surface area were involved in the analysis. It is a good, but rather challenging idea to estimate the fraction of insoluble species in the suspension. A brief comment and a reference were included in the text on this.

Line 331, how are the intercepts on the DTT vs AA in both plots of Fig 2 explained?

Response 7: The intercepts of the regression lines in Fig. 2a and 2b can be related to and can indicate the minimal response of the different methods to the sample types or some specific components. Their interpretation, particularly in Fig. 2a, is not straightforward and definitely needs additional investigations. We would like to avoid drawing specific conclusions here.

Lines 335 to 343, the paragraph is not very clear. Why are the assays coherent based on both the mass and volume normalized data? Also, I assume the statement that AA is more sensitive than DTT is because the slope of DTT vs AA is less than 1 (might state this explicitly). But what if DTT is just more sensitive to a more chemical species? Reference to sources here (SOC, BB) seems out of place, it has not been discussed yet.

Response 8: The statement on the coherency was confined only to the OP normalised to sampled volume (line 335 of the original MS). By this, we intended to express the significant correlation between the AA and DTT methods. The section was reformulated to be more specific and to increase its clarity. The discussions related to BB and SOC sources were replaced to section 3.4 Oxidative potential and aerosol sources.

Source apportionment does not show any secondary biogenic SOA? Also, can PMF be performed here for a given location and season given the limited number of samples?

Response 9: Separate biogenic SOA factor did not appear in the source apportionment since the available SOC concentrations could not be unfortunately included in the calculations due to their smaller count and larger (up to 30 % – 50 %) relative uncertainty. The PMF modelling for a given location and separate seasons was not attempted because of insufficient number of samples for these specific cases. Nevertheless, we utilised the SOA and biogenic OC as supporting data mainly in the regression analyses. These two quantities were derived earlier by an EC tracer method and a coupled radiocarbon-levoglucosan marker method, respectively from the same samples (Salma et al., Fossil fuel combustion, biomass burning and biogenic sources of fine carbonaceous aerosol in the Carpathian Basin, *Atmos. Chem. Physics*, 20, 4295–4312, 2020; Salma et al.: Secondary organic carbon and its contributions in different atmospheric environments of a continental region and seasons, *Atmos. Res.*, 278, 106360, 2022). Additional PMF test runs to the base case were discussed in sections 2.4 Mathematical models and 3.6 Limitations and later possibilities, which were briefly extended.

Line 424 and throughout. Is intersect the proper term for the regression intercept? Maybe either is acceptable?

Response 10: The term intersect was replaced by intercept in all cases.

Discussion of slopes of DTT or AA vs PM<sub>2.5</sub> mass concentration. Isn't the slope in the plots of Figure 4 equal to DTT<sub>m</sub> and AA<sub>m</sub>? This could be explicitly discussed and aid in the interpretation. Also, contrast this DTT<sub>m</sub> and AA<sub>m</sub> to the data in Table 1 which give a different trend in AA<sub>m</sub> (I believe), likely due to the intercept not affecting AA<sub>m</sub> from the slope method. This should be discussed.

Response 11: Figure 4 shows the OP obtained by the AA and DTT assays and normalised to sampled air volume ( $OP_{AA,V}$  and  $OP_{DTT,V}$ ) and not to the PM mass. The former option or atmospheric quantity is considered to have a closer relationship to human exposure.

Moreover, the correlation of PM-mass normalised OP data with PM mass would raise further questions. The slopes of the regression lines were explicitly given in lines 489–496 of the original MS. They were further discussed now.

Does it make sense that the “toxicity” based on AA (ie, the slope in Fig 4b) is the least for the city center and higher at the rural site? The term toxicity is used loosely here since these acellular assays really do not provide insight on actual toxicity. One might consider this when using the term toxicity throughout the manuscript. The AAm spatial trend is opposite other studies from Europe and most other regions; cities have higher mass normalized OP than rural areas. Are there studies that show the opposite, as found for AAm in this study? Do the authors think this is actually true from a health perspective; ie, the toxicity of rural aerosol, which is largely due to the influence of long-range transported dust, is more toxic than urban aerosols?

Response 12: The term toxicity was replaced by OP or larger sensitivity of OP determined by a particular assay in many relevant cases in the MS. The OP values normalised to sampled volume tend to increase from the rural area to the urban sites as indicated in Fig. 1, which is in line with other European studies. The situation could be complicated with the different contributions of BB and road vehicle sources at the different sites and seasons. See also response 11.

On both scatter plots, why not make the symbols, regression line and the  $r^2$  colors match? This would make it easier to interpret.

Response 13: The experimental data points have lighter shade of a colour, whereas the corresponding regression lines and coefficients of correlation have the darker shade of the same colour to separate the measured data (symbols) from the modelled quantities (lines and numbers).

Line 562, could one find a better word than pronouncedly?

Response 14: The word was deleted.

Finally, two different acellular assays were used. They respond differently to chemical species in the aerosol particles, and so respond differently to sources, which results in differing interpretations of the PM<sub>2.5</sub> air quality in the various regions. It is concluded that more than one assay is needed to get a full picture of air quality. But more could be added. Are these two assays optimal for that goal, or can the authors suggest other pairs based on published findings. If not, what test should be done to select optimal assays? It seems like more interpretation could be gleaned from this study.

Response 15: More precisely, we suggested that multiple, at least two OP assays are to be deployed to get a more holistic picture (line 341 of the original MS). Other assay combinations are also possible; however, most researchers select the AA and DTT assays if only two methods are utilised. A reasonable compromise among the number of the assays,

time or labour demands of the experiments and expected results should be reached. Our next research project is to deal with these issues in a systematic manner, and the preliminary results indicate that further firm and conclusive interpretation are expected in this field. A brief extension was amended on this in section 3.6 Limitations and later possibilities.

Imre Salma  
corresponding author